The Status of Coronals in Standard American English

An Optimality-Theoretic Account

Dissertation zur Erlangung des Doktorgrades

der Philosophischen Fakultät der Universität zu Köln

vorgelegt von

Sybil Scholz

geb. in

Düsseldorf

Erstgutachter: Prof. Dr. Jon Erickson
Zweitgutachterin: Prof. Dr. Susan Olsen

Köln 2003

For

Aleksandra Davidović

CONTENTS

ACKNOWLEDGEMENTS

I would like to thank Samantha Hume, Wolf-Dietrich Bald, Ruth Baumert, Ute Breidenbach, Jon Erickson, Melinda Möller, Susan Olsen, Martha Wassen, and Bella Palmera for their various contributions, their help and emotional support.

PREFACE

List of abbreviations, list of constraints, typographical conventions, and phonetic symbols

ABBREVIATIONS

AAVE	African American Vernacular English
AP	Articulatory Phonology
AT	articulator theory
ATR	advanced tongue root
CON	universal set of CONSTRAINTS
DP	Declarative Phonology
EPG	electropalatography
EVAL	EVALUATOR
GA	General American
GEN	GENERATOR
GP	Generative Phonology
GT	Generative Theory
IPA	International Phonetic Alphabet
MP-rule	morphophonemic rule
MSC	morpheme structure condition
MT	Markedness Theory
NGP	Natural Generative Phonology
NP	Natural Phonology
OCP	Obligatory Contour Principle

OT	Optimality Theory
P-rule	phonological rule
PPT	Principles & Parameters Theory
PT	place of articulation theory
PW	phonological word
ROA	Rutgers Optimality Archive
RP	Received Pronunciation
RTR	retracted tongue root
RU	Radical Underspecification
SAE	Standard American English
SBCSAE	*Santa Barbara Corpus of Spoken American English*
SBE	Standard British English
SLH	Strict Layer Hypothesis
SPC	surface phonetic constraint
SPE	*Sound Pattern of English* (Chomsky & Halle 1968)
SR	surface representation
SSE	Standard Scottish English
SSP	Sonority Sequencing Principle
TCRS	Theory of Constraints and Repair Strategies
UG	Universal Grammar
UR	underlying representation
UT	underspecification theory

CONSTRAINTS (25) AND THEIR DESCRIPTION

ASSIM/LAR	assimilate laryngeal features
ASSIM/PLACE	assimilate place features
DELETE/COR	delete coronal segment
DEP–IO	output is dependent on input
DEP-IO/ONSET	dependency with respect to segments in onsets
DEP–IO/POA	dependency with respect to place of articulation

EPENTHESIZE	epenthesize segment
FAITHFULNESS	faithfulness to input
*GESTURE/COR	no coronal gesture
*GESTURE/DOR	no dorsal gesture
*GESTURE/LAB	no labial gesture
IDENT-IO	identity between input and output correspondents
IDENTIO/FONSET	identity with respect to features (F) of segments in onsets
IDENTIO/FSÝL	identity with respect to Fs of segments in stressed syllables
IDENTIO/LAR	identity with respect to laryngeal features
IDENTIO/MOA	identity with respect to manner of articulation
IDENTIO/POA	identity with respect to place of articulation
MARKEDNESS	no marked structures
MAX–IO	input is maximally present in output
MAX–IO/ONSET	maximality with respect to segments in onsets
MAX–IO/POA	maximality with respect to place of articulation
OBS/VOI	an obstruent must be voiceless
OCP(F)	no adjacent identical elements on a given tier
SONORITY	onsets must increase and codas must decrease in sonority
SON/VOI	a sonorant must be voiced

CONVENTIONS AND SYMBOLS

The phonetic symbols I will use for the transcription of consonants are the common ones taken from the International Phonetic Alphabet (IPA) with affricates written as digraphs (International Phonetic Association 1999). For the vowels I chose the symbols following Wells (1982: xviii).

C	constraint (C_1, C_2 = constraint one, constraint two)
C	consonant (C_1, C_2 = consonant one, consonant two)
V	vowel
G	glide

<table>
<tr><td>L</td><td>liquid</td></tr>
<tr><td>N</td><td>nasal</td></tr>
<tr><td>F</td><td>fricative</td></tr>
<tr><td>A</td><td>affricate</td></tr>
<tr><td>P</td><td>plosive</td></tr>
<tr><td>ʔ</td><td>glottal stop</td></tr>
<tr><td>#</td><td>word boundary</td></tr>
<tr><td>+</td><td>morpheme boundary</td></tr>
<tr><td>| |</td><td>perceptual specification</td></tr>
<tr><td>[ ]</td><td>articulatory specifications or phonetic material</td></tr>
<tr><td>/ /</td><td>perceptual input or phonemic material</td></tr>
<tr><td>< ></td><td>orthographical material</td></tr>
<tr><td>[±coronal]</td><td>SPE-type distinctive feature (part of feature bundle)</td></tr>
<tr><td>CORONAL</td><td>non-linear class node (part of feature geometry)</td></tr>
<tr><td>[±anterior]</td><td>non-linear terminal feature (part of feature geometry)</td></tr>
<tr><td>'</td><td>minute (25' 30" = 25 minutes and 30 seconds)</td></tr>
<tr><td>"</td><td>second (25' 30" = 25 minutes and 30 seconds)</td></tr>
</table>

CHAPTER 1

Introduction: procedure and organization

1.1 DATA: THEIR STATUS WITHIN GENERATIVE THEORY

As I am not a native speaker of English, the data which appear in the examples of this dissertation stem from a computer corpus of spoken Standard American English (SAE)—the *Santa Barbara Corpus of Spoken American English* (*SBCSAE*). The corpus consists of a large body of machine-readable natural speech data, which are used as a basis for an analysis of a class of sounds—the coronal consonants. In contrast to other classes of sounds, coronals display truly astonishing properties. The theoretical framework for my investigation is Optimality Theory (OT), which is the most recent theoretical development within generative theory.

Depending on the theoretical framework within which a linguistic study is situated, naturalistic corpus data have quite a variable status. Such data are collected, for example in the case of the *SBCSAE* (cf. chapter 2.1 below), by making audio-recordings of informants' spontaneous speech in natural settings, e.g. at home, often without their prior knowledge of being recorded. This means the danger of falling into the trap of the *Observer's Paradox* is avoided by researchers. The Observer's Paradox states that, in order to study human communication, data must be collected from informants and the very act of collection, i.e. some form of observation, may affect the phenomena under study (Labov 1972b: 209). In the case of recordings of broadcast news, for example, the effects of observation approach zero. It must be assumed, however, that the Observer's Paradox arises in any kind of conversation when the participants know that they are being recorded, especially where microphones are within plain sight. Here one might expect informants to believe that they should monitor their speech but sometimes they also forget to do so. Naturalistic corpus data differ from data collected from experimental studies in that they do not yield introspective well-formedness judgements. An obvious advantage of experimental studies is that the set-up of the experiment enables linguists to elicit very specific data from the subjects tested. Thus output data are controlled, and accidental gaps can be avoided. On the other hand, most recently compiled corpora have the advantage of providing huge amounts of speech samples, which likewise minimizes the risk of having a large number of accidental gaps in them.

Two polar positions concerning the use of data can be distinguished: either linguists begin with the data and try to devise a theory on the basis of the discovered items and their arrangements; or data are rejected as fragments of performance which by definition cannot yield an adequate model of linguistic structures as represented in the minds of native speakers. These two positions can be attributed to American structuralism and the standard model of generative grammar respectively. Both can be considered too radical and neither will be followed in this investigation.

Within the framework of American structuralism, descriptive accounts of language are corpus-based in the sense proposed by Harris:

> [T]he analysis of a particular corpus becomes of interest only if it is virtually identical with the analysis which would be obtained in like manner from any other sufficiently large corpus of material taken in the same dialect. […] When this is the case, the analyzed corpus can be regarded as a descriptive sample of the language. (Harris 1951: 13)

In this sense my investigation is corpus-based: if the recorded material turns out to be insufficient, the data can—in theory—be supplemented with further native speaker data of the same language variety at any time. This type of procedure is then made unnecessary when additional data will not yield new results.

In contemporary corpus linguistics, two main approaches to corpus work have recently been discussed under the headings "corpus-based" and "corpus-driven". The corpus-based approach "uses corpus evidence mainly as a repository of examples to expound, test or exemplify given theoretical statements" (Tognini-Bonelli 2001: 10). This is, again, the methodology that I subscribe to in my dissertation. In contrast, a corpus-driven approach does not content itself with backing up pre-existing theoretical notions and categories. Instead, the "theoretical statement can only be formulated in the presence of corpus evidence and is fully accountable to it" (Tognini-Bonelli 2001: 11). It is self-evident that an investigation conducted within the theoretical framework of OT cannot satisfy the methodological requirements of a corpus-driven approach.

It should be noted that my analysis is not corpus-restricted either, as in the case of a historical study, for example, which deals with a limited and unchangeable set of historical data. Total reliance on a corpus, on the other hand, and complete negligence of theoretical notions beyond corpus issues will undoubtedly lead to scepticism and even criticism from generative grammarians because their focus is on the introspective qualities of native speakers, i.e. their intuitive knowledge of language. According to the tenets of standard generative theory—OT being a successor theory of generative theory—only a thorough investigation of native speaker competence can shed light on the language system as a whole.

However, it is worth mentioning that even the central figure among generative linguists, Noam Chomsky, does not completely rule out external language data as a means of verification or falsification of hypotheses, although he undoubtedly ascribes the highest degree of reliability to native speakers' (internal) grammaticality and acceptability judgements:

> What is given to the linguist is finite arrays of data from various speech communities, on the basis of which the linguist will attempt […] to discover UG […]. (Chomsky 1986: 11)

And in another place he states somewhat cryptically:

> Research typically begins with examples of structured expressions or, more precisely, with judgements by speakers (or other evidence) […].
> (Chomsky 1988a: 60)

In this context, Chomsky does not clarify what exactly "other evidence" might mean. His obvious reluctance to use corpus data is reflected in his critique of American structuralism (1988b: 523) and specifically of Harris' view of accounting for language, which equates well compiled corpus data with a linguistic description:

> The stock of recorded utterances constitutes the corpus of data, and the analysis which is made of it is a compact description of the distribution of elements within it. (Harris 1951: 12)

According to Chomsky such an analysis attains only the lowest level of adequacy: "observational adequacy".

> [A] grammar that aims for observational adequacy is concerned merely to give an account of the primary data (e.g., the corpus) that is the input to the acquisition device […]. (1964: 29)

Chomsky goes on to claim that data should not be the focus of the investigation. A grammar should not take as its task the description of the input to the language acquisition device but instead, it should at least account for the output of the device, i.e. native speaker knowledge ("descriptive adequacy"). Ideally even, a grammar should provide a principled description of the internal structure of the language acquisition device ("explanatory adequacy"), which enables linguists to construct individual grammars for each language and predict the data from the theory, and not vice versa (ibid.).

In the light of much criticism of his position, it is worth mentioning that about twenty years later Chomsky's attitude towards data appears to be somewhat altered:

> [T]he problem always exists of determining how given experimental evidence bears on theories that involve crucial idealizations. From these inescapable contingencies of rational inquiry one cannot, of course, conclude that experimental data—in this case, studies of performance—are irrelevant to the postulated theories, nor has this absurd proposal ever been advanced, to my knowledge. […]. Still, there persists much misunderstanding about what has been proposed. As a case in point, consider the paper "Logic and the Theory of Mind" by Carol Fleisher Feldman and Stephen Toulmin […]. Feldman and Toulmin assert that I have dealt with the "evidential problem in a dismissive manner", taking the "extreme position" that the "language behavior of laboratory subjects … can neither verify nor falsify the 'psychological reality' of structures in Chomskian grammar", thus regarding "observable data about behavior" as "unsuitable evidence" in contrast to "'intuitions' of a native speaker" which are the "only appropriate substitute". […]. But that experiment and observation of behavior can provide "suitable evidence" is a truism that I (and others) have never questioned and have repeatedly stressed. (1983: 50-1)

In concluding this section, I want to stress that there is of course absolutely no doubt about the fact that Chomsky does value judgements[1] derived from native speakers' intuitions very highly, but, on the other hand, his explicit statements concerning studies of performance like the ones above must be taken seriously and should not be disregarded, as has frequently been done by his opponents (e.g. McEnery & Wilson 1996: 9).

[1] Cf. the relatively recent work of Schütze (1996), who presents a critical evaluation of the use of grammaticality judgements for linguistics, including the dangers of unsystematic data collection. He ascribes variation in the judgement process to a variety of factors, e.g. biological, cognitive, and social differences among informants, differences in elicitation methods, and differences in the materials being judged by informants.

1.2 METHOD

For the analysis of SAE alternations (assimilation, deletion, epenthesis, and dissimilation) and the participation of coronals in these alternations, I will appeal to what is commonly called *internal evidence*, i.e. evidence based on considerations from within the language (or variety) system, and *external evidence*, i.e. evidence gathered outside the language (or variety) system, e.g. the facts of speech perception and production, language acquisition, and typology (Anderson 1981; Paradis & Prunet 1991: 3).

The method used is that of Optimality Theory, as sketched in Archangeli's net metaphor (1997; cf. quote from Popper in section 1.3 below). She likens pre-OT research in generative phonology (GP) to efforts at devising the ideal net which catches all the fish of certain types in a certain area but nothing else, i.e. the undesired fish are allowed to escape. Transferred to grammar, this is a model which accounts for all expressions that belong to a specific language or variety and which distinguishes these from expressions that do not belong to that respective language or variety. The perfect net, however, may be impossible to construct (1997: viii). Thus it followed that GP needed filters and later constraints in addition to the net in order to account for the surface phenomena.

OT shifts its focus of attention away from the net to the construction of the ideal separator, i.e. the set of ranked constraints. This implies that the net catches everything in the area, allowing absolutely nothing to escape, regardless of what is in the water; consequently, all the work is left to the separator (Archangeli 1997: ix). A minor flaw in this scenario is, as LaCharité & Paradis (2000: 230) point out, that previous models of GP catch fish that are already present in the environment while OT actually creates the undesirables (GEN) that the separator (EVAL) subsequently has to remove (for an overview of OT architecture cf. chapter 4.1.2 below). Further critical comments on unresolved issues are added in chapter 9 below.

In my examination of the behavior of coronals in Standard American English (for clarification of what ' Standard American English' is, cf. chapter 2 below), I will use data from the *SBCSAE* for different types of alternation (assimilation, deletion, epenthesis, and dissimilation) to investigate the possibility of accounting for the sometimes controversial behavior of coronals with a suitable subset of (universal) constraints and a sensible ranking of these constraints. By sensible I mean a ranking that does not make controversial claims.

1.3 THEORY: WHAT VERSION OF OT?

> Die Theorie ist das Netz, das wir auswerfen, um "die Welt" einzufangen,—sie zu rationalisieren, zu erklären und zu beherrschen. Wir arbeiten daran, die Maschen des Netzes immer enger zu machen. (Popper 1973: 31) [59]

A fundamental aspect in which theories of language differ is how linguistic items are represented within the theory. This is also true of phonology as a subdiscipline of linguistics. Ideally, a phonological theory should supply representations of all forms or processes that are possible in human language while it should exclude representations of forms or processes that do not exist in human language—admittedly an ambitious goal.

For example, one type of alternation that quite frequently occurs in the languages of the world is the voicing of intervocalic consonants: in Standard American English (cf. chapter 8.2.1.2 below), /t/-voicing or the so-called flapping of /t/ is a widespread phenomenon. Within the theoretical framework of standard generative phonology, Chomsky and Halle's (1968) *Sound Pattern of English* (*SPE*), phonological rules (P-rules) serve as powerful core representational devices for the statement of various process types. However, these representations have proved to be not sufficiently explicit about process types that are either extremely unlikely or even impossible: the diametrically opposed process—the devoicing of interconsonantal vowels—is such a phenomenon hardly attested in natural languages (cf. Schachter 1969: 343-4 for a very early criticism of the *SPE* model). From an *SPE* point of view, however, this rule is just as costly, plausible, and natural as the rule voicing intervocalic consonants, provided that cost and naturalness are measured in terms of feature specifications needed to formulate a rule (Chomsky & Halle 1968: 400). Thus the representation does nothing to rule out the unrealistic process of devoicing interconsonantal vowels.

A further requirement for a theory is economy and simplicity of description—a principle also known as Ockham's Razor: *pluralitas non est ponenda sine necessitate*. It was appropriated for linguistic purposes in the 20th century by Bertrand Russell within the broader field of the philosophy of language in order to eliminate lengthy descriptions and unnecessary terms from analyses. The economy principle can be viewed as a methodological rule for selecting a theory or hypothesis from an array of competing ones. Accordingly, the theory or the hypothesis which supplies the simplest, most economical account of the facts is given priority. A claim often maintained in linguistics is that the simplest and most economical grammatical statements should correspond to the most frequently occurring forms and processes in the languages of the world (e.g. Trubetzkoy 1939; Chomsky & Halle 1968: chapter 9; Hooper 1976; Stampe 1979): this is where the notions of universality, markedness, and naturalness (cf. chapter 4.2 below) enter the picture.

In terms of simplicity and accuracy, autosegmental theory for instance, is an improvement over its predecessor *SPE*, because autosegmental spreading is a 'better' representation of the process of assimilation than changing each individual feature in feature matrices if, for example, a whole group of features (e.g. labial, coronal, or dorsal) is affected. Within a nonlinear framework such a complex process can be represented just as simply as if a single feature were involved: this is achieved by a hierarchical feature representation—distinguishing between terminal features, such as [continuant], [nasal], [anterior], [ATR], and nonterminal class nodes, such as the PLACE node—and subsequent autosegmental spreading of a major class node. Nonlinear representations are assumed in most OT literature on phonology (cf. chapter 4.3.4 below).

A linguistic theory is a model of language or a model of one of its components, and additionally a set of rules, laws or constraints (OT) that relates the representations (entities) and the levels of description within the model to our observations. A theory merely exists as a set of conceptualizations in the minds of human subjects and possesses no reality whatsoever beyond this. Strictly speaking, a theory is always preliminary, a mere hypothesis, because it can never be verified.

> Unsere Wissenschaft ist kein System von gesicherten Sätzen, auch kein System, das in stetem Fortschritt einem Zustand der Endgültigkeit zustrebt. Unsere Wissenschaft ist kein Wissen (episteme): weder Wahrheit noch Wahrscheinlichkeit kann sie erreichen. (Popper 1973: 223) [278]

Karl Popper's important book *Logik der Forschung*, which contains the fundamental principles of critical rationalism, rejects the utopian ideal of ever proving a scientific theory. Instead, a theory must in principle be falsifiable through observations, otherwise it is not a theory; this is known as the falsifiability criterion (1973: 47-59

[78-92]). Frameworks must be subjected to constant criticism and, if necessary, be modified, e.g. by rescue procedures, so-called auxiliary or ad-hoc-hypotheses, or else be abandoned if inconsistencies and contradictions grow rampantly, and eventually negative evidence turns out to be overwhelming. In such a situation, the time is ripe for the development of a new theory. Popper presents an example from classical mechanics: Newton's laws of motion and universal gravitation were substituted for by Einstein's theory of relativity with its main claims that all motion is relative to a frame of reference and that time and space are also relative rather than absolute concepts. Einstein's theory more accurately accounts for the observed facts than Newton's because it correctly predicts the formerly inexplicable elliptical orbits of the outer planets in our solar system. The new theory introduced new physical effects and opened up new possibilities for testing and for falsifying it (1973: 51 [83]).

Returning to the phonological microcosm, ideally a phonological theory, too, provides falsifiable assumptions, which are visible in representations and their interactions. These representations are supposed to model the structure of linguistic knowledge in the human brain. According to Chomsky (1964: 28), such a theory meets the condition for descriptive adequacy, but it should be noted that not all (adequate) descriptive devices are of equal value. A higher level of adequacy can be attained if a theory provides representations that lead to explanations of why the observed facts are as they are and why the representations are thus and not otherwise—a claim totally in line with OT's demands on a theory.

> [A] linguistic theory that aims for explanatory adequacy […] aims to provide a principled basis, independent of any particular language, for the selection of the descriptively adequate grammar of each language. (Chomsky 1964: 29)

Such a theory will predict, to a certain degree at least, which forms and processes are to be expected and which are not (Popper 1973: 32 [60]). In addition, it seems reasonable to claim that a high explanatory content of theoretical formulations will guarantee and enhance Popper's falsifiability criterion, by inviting linguists to propose different explanations.

In phonology, a way of achieving explanatory adequacy is through the grounding of representations in phonetics (cf. Archangeli & Pulleyblank's 1994 study *Grounded Phonology*, discussed in chapter 4.3.3). I claim that by allowing such factors as vocal tract anatomy, acoustics, and perceptual phenomena to contribute empirical evidence to explanations, phonological theory is capable of producing representations that account for phonological facts in a nonarbitrary fashion. One of the main aims of phonology is to identify the distinct but interacting forces that underlie both the crosslinguistically common and the language-particular sound systems which are revealed through the analyses of language data.

Phonology thus conceived is an inherently interdisciplinary field of study; subject areas that are commonly attributed to phonetics are automatically subsumed under phonology. That is, articulatory properties of speech sounds as produced by the vocal organs of the speaker, acoustic properties of the physics of speech sounds as transmitted between speaker and hearer, and auditory properties of the perceptual responses to speech sounds, usually with respect to the hearer, are all relevant factors within a unified phonetics-phonology model.

(1.1) Extralinguistic factors contributing to phonological theory

- constraints on speech production and perception (distribution of sound segments in the system)
- cognitive mechanisms employed by language learners to organize the enormously rich content of the ambient speech signal into a functionally coherent linguistic system (language acquisition, learnability)
- functional principles underlying the interaction of phonetic and cognitive forces that effect changes in sound systems within and between speech communities, and over a period of time (alternations, sound change)

A large amount of literature is concerned with exactly these issues (e.g. Anderson 1981; Keating 1985; 1988): should both phonetics and phonology be regarded as totally autonomous disciplines, is phonetics a branch of linguistics (or is it closer to the natural sciences), or is there a way of integrating the two in a unified grammar model? In the sense that phonetics is concerned with the phonetic properties of a specific language (English) or even a variety thereof (SAE), and with the question of how the sounds function within the pronunciation system (of that variety), I suggest that the two in fact work hand in hand and I propose the use of a unified OT-model of phonetics and phonology for my investigation of coronal consonants in SAE (cf.

also Hayes 1995; 1996; Flemming 1997; Boersma 1998; Kirchner 1998; Steriade 2000; Flemming 2001; 2002).

The question of the explanatory adequacy of theoretical statements can be linked directly to the notions of psychological reality and cognitive adequacy. A condition on OT could be the demand that as a theoretical construct OT will only be psychologically real if some empirical correlates are connected with it. In the true rationalistic tradition, Chomsky values most highly not the experimental or empirical findings, which for him count as performance data, but the explanatory force of the hypotheses made. Linguistic phenomena are psychologically real by virtue of being part of the internalized grammar, which is subject to introspection. The linguistic competence of native speakers is usually tested with judgements on grammaticality that serve as verificational or falsificational data for the theory in question. However, since I am not a native speaker and therefore cannot supply data myself, it is legitimate for me to use the *SBCSAE* as my data base for the investigation of coronals.

1.4 ORGANIZATION OF THE DISSERTATION

The rest of the dissertation is organized as follows. In chapter 2, I briefly introduce the corpus (*SBCSAE*) which I used as my data base. After a description of the type of English contained in the corpus, I attempt to give a definition of *Standard American English*.

In chapter 3, I provide a classification of coronal sounds as well as an overview of how the feature emerged and gained importance during its history. The chapter mainly deals with the phonetic properties of coronal sounds. Acoustic and perceptual properties are discussed, but the focus is on the *articulatory* dimension of coronals. Articulatory Phonology (AP), a gestural model developed by Browman & Goldstein (1986; 1989; 1990; 1992), is presented as an example hereof and discussed in greater detail since AP gives a unifying and explanatory account of various connected speech alternations that require a number of separate and sometimes quite arbitrary phonological rules in featural phonology.

This naturally leads to the question of how coronals are represented within a phonological theory. In chapter 4, I first provide a short introduction to Optimality

Theory. After discussing the relevance of the notions of *markedness* and *naturalness* for an OT grammar, I show which representations and levels of analysis are favored in OT frameworks. I also introduce the Grounding Hypothesis (Archangeli & Pulleyblank 1994) and the extension of grounding conditions into OT as constraints.

Chapter 5 focuses on the *other* dimension of the communication event, not on the articulatory side of the speaker but on the *perceptual* domain of the hearer and its significance for a unified theory of phonetics and phonology. Functionalist approaches are set off against formalist approaches to phonology, and three different models of representing phonological contrast are presented. I argue that functional notions, e.g. those which refer to the requirements of speech perception, must be integrated into a model of phonology that aims at explanatory adequacy. In my view, OT has the potential to do so in that the possibly contradictory requirements of speech perception, such as the principle of *maximization of perceptual distinction*, and of speech articulation, such as the principle of *minimization of articulatory effort*, are formalized as conflicting and violable OT constraints.

In chapter 6, the distributional properties of coronals are discussed with respect to the SAE segment inventory and SAE syllable structure.

After giving a brief overview of pre-OT analyses of phonological alternations in chapter 7, I propose an OT model that attempts to unite both phonetic and phonological aspects of alternations. I also introduce various sets of OT constraints that are needed for an OT analysis of SAE alternations.

Employing the grammar model of the previous chapter, an analysis of SAE alternations, finally, is the subject of chapter 8. The two main, opposing types of alternation that coronals in SAE participate in are examined. These are weakening or lenition phenomena, i.e. *assimilation* and *deletion*, and strengthening or fortition phenomena, i.e. *epenthesis* and *dissimilation*. Coronals are notorious for their contradictory behavior, especially in alternations; I want to show that this type of behavior can be accounted for within a phonetically grounded OT framework.

Chapter 9 is intended to give a short overview of residual or unresolved issues with respect to the proposed OT model. This 'critique' includes the question of *markedness* as an explanatory notion, the status of universals, and problems connected with OT's basic machinery, i.e. the LEXICON, GEN, and EVAL.

CHAPTER 2

The SAE corpus data

2.1 THE ENGLISH OF THE *SBCSAE* (PART I)

As early as 1991, Wallace Chafe, John du Bois and Sandra Thompson announced a "new corpus of spoken American English" (Chafe et al. 1991: 64). From then it took almost a decade to the actual publication of the first part of the *Santa Barbara Corpus of Spoken American English* (*SBCSAE*) in an interactive computer format. In February 2000, Part I of the *SBCSAE* came out on three CD-ROMs containing 14 speech files with 50 different speakers amounting to roughly 70,000 words and 5½ hours of speech.

According to information given by Lenk (1998), who worked with the unpublished *SBCSAE* at the University of California at Santa Barbara (including the still unpublished Parts II and III), a total of 46 speech events was recorded, adding up to about 15 hours of face-to-face conversation (cf. figure 2.1 below). Having about 200,000 words, the compilers of the new corpus attempted to achieve comparability with the face-to-face conversation portion of the *London-Lund Corpus of Spoken English* (Chafe et al. 1991: 68). The face-to-face material of the *LLC* contains 220,000 words of spoken British English (Svartvik & Quirk 1980: 12) and was published in its printed, computerized versions in 1980.

Lenk reports that originally a corpus of 1 million words amounting to 100 hours of speech in 200 files was planned (1998: 8), but which was finally reduced to one fifth of this size. Also, Chafe et al. had originally announced

> that researchers will be able to see a transcription on their micro-computer screen and to hear it at the same time. Moreover, users will be able to search the written transcripts for relevant key words using standard English orthography and then hear the corresponding passages, thus gaining access to what is in effect an auditory concordance. (1991: 66)

Unfortunately, this ambitious goal of aligning the speech files with the corresponding transcripts to make the life of researchers easier was not realized.

(2.1) Titles of all 46 *SBCSAE* transcripts (after Lenk 1998: 213-4). Titles marked with * will not be included in the final version of the *SBCSAE*; for titles marked with # the final decision on their inclusion was still pending in July 1997; titles in boldface form Part I of the *SBCSAE*.

Actual Blacksmithing
Africa *
American Democracy is Dying
Ancient Furnace
Appease the Monster
Atoms Hanging Out
Attorney
Bank Products
Beaten On a Regular Basis
A Book About Death
Bridge Story *
Car Sales *
Charades #
Conceptual Pesticides
Cuz
Deadly Diseases
Doesn't Work in this Household
The Egg Which Luther Hatched
Farmtalk #
Fear
Fiesta Party
God's Love
Guilt
Hey Cutie Pie
Howard's End #
A Hundred Million Dollars
Hypochondria *
Lambada
Letter of Concerns
Lunch *
Mickey Mouse Watch
One Man Dialogue #
Philosophical Differences
Raging Bureaucracy
This Retirement Bit
Runway Heading
Small Claims Court
Tapedeck
Tell the Jury That
A Tree's Life
Try a Couple Spoonfuls
Vet Morning
Vision
Wonderful Abstract Notions
X Units of Insulin *
Zero Equals Zero

In view of the speed of developments in the field of computer technology along with the compilation of ever larger computer corpora, it is obvious that a corpus of the rather small size of about 70,000 words (or even a total of 200,000 words) can hardly compete with a corpus such as the *British National Corpus* with an estimated 100 million words (including roughly 10 million words of spoken English) which legitimately aims at representing the universe of contemporary British English. In this sense, a corpus as small as the *SBCSAE* cannot claim to characterize the state of contemporary American English; with 14 25-minute recordings, it must be regarded rather as a more or less random collection of spoken discourse. On the other hand, it is worth noting that a corpus which consists primarily of audio recordings—files in the *Windows wav*-format take up extremely large amounts of computer storage capacity—can never be expected to be as comprehensive as a corpus that merely contains the transcripts of spoken discourse without the actual recordings. For the present phonological analysis, however, the corpus proved large enough.

The speech files of Part I are accompanied by transcript files in which utterances of speakers are time-stamped with respect to the corresponding audio recordings. The transcription of the spoken material into standard English orthography—enriched with the symbols of discourse transcription as printed in Schiffrin (1994: 422-3)—allows for the search for relevant linguistic units in the written transcript files with a suitable computer program. I have chosen *WordSmith Tools* (version 3.0), which offers the following basic functions (tools):

- *Concord* creates concordances, which are listings of the occurrences of a particular feature (e.g. a word) or combination of features in a corpus; each hit is displayed with a specifiable amount of context (0-25 words to the left of the search item, 0-25 words to the right)
- *WordList* generates lists of all words or word clusters in a text, either in alphabetical or in frequency order; it provides statistical means, e.g. total number of words, length of words, number of sentences
- *KeyWords* identifies the key words in a given text

Search-words or phrases, of which concordances are made in the chosen text files, are annotated with a time count in seconds in the transcripts of the corpus. The annotated time count facilitates listening to the corresponding audio sections in the speech files. This task is managed by the computer tool *VoiceWalker* (version 2.0)—a discourse transcription utility which functions as a digital CD deck and is controlled from the computer keyboard. *VoiceWalker* combines the high-quality stereo sound of a high fidelity tape deck with a far more precise control over audio playback than a specialized foot-pedal tape machine can provide. In this way, a kind of auditory concordance is achieved.

In addition to the transcript and speech files, the CD-ROMs contain a table of contents with a list of the conversational topics (cf. figure 2.2 below), a *speaker table.doc* (cf. figure 2.3 below) and a so-called *readme* file, which, in my opinion, all give only the most superficial information about the nature of the corpus: the recordings are claimed to be representative of regional, social and ethnic variation in an array of situational contexts such as conversation, gossip, arguments, job talk, card games, city council meetings, sales pitches, classroom lectures, political speeches, bedtime stories, sermons and weddings. Names and phone numbers have been altered in order to preserve the anonymity of the speakers.

(2.2) List of the 14 speech files of the *SBCSAE* (Part I) on three CD-ROMs

sbc1_1/speech	sbc0001.wav	Actual Blacksmithing (~25')
	sbc0002.wav	Lambada (~24')
	sbc0003.wav	Conceptual Pesticides (~26')
	sbc0004.wav	Raging Bureaucracy (~19'30")
	sbc0005.wav	A Book About Death (~20'30")
sbc1_2/speech	sbc0006.wav	Cuz (~27')
	sbc0007.wav	A Tree's Life (~23')
	sbc0008.wav	Tell the Jury That (~25'30")
	sbc0009.wav	Zero Equals Zero (~25')
	sbc0010.wav	Letter of Concerns (~15'30")
sbc1_3/speech	sbc0011.wav	This Retirement Bit (~20'30")
	sbc0012.wav	American Democracy is Dying (~25'30")
	sbc0013.wav	Appease the Monster (~27'30")
	sbc0014.wav	Bank Products (~28'30")

In my view, it is problematic that a commentary on the specific data collection procedures is not included in the documentation that accompanies the corpus. How the goal of a representative sampling was achieved, how regional, social and ethnic varieties are dealt with, what demographic techniques were employed, simply remains unexplained. For this reason, the corpus user is left with the more or less unreliable information given in the 1991 preview article by Chafe et al. regarding details on questions of standard vs. nonstandard, accent, register, style, etc. of the recordings included in the *SBCSAE.*

The so-called *speaker table.doc*, which is included in the electronic version of the corpus, is a list of the 50 recruited adult speakers (aged 16 to 90 years, average age is 36) with eleven speaker parameters, including their pseudonyms and some information about their gender, age, dialect, social class, and ethnicity. The percentages in figure 2.3 below are calculated on the basis of this list. The sampling resulted in approximately equal numbers of both genders and of each of the three age groups, but when it comes to regional, social, and ethnic factors, there is an obvious 75% predominance of the white educated speaker of General American (GA)—provided that the assumption which is explained in section 2.2.1 holds and the Western accent region is largely congruent with the region where GA is predominantly spoken (Wells 1982: 471), and provided that the native origin of speakers recorded in the West is actually this area. Since the design criteria of the corpus are not made explicit anywhere, one can only speculate on possible reasons for the unbalanced selection procedure.

(2.3) Table with five of eleven speaker parameters[2]

gender	28	female
	22	male
age	17	16-26 years
	15	27-40 years
	14	41-90 years
	4	unclassified
dialect	20	Western
	8	Northern
	3	Southern
	2	unclassified
education	35	college or higher
	9	high school
	6	unclassified
ethnicity	37	European American
	7	Hispanic
	2	African American
	2	Native American
	1	Japanese
	1	unclassified

The table obviously does not constitute a representative cross-section of the American population; 20 speakers, i.e. 40% of all speakers, are from California alone.

[2] The other six parameters are: speaker id, name used in transcript, state of birth, state of residency, years of education, and occupation.

A few more details on the sampling can be found in the article "Towards a new corpus of spoken American English" (Chafe et al. 1991). On seeing the title of the article, one question immediately comes to mind: what accent or variety is meant by "American English"? Chafe et al. restrict the term to Standard American English and much of their definition coincides with my discussion presented in section 2.2 below:

> Spoken SAE encompasses formal and informal styles, […] and a host of regional, class, gender and ethnic accents. What unifies it is a shared set of grammatical rules and structures, which among spoken varieties of American English come closer than any other to the written variety as used in journalism, […]. Because SAE encompasses a variety of accents, it is not the sole property of any particular group in American society. (1991: 69)

This may be the reason why few speakers of other than European American ethnicity were included.[3] Chafe et al. also characterize SAE rather vaguely as something like network American, "the variety of English used in spoken and written journalism" (ibid.). The standard variety was chosen, they say, in order to enable comparison with other existing corpora of a standard variety, and because speech corpora could not be made large enough to represent more than a single variety.

In short, the well-meant claim that the *SBCSAE* is devised to reflect the social, geographic, and ethnic diversity of the American population (1991: 71) appears somehow weakened in connection with the following qualification: "as far as is possible within the rather severe limitations on corpus size" (ibid.). The quotation below even contradicts the data given in figure 2.3 above:

> We have decided to construct a corpus that is relatively homogeneous with respect to Standard American English but varied with respect to such features as age, sex, region, educational level and occupation, […]. (ibid.)

To say the least, the numbers for region, education, and ethnicity seem to be somewhat off balance with a ratio of 75% to 25% in each category.

[3] However, while it is true that African American Vernacular English (AAVE), for example, does not pass as a standard variety, it is not true that all African Americans necessarily are speakers of the AAVE variety. Thus, Chafe et al.'s criteria for the selection of their informants appear to be not properly thought-out.

2.2 STANDARD AMERICAN ENGLISH

As mentioned above, the variety used by the subjects recorded for the *Santa Barbara Corpus of Spoken American English* is classified as General American (GA), which raises questions about regional and social varieties. Wells (1982) gives a description keeping reservations about the dialectological status of General American in mind:

> In the United States, it is true not just of a small minority, but of the majority, that their accent reveals little or nothing of their geographical origins. They are the speakers of General American […]. This is a convenient name for the range of United States accents that have neither an eastern nor a southern colouring; dialectologically, though, it is of questionable status. (1982: 10)

The area where unmarked GA accents are mainly found is roughly congruent with the area termed *the West* in dialectological studies. I am aware of the fact that dialectologically, where all components of the grammar—syntax, the lexicon and pronunciation—are investigated, such a statement is of course highly questionable. With a focus on phonological distinctions and more specifically on coronal consonants, it is viable since there is hardly any variation with respect to English consonant inventories.

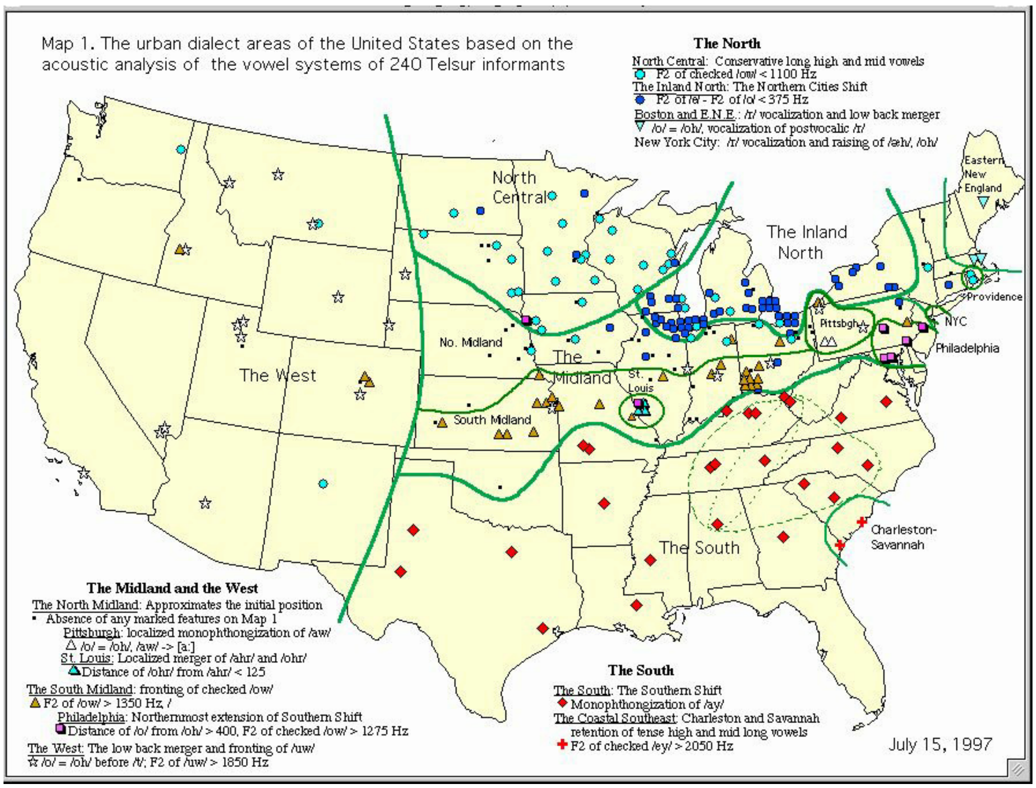
2.2.1 Western as the default variety

Traditionally, regional varieties are determined by the use of different lexical items in the relevant areas, and isoglosses are established that delimit the regional boundaries of the respective varieties. Recently, more emphasis has been placed on phonology for the demarcation of these boundaries (Labov 1991: 2; Wolfram & Schilling-Estes 1998: 91). Where phonology is concerned, the crucial role in distinguishing sound inventories is played mainly by the vowel systems; neither the consonant systems of the Englishes world-wide (Giegerich 1992: 43) nor the consonant frequencies vary significantly (Wang & Crawford 1960: 138).

On the basis of two types of phonetic change which result in a considerable diversity of vowel systems in American English, Labov identifies three major types of accents. One of the two phonetic processes is chain shifting, which is responsible for the widely contrasting vowel systems of two regional areas: the North (Northern

Cities Shift) and the South (Southern Shift).[4] The general principles which govern chain shifts are the raising of tense or peripheral vowels (with respect to their location on the vowel trapezium), the lowering of lax or nonperipheral vowels, and the fronting of back vowels. By way of a perpetual rotation of phonetic features, e.g. [tongue height], phonologically relevant distinctions in the affected vowel systems are maintained.[5]

(2.4) Map of the three major dialect regions of the US



URL: http://www.ling.upenn.edu/phono_atlas/NationalMap/NatMap1.html

[4] More generally even, in his article "The Three Dialects of English" Labov claims a tripartite division for all Englishes, not just the English spoken in the US (1991: 4), with accidental dialect similarities between the Northern Cities Shift in the US and its British relatives in the north of England as well as between the Southern Shift and the dialects of e.g. southern England.

Starting out from the same basic assumption that accents of English differ mainly in terms of their vowel systems, Giegerich also recognizes three main accent varieties. For him, the inventories of GA, RP and SSE (Standard Scottish English) represent very radical differences in the vowel phonology of English, the three pivotal points in an accent typology (1992: 43).

[5] I will not go into detail concerning the vowel systems because a further analysis of these is irrelevant for an investigation of coronal consonants.

The third major accent type which is of special interest here is defined chiefly by the absence of chain shifting and—stated affirmatively—by the presence of the other phonetic process type: merger. Compared to chain shifting, merger works in quite the opposite way. Formerly or elsewhere distinctive features have been neutralized, and distinctions within the system have been lost. The so-called "Third Dialect" (Labov 1991: 30) is characterized by the merger of the low /ɑ/ and the back /ɔ/ vowels as in the pair <cot> and <caught>; the lax vowel [ɑ] appears in both instances, making the two homophones. Geographically the "third region […] covers a vast […] area of the western United States with no particular urban center." (Labov 1991: 31)

The map above is based on a current nationwide telephone survey (Telsur) and is taken from the paper "A National Map of the Regional Dialects of American English" which Labov, Ash and Boberg published in 1997 on the internet. The three major regions identified on the map—the North, the South, and the West—correspond to the three accent areas presented in Labov's 1991 article "The Three Dialects of English".[6]

The map also illustrates the fact that vocalic variation is decisive for the differentiation of accents. Consonantal variation merely plays a marginal role, and coronals are not affected categorically. A brief list of the few significant cases is as follows:

- Postvocalic /r/ is present in Western speech whereas Northern and Southern speech is frequently nonrhotic, although even in these areas the pronunciation of postvocalic /r/ seems to become more and more prestigious (Cassidy 1983: 203, 206).
- In Western speech, contrast with respect to the feature [voiced] is frequently suspended in the coronal plosives /t/ and /d/. If they appear in the vicinity of other voiced segments and at the onset of an unstressed syllable, the stops are typically flapped, i.e. voiced[7] (Kahn 1976: 56-61). In the North, due to the model function of British English, contrast is frequently retained. However, there is a tendency towards the more prestigious American pronunciation of

[6] For maps with a very similar regional division cf. Cassidy (1983: 202) and Wolfram & Schilling-Estes (1998: 121).

[7] Some scholars represent this voiced sound as a [d] (Webster's Third 1961: 41a), some call it a "voiced t" (Oswald 1943; Sharf 1960; Wells 1990: 703; Kenyon 1994: 126), and some an alveolar tap or flap [ɾ] (Trudgill & Hannah 1994: 42). In chapter 6 this sound will be subject to further investigation.

<dirty> as [dɜrɾɪ] (sbc0006.wav: 14'33") with a voiced flap instead of a voiceless [t] sound.

- Neutralization of contrast in the pairs /t/ ~ /θ/ and /d/ ~ /ð/ as in /tɪn/ for <thin> and /dɛn/ for <then>, and fricative stopping as in /bɪdnɪs/ for <business> which was thoroughly investigated by Schilling-Estes (1995); both these features are typical of the South and do not occur in the North or West.

The second edition of the *Random House Unabridged Dictionary*, a self-acclaimed 'authoritative' American dictionary, has the following definition of the term "General American":

> [A]ny form of American speech considered to show few regional peculiarities, usually including all dialects except for Eastern New England, New York City, Southern, and South Midland […]. (1993: 795)

For the purpose of this investigation and in accordance with Wells' definition of GA—a variety devoid of Northern and Southern pronunciation characteristics (1982: 10)—I claim the Western accent family to be a 'neutral' or default accent, because the region where GA is predominantly spoken (Wells 1982: 471) is largely congruent with the Western speech area as delineated by Labov (1991).

Obviously, within a large geographical region such as the West, the accents of individual speakers (idiolects) vary considerably.[8] Instead of assuming a homogeneous accent within certain isoglosses, it is preferable to imagine each idiolect located on an accent continuum with idiolects sharing salient features of the area in question—for my purpose consonantal features. For this reason, I will limit my classificatory attempts to assuming, rather roughly, a Western accent family which corresponds to the cover term GA.

Chafe et al.'s (1991) claim that the speakers included in the *SBCSAE* are all speakers of GA or SAE not only raises the question about the regional or rather non-regional character of this specific variety, by referring to a 'standard' they also hint at sociolinguistic aspects of a specific variety.

[8] The San Francisco and Los Angeles metropolitan areas, for example, do not participate in the low back vowel merger which is characteristic of the Third Dialect (cf. map by Wolfram & Schilling-Estes 1998: 121). But again, only the vowels are affected by the variation, and so it seems justified not to elaborate on this point any further. Moonwoman's (1992) dissertation about sound change in San Francisco English and Decamp's (1971) article on the pronunciation of English in San Francisco substantiate this decision.

2.2.2 Standards: formal, informal, vernacular

The *Random House Unabridged Dictionary* has the following definition under the entry "Standard English":

> [T]he English language in its most widely accepted form, as written and spoken by educated people in both formal and informal contexts, having universal currency while incorporating regional differences. (1993: 1857)

The dictionary definition sets written and spoken language on a common footing and subsumes both formal and informal contexts under the label Standard English. It allows for regional variation, but excludes social variation by explicit reference to the 'educated' speaker. It is obvious from this definition that one social group is given the authority to determine for all speakers how language is best used.

In the sociolinguistics literature, 'standard' is anything but a straightforward notion, its very meaning can be observed as oscillating between the poles 'uniformity' and 'level of excellence'. For the sake of clarification of this issue, quite a few linguists are involved in a "standard English debate" at this time (cf. e.g. Cameron 1995: 81-2; Thomas & Wareing 1999: 151-71; Bex & Watts 1999: *Standard English: The Widening Debate*; Milroy & Milroy 1999: *Authority in Language: Investigating Standard English*). Attempts at formulating a precise linguistic definition of 'standard' are constantly influenced by ideological notions about standard English, which is claimed to be the 'correct' form of the language, whereas other varieties are supposedly 'incorrect'. This is why Lippi-Green speaks of a "standard language myth" (1997: 53) and Cameron goes so far as to analyze the fear of losing a standard and correct usage metaphorically. In the chapter "Grammar as Moral Metaphor" of her book *Verbal Hygiene* she relates the rejection of nonstandard varieties to anxieties about disorder and lawlessness in society:

> A panic about grammar is therefore interpretable as the metaphorical expression of persistent conservative fears that we are losing the values that underpin civilization and sliding into chaos. (1995: 95)

Frequently discussions about standard and nonstandard refer exclusively to lexis and syntax, in other words to Trudgill's definition of dialect, and not to accents, which are related to pronunciation (1995: 5). He illustrates his claim with two triangular diagrams which are supposed to represent the situation in Great Britain. While both triangles show that it is impossible to neatly separate regional and social variation, only the dialect triangle shows the distribution of standard English. The accent triangle is solely concerned with variation in pronunciation and presents the location of the most prestigious accent in Great Britain—Received Pronunciation (RP)—at the very top (1995: 29-30). In his recent paper with the telling title "Standard English: What it Isn't" Trudgill explicitly formulates the radical view that accents should be excluded from the definition of standard English:

> There is one thing about Standard English on which most linguists, or at least British linguists, do appear to be agreed, and that is that Standard English has nothing to do with pronunciation. (1999: 118)

This view does not go unchallenged by Milroy, for example, who claims RP to be acknowledged as the standard English accent by professional linguists (1999: 173-4). In the US she identifies 'network American' as the standard English pronunciation, although as a mainstream accent this variety and RP (spoken only by the highest class) "are horses of a very different colour". She also concedes that "in reality many high profile American broadcasters speak with regional accents" (note on page 174).

For American English the situation seems to be different. Wolfram and Schilling-Estes distinguish between three varieties: a formal standard, an informal standard, and vernacular speech, which is sometimes also called a nonstandard variety. The so-called formal standard variety of English is commonly linked to written, not to spoken language. Apart from some minor exceptions with respect to national variation (American and British), Quirk et al. also emphasize a striking world-wide uniformity in syntax, lexis and orthography with a tendency toward even greater homogeneity under the influence of ever closer world communication:

> [O]ne can frequently go on for page after page without encountering a feature which would identify the English as belonging to one of the national standards. (1985: 19)

Wolfram and Schilling-Estes even equate formal standard English with "prescriptive standard English" (1998: 10), since it is this variety that is codified in grammars and guides to usage. It is also perpetuated in formal institutions which claim the right to authorize a certain usage over another: e.g. respected schools and universities, government agencies, political parties, the press, or in jurisprudence. Apparently, it is a very conservative variety, as advocates of it have trouble tolerating change within the language. According to Wolfram and Schilling-Estes, it is a variety which is virtually not maintained in natural spoken language, but only in the most formal contexts (ibid.).

Where pronunciation comes into play the situation is considerably different. Although there is allegedly only one formal standard spoken by educated speakers in Great Britain (RP), an informal standard pronunciation used in every day natural speech allows for a good deal of variation and thus cannot be regarded as a monolithic entity, regardless of the speech community concerned:

> [E]xperts and laypersons alike have just about as much success in locating a specific agreed spoken standard variety in either Britain or the United States as have generations of children in locating the pot of gold at the end of the rainbow. (Milroy 1999: 173)

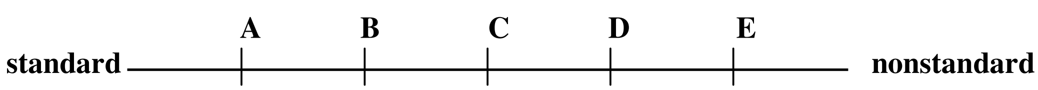
However, since the very term 'standard' does not seem to be consistent with a high degree of variability, Milroy & Milroy suggest "speak[ing] of standardisation as a historical process" motivated by social, political, and commercial needs; in short it is an "ideology" directly opposed to variation and change, and promoted chiefly through the writing system, which is standardized much more easily than pronunciation (1999: 19).

As to the situation in the US, many linguists seem to agree that in American English there is no variety similar to RP. As Romaine claims: "American English has pluricentric norms for pronunciation with the different regions having their own standards used by educated speakers." (1998: 39; for a similar statement cf. Wolfram & Schilling-Estes 1998: 11). Just how many standards should be recognized is open to question. Bauer, for example, finds it reasonable to distinguish standards as local as San Francisco English and Seattle English (1994: 3). Fromkin, Rodman and Hyams even claim that a uniform standard US pronunciation is nothing but an 'idealized' spoken variety:

> SAE [Standard American English] is an idealization. Nobody speaks this dialect; and if somebody did, we would not know it, because SAE is not defined precisely. […]. It used to be the case that the language used by national newsbroadcasters represented SAE, but today many of these people speak a regional dialect, or themselves "violate" the English preferred by the purists. (2003: 455)

While the determination of a formal standard, especially when applied to spoken English, is not an unambiguous enterprise, a definition of an informal standard is even more difficult to provide. Nevertheless, it seems to be this term which is mainly applied to spoken language and which best fits the data. In order to account for informal standard English, Wolfram and Schilling-Estes propose a continuum of standardness ranging from Speaker A using few, if any, nonstandard forms to Speaker E using many (cf. figure 2.5 below). Assuming that Speaker A will be rated as a speaker of formal standard English and Speaker E will be regarded as a speaker of a nonstandard or vernacular variety, the informal standard variety is located somewhere between Speakers B and D. Indeed one can say that an informal standard, too, is a pluralistic notion in the sense of multiple[9] norms (cf. Romaine's claim above) since there is no clear dividing line between formal and informal standard but a considerable amount of overlap, especially with respect to spoken language.

(2.5) Continuum of standardness (Wolfram & Schilling-Estes 1998: 11)



It is also important to note that such ratings are of course a highly subjective matter because they depend on assessments that members of the speech community make about other speakers' degree of standardness.[10] In sum, the informal standard is determined by multiple norms of acceptability and defined negatively, rather by what it is not than by what it is, e.g. by the avoidance of syntactical, lexical, or pronunciational features that are regionally or socially conspicuous. This general difficulty in arriving at something like a clearcut definition of standard is reflected in the title of Trudgill's 1999 paper mentioned above—"Standard English: What it Isn't".

---

[9] The situation is even more complicated when two national varieties are compared with respect to an informal standard in spoken language. The results of a recent study of face-to-face and telephone conversations confirm the widely held view that the standard in spoken American English is more relaxed and informal than in spoken British English (Helt 2001: 171-83).

[10] These attitudes towards language are commonly investigated by linguists under the heading 'folk dialectology'. To elicit attitudes, informants from a certain state of the US are presented with a geographical map of the US and asked to rate all states according to the criterion where the most 'correct' English and the least 'correct' English is spoken. The findings from perceptual dialectology in the US are commonplace: the North and the South have the lowest rankings for 'correctness' and 'pleasantness' (see e.g. Preston 1993: 333-77; Hartley & Preston 1999: 207-38). It has long been claimed that folk notions such as these have to be taken seriously since they may be linguistically relevant as directors and triggers of language change (Hoenigswald 1966: 19).

Finally, vernacular or nonstandard varieties, which also mainly refer to spoken language, are characterized by the presence of socially disfavored linguistic structures. The assessment 'disfavored' of course rests on 'authoritative' speaker judgements, which are essential in determining social unacceptability and which commonly result in stereotyping speakers who use these stigmatized forms as 'uneducated'. In other words, overt prestige is assigned to all standard varieties and vernacular varieties are stigmatized, not only by higher-status groups but across a range of social classes. Along with these widely accepted norms, however, there exists another set of norms which is based on solidarity among more locally defined groups irrespective of their social status. Within a local culture, otherwise stigmatized linguistic structures may carry covert prestige in the sense that these forms stand for the feeling of a common bond. Thus a vernacular variety is bound to be stigmatized in one social context and typically receives covert prestige in another setting (Wolfram & Schilling-Estes 1998: 159).

What has been neglected so far is that language not only varies in use according to speaker variables such as regional and social background, gender, class, and ethnicity (see above), but also according to situational variables such as e.g. medium, style, and subject matter, commonly called register. Variation according to speaker means that speakers use one variety and use it all the time, in other words: the choice of a particular variety (e.g. regional or social) is predetermined for each speaker.[11] Variation according to situation[12] on the other hand means that speakers have a range of varieties and choose between these at different times depending on their perception of the demands of the situation.

The variety which seems to come closest to the corpus data is a spoken informal standard American English as it is used by newsreaders on radio and television networks in the US.

> In the US, the standard includes most if not all avowedly educated regional accents and usage, as well as—for teaching purposes—an accent loosely referred to (amid scholarly controversy) as both *General American* and *Network Standard*, the presumed usage of radio and television newscasters, […].
>
> (McArthur 1998: 117)

[11] This of course does not mean that the speaker is not capable of learning an additional variety.

[12] Halliday et al. (1968: 141) and Quirk et al. (1985: 16) speak of "varieties according to users" and "varieties according to use" in this context.

Network American—some linguists use this term synonymously with Standard American English (SAE)[13]—is an informal mainstream variety, and as such it is subject to some degree of inconsistency. It seems acceptable for American speakers to have at least a residue of regional, if not social, dialect in their speech.

The following chapters will present a classification of coronals and an analysis of their distributional and alternational properties in SAE within the theoretical framework of OT. The variety which I chose as data for my dissertation is SAE, as defined by Chafe et al. (1991) and as contained in the *SBCSAE*. I have also showed that coronals—being consonants—are not really affected even given the notion *variety*: it is the vowels that are responsible for the greatest amount of variation among the varieties of English, whereas the consonants, and especially the set of coronals, are quite stable across the different varieties of English. As a source of reliable examples for this dissertation, the use of the *SBCSAE* is not objectionable regardless of its small size, its unbalanced sampling method, and the rather vague varietal classification of its data (cf. section 2.1 above), since a specific class of consonants—the coronals—is the subject of this investigation.

[13] Wolfram and Schilling-Estes (1998: 282) for example use this term to emphasize the fact that a national variety (US) is meant, and not an English generally spoken on the network.

CHAPTER 3

Classification of coronals

3.1 HISTORY OF THE FEATURE [CORONAL]

Most current research assumes phonetic features to be the minimal units of phonological analysis and sees in them a means of explaining how the sound systems of languages are organized. Although the feature [coronal] as part of a feature system and embedded within a fully-fledged feature theory first appears in the *Sound Pattern of English* (*SPE*), there is some earlier mention of it.

3.1.1 Pre-*SPE* feature systems

Components of speech sounds or features have a long history in the description of sound inventories of languages. As can be read in Whitney's *Sanskrit Grammar* (1889), the earliest mentionings of phonetic features date back to the time of ancient India of about 2500 years ago. Hindu grammarians' interest in the sacred language Sanskrit was aroused by their study of traditional religious texts. The central figure here was the distinguished grammarian Pānini[14], whose prescriptive account of Sanskrit grammar is "cast into the highly artful and difficult form of about four thousand algebraic-formula-like rules" (Whitney 1889: xiii) and developed into *the* authoritative norm of correct Sanskrit speech. To describe the sounds of their language, the Hindu grammarians used precisely defined terms, some of which correspond exactly to the terminology of modern phonetics. Not only did they

[14] The influence which the detailed analyses of Indian grammarians such as Pānini had on modern linguistics is still visible today from the use of Sanskrit terminology; well-known examples are *sandhi* for assimilatory phenomena, or *dvandva* (copulative) and *bahuvrihi* (exocentric) for certain types of compounds. Less well-known is perhaps the fact that Pānini's grammar of Sanskrit can be regarded as a precursor to generative phonology with respect to rule ordering relations: It uses the terms *siddha* (effected) and *asiddha* (not effected) for phonological rules depending on which rule applies first (Joshi & Kiparsky 1979: 223-50).

distinguish vowels and consonants, but also classes of consonants as for instance stops (*sparça*), fricatives (*ūsman*) and semivowels (*antahsthā*), voiced (*ghosavant*) and voiceless (*aghosa*) consonants, and five types of place features: labial (*osthya*), dental (*dantya*), retroflex (*mūrdhanya*), palatal (*tālavya*), and guttural (*kantha*) (Whitney 1889: 13-18). To give an example of how close the ancient labels match modern articulatory terminology, I quote the ancient Hindu grammarians' description of *palatal* from Whitney:

> They [the sounds] are called **tālavya** *palatal*, and declared to be formed against the palate by the middle of the tongue. (1889: 16)

Much more recently, the Prague School phonologist Nikolaj Trubetzkoy in *Grundzüge der Phonologie* (1939) proposed a set of distinctive features in order to describe the functional role of sounds within a linguistic system. He intended these features, e.g. labial, dental, or velar, to spell out the phonologically relevant oppositions between the meaning-distinguishing units of language—the phonemes. Trubetzkoy's description of distinctive features ("distinktive Schalleigenschaften" 1958: 82) is articulatory in nature.[15] However, it is not as systematic and rigidly defined as the later feature theory presented by Jakobson, Fant & Halle (1951) and Jakobson & Halle (1956).

Where Trubetzkoy's focus is clearly on the phoneme, i.e. on a language-specific functional unit, Jakobson, Fant & Halle (1951) construct a rigorous feature theory, which makes the distinctive feature itself the basic unit of phonological analysis.[16] In contrast to Trubetzkoy's features, their set consists of only twelve binary features which are characterized by an acoustic description in terms of spectral energy in formant structure. In addition to these twelve features, there are the three "prosodic" features [force], [quantity], and [tone] (Jakobson & Halle 1956: 29-33). The sum of fifteen features is claimed to universally underlie the "entire lexical and morphological stock" of all languages of the world (Jakobson, Fant & Halle 1951: 40).

---

[15] With respect to consonantal place features, Trubetzkoy makes a distinction between "labial", "apical", and "dorsal", because lips, tongue tip, and tongue body accordingly are the articulators most easily moved to produce an obstruction to the airflow in the oral cavity (1958: 114-5). A coronal feature is not part of his set of place features.

[16] This is a very modern approach, because phonological processes, including different types of assimilation, are better described as involving only parts of a segment, e.g. a certain feature, than the whole segment—a discovery that led to the recognition of natural classes of segments (cf. chapter 4.3.1 below).

(3.1) The set of twelve binary acoustic features (after Jakobson, Fant & Halle 1951: 40)

[vocalic] vs. [nonvocalic]	[consonantal] vs. [nonconsonantal]
[compact] vs. [diffuse]	[tense] vs. [lax]
[voiced] vs. [voiceless]	[nasal] vs. [oral]
[discontinuous] vs. [continuant]	[strident] vs. [mellow]
[checked] vs. [unchecked]	[grave] vs. [acute]
[flat] vs. [plain]	[sharp] vs. [plain]

Although the majority of phonological theories has adopted speaker-oriented articulatory parameters as a basis for phonetic description, an acoustic or perceptual feature characterization defined in terms of the differentiating physical qualities of speech sounds—and hence excluding an articulatory cavity feature such as [coronal] by definition—has notable merits. Just as articulatory features are created to express the functionally relevant distinctions with respect to speech production, acoustic properties like formants are measured not because they are believed to be objects of linguistic reality in their own right, but because they serve as significant cues in human speech perception for discriminating utterances. Jakobson, Fant & Halle state their belief in the primacy of the perceptual dimension for the determination of the distinctive features quite explicitly:

> The closer we are in our investigation to the destination of the message (i.e. its perception by the receiver), the more accurately can we gage the information conveyed by its sound shape. This determines the operational hierarchy of levels of decreasing pertinence: perceptual, aural, acoustical and articulatory (the latter carrying no direct information to the receiver). The systematic exploration of the first two of these levels belongs to the future and is an urgent duty. (1951: 12)

For one thing, a set defined along acoustic and auditory dimensions acknowledges the communicative dimension of language as a code shared among speaker *and* hearer; for another, it allows for the description of vowels *and* consonants with a single set of features: The feature opposition [grave] vs. [acute], for example, distinguishes back vowels and labial and velar consonants with their energy concentrated in the lower frequencies from front vowels and alveolar and palatal consonants with a greater concentration of energy in the upper frequency ranges.[17]

[17] It is a curious fact, which has resulted in a considerable amount of dispute, that the acoustic feature [grave] was replaced by the articulatory feature [coronal] in *SPE*, i.e. [–coronal] = [+grave]

This type of generalization is hardly achieved in articulatory terms. Nevertheless, articulatory terminology is predominant in the literature. This may be due to the fact that the production of speech and the observation of the movements and positions of articulatory organs are felt to be the most basic aspects of speech.

To the best of my knowledge, the term *coronal*[18] *articulation* is first used by the neogrammarian Eduard Sievers in *Grundzüge der Phonetik*:

> Coronale Articulation: die Articulation wird durch den vorderen Zungensaum bewirkt, welcher sich als eine mehr oder weniger scharfe Kante dem Gaumen entgegenstellt. (1901: 59)

Sievers includes sounds produced with the tongue tip (apical articulation) in the class of coronals which contains dental, postdental, alveolar, and "inverted" (i.e. retroflex) sounds (1901: 59-62). In *Language* (1933), Bloomfield presents a very similar taxonomy of sound features, which may well have been inspired by Sievers, since *Grundzüge der Phonetik* (1901) appears in Bloomfield's reference list. Bloomfield clearly distinguishes apical[19] from coronal articulations (1933: 98-9), which however is only a minor divergence from Sievers' set.

3.1.2 The feature [coronal] within the *SPE* model

Chomsky & Halle devise a finite set of 25 phonetic features, plus eight "prosodic features" (1968: 298-300), all of which are defined in articulatory terms.[20] Similar to Jakobson, Fant & Halle (1951), the authors of *SPE* put considerable emphasis on the universal validity of their feature inventory, which they regard as an adequate means for the description of the sound systems of the world's languages.

> The total set of features is identical with the set of phonetic properties that can in principle be controlled in speech; they […] are therefore the same for all languages. (Chomsky & Halle 1968: 294-5)

---

and [+coronal] = [–grave], but with conflicting claims as to the patterning of palatals (cf. section 3.2.2 below).

[18] The class of sounds which is defined by this term was recognized even earlier, e.g. by the French phonetician Paul Passy who identified a class of "linguales" ("formées avec la pointe ou la face de la langue et les dents ou les alvéoles") which comprises dentals, postdentals, alveolars, and postalveolars (1890: 88-9).

[19] In the *SPE* feature system, apical reappears as the distinctive feature [–distributed], while [+distributed] corresponds to a laminal articulation (Chomsky & Halle 1968: 312-4).

[20] On closer examination however, features like [sonorant] and [strident] at least suggest an acoustic or a perceptual foundation: "Strident sounds are marked acoustically by greater noisiness than their nonstrident counterparts." (Chomsky & Halle 1968: 329)

The features [high], [low], and [back] relate to the placement of the dorsum of the tongue during the production of sound segments (cf. also figure 3.3 below). Although they are articulatory in nature, they can be used for the description of vowels and consonants alike, similar to Jakobson et al.'s features [diffuse] vs. [compact] and [grave], which Chomsky & Halle claim to be earlier versions of their new features. According to Chomsky & Halle, Jakobson et al.'s feature set had to be revised and expanded in order to account for the increasing number of different sound systems which had to be described within one single framework (1968: 306).

(3.2) The feature matrix of *SPE* for the English consonant inventory (a selection of features)

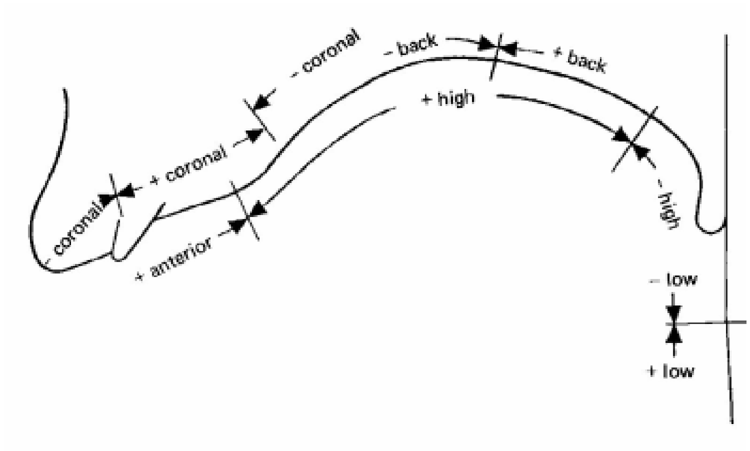
		p	b	t	d	k	g	f	v	θ	ð	s	z	ʃ	ʒ	tʃ	dʒ	m	n	ŋ	l	r	j	w	h	ʔ
Major	[sonorant]	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	+	+	+	+	+	+	–	–
class	[vocalic]	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
features	[cons.]	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	–	–	+	+
Cavity	**[coronal]**	**–**	**–**	**+**	**+**	**–**	**–**	**–**	**–**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**–**	**+**	**–**	**+**	**+**	**–**	**–**	**–**	**–**
features	[anterior]	+	+	+	+	–	–	+	+	+	+	+	+	–	–	–	–	+	+	–	+	+	–	+	–	–
Tongue	[high]	–	–	–	–	+	+	–	–	–	–	–	–	+	+	+	+	–	–	+	–	–	+	+	–	–
body	[low]	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	+
features	[back]	–	–	–	–	+	+	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–	–	+	–	–
Manner	[cont.]	–	–	–	–	–	–	+	+	+	+	+	+	+	+	–	–	–	–	–	+	+	+	+	+	–
features	[del. rel.]	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	+	–	–	–	–	–	–	–	–	–
Source	[voice]	–	+	–	+	–	+	–	+	–	+	–	+	–	+	–	+	+	+	+	+	+	+	+	–	–
features	[strident]	–	–	–	–	–	–	–	–	–	–	+	+	+	+	+	+	–	–	–	–	–	–	–	–	–
Prosodic	[stress]																									
features	[pitch]																									

One of *SPE*'s 33 universal features is [coronal]; its concise definition reads as follows (cf. also figure 3.3 below):

> Coronal sounds are produced with the blade of the tongue raised from its neutral position; noncoronal sounds are produced with the blade of the tongue in the neutral position. (Chomsky & Halle 1968: 304)

It has since been remarked in the literature (Ladefoged 1980: 491; Lass 1984: 84) that this definition is phonetically questionable. Evidently, the tongue blade is not at rest in its neutral position during the production of a velar /k/, for example, which is marked [–coronal], since the constriction formed by the body of the tongue (dorsum) *does* have an influence on the adjoining blade (lamina). An articulatory feature characterization—even as sketchy as in *SPE*—is more in line with the speaker-based approach of generative grammar than an acoustically or perceptually determined set.[21]

(3.3) A representation of binary *SPE* features in terms of an articulatory diagram focusing on [coronal] (after Ladefoged 1971: 101)



SPE features are termed *phonetic*, when in fact they do not necessarily correspond to straightforward phonetic properties and are too abstract in nature, as I will show below. This problem was recognized early on by Chomsky & Halle themselves in the epilogue of *SPE* and addressed in a section with the title "*some unresolved problems*":

> The problem is that our approach to features, to rules, and to evaluation has been overly formal. […] In particular, we have not made any use of the fact that the features have intrinsic content. (1968: 400)

It is of course common knowledge that any attempt at structuring the continua of sound waves and articulatory movements by setting up linguistic categories and classifying phenomena according to these categories constitutes an abstraction.

> Since the sound matter of language is a matter organized and formed to serve as a semiotic instrument, not only the significative function of the distinctive features but even their phonic essence is a cultural artifact. […] Where nature presents nothing but an indefinite number of contingent varieties, the

[21] For a phonetically more detailed description of the coronal articulators see section 3.2 below.

> intervention of culture extracts pairs of opposite terms. The gross sound matter knows no oppositions. (Jakobson 1949: 321)

There is justification for abstraction, however, if a class of sounds, defined by the feature [coronal] for example, can be shown to have certain functional characteristics in common with respect to patterns of distribution or assimilatory phenomena.

However, what makes the definition of *SPE* features vague is their double function: a manner feature like [tense], for instance, is both a universal descriptive label referring to articulatory properties (tenseness of supraglottal musculature) and a language-specific classificatory label for identifying a functional class of vowels in the phonology of English. In this function, the feature system defines the natural classes of a specific sound inventory and provides a basis for characterizing and evaluating the different phonological processes of the language in question. It has been a controversial issue as to what extent a single set of features is adequate for both levels of description: the phonetic substance and the phonological form. For a more detailed investigation of this question, it is necessary to give an outline of the function of features within the *SPE* model.

The input to the phonological component of the grammar (cf. figure 3.4 below) are structurally analyzed strings of segments called "formatives" (i.e. lexical entries, e.g. morphemes; Chomsky & Halle 1968: 164) which are provided by the lexicon.[22] Formatives consist of segments, and the segments in turn are regarded as structureless bundles[23] of "phonological features". These features, which represent the phonologically relevant oppositions of a certain lexical entry, are strictly binary if they supply nonredundant information, otherwise they are marked with a zero.[24] Thus the features subclassify the sound inventory into e.g. coronals and noncoronals, anteriors and nonanteriors etc., which are claimed to be natural classes in the sense that they are phonetically motivated. This claim was subject to immediate criticism, and in order to formalize the notion of phonetic naturalness, a set of universal marking conventions was invoked in the epilogue of *SPE*. The theory of markedness (cf. chapter 4.2 below) recognized the necessity of differentiating "expected" and

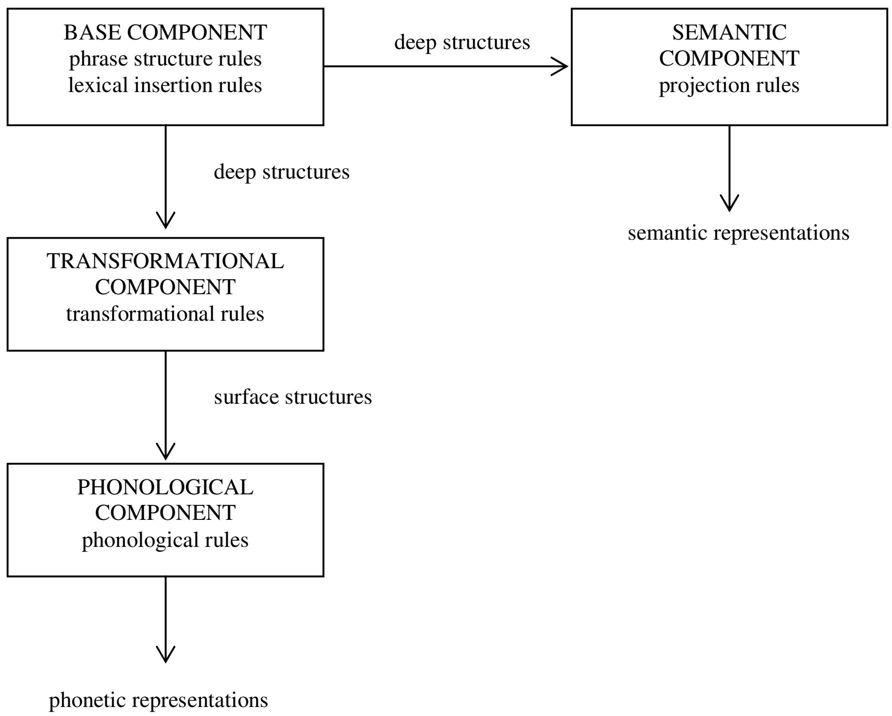
[22] Within the 1968 *SPE* model, the lexicon is still an integral part of the syntactic component of the grammar (Chomsky & Halle 1968: 164).

[23] In *SPE*, features are grouped under the five headings "major class", "cavity", "manner", "source", and "prosodic features" (cf. figure 3.2 above), but this has no bearing yet on the structure of the segment itself: "This subdivision of features is made primarily for purposes of exposition and has little theoretical basis at present." (Chomsky & Halle 1968: 300)

[24] Other types of features contained in a lexical entry are e.g. syntactic and semantic features.

"natural" cases of feature configurations and rules from "unexpected" and "unnatural" ones and aimed at accommodating the intrinsic content of features by grounding them in phonetics. This goal was to be achieved through labeling the former "unmarked", while the latter were labeled "marked" (Chomsky & Halle 1968: 402), but as it turned out, the *SPE* marking conventions were not derived from independent phonetic principles either (cf. discussion in chapter 4.2 below).

(3.4) A generative model of grammar (standard theory type)



In the phonological component of the grammar, there is a set of rules, termed lexical redundancy rules[25] by Chomsky & Halle (1968: 171), which operate in a structure building mode on the underlying representations of lexical entries (cf. figure 3.5 below) and provide the redundant feature specifications in the shape of scalar "phonetic features", i.e. not necessarily binary features.

[25] Lexical redundancy rules also have the function to state well-formedness conditions on underlying representations of lexical entries (cf. chapter 4.1 below).

(3.5) *SPE*-type underlying representation of <cat> with phonological segments as unstructured complexes of distinctive features

k	æ	t
+ consonantal	– consonantal	+ consonantal
– sonorant	+ sonorant	– sonorant
– coronal	0 coronal	+ coronal
– anterior	0 anterior	+ anterior
– continuant	0 continuant	– continuant
– voice	0 voice	– voice
0 low	+ low	0 low
0 back	– back	0 back
0 ATR	0 ATR	0 ATR

These phonetic features may have numerous possible gradings, each of which is expressed by an integer in the phonetic specification (cf. figure 3.6 below). The specific context in which the segment occurs supposedly determines the exact integral value, e.g. [5 low] for the vowel /æ/ in <cat>, but mostly *SPE* recognizes only "two positions along a phonetic scale, in which case we may use the symbols + and – instead of integers to indicate phonetic values" (Chomsky & Halle 1968: 164).

(3.6) *SPE*-type surface representation of <cat> with fully specified phonetic segments

k^h	æ	t
+ consonantal	– consonantal	+ consonantal
– sonorant	+ sonorant	– sonorant
– coronal	– coronal	+ coronal
– anterior	– anterior	+ anterior
– continuant	+ continuant	– continuant
– voice	+ voice	– voice
– low	+ low	– low
+ back	– back	– back
– ATR	– ATR	– ATR

Phonetic representations (cf. figure 3.6 above) constitute the output of the phonological component and are automatically converted into articulated utterances by the phonetic component, i.e. by universal phonetic rules which in principle can interpret a phonetic representation of any language. This predictable component is technically not part of the grammar within the *SPE* model. Consequently, certain properties of the signal, for example articulation, timing, and coarticulation, belong to an extragrammatical phonetics (Chomsky & Halle 1968: 295). In this respect, the traditional view that phonetics is outside the domain of linguistics is supported by Chomsky & Halle.

Three years after the publication of *SPE*, Ladefoged proposed a modified feature system of 26 articulatory and auditory features, 20 of which are strictly binary, the rest used in a nonbinary mode. Two innovations are worth mentioning here: the feature [coronal] is substituted for by the binary feature [apical] in order to distinguish between sounds made with the tongue tip and sounds made with the tongue blade (1971: 44). More importantly, Ladefoged suggested a multivalued feature [articulatory place] which contains several independent and equidistant targets located on a place continuum along the oral cavity (1971: 91-4). Ladefoged's arguments for a nonbinary feature [articulatory place] are the same as for the feature [vowel height]: both are regarded as scalar rather than binary features. These arguments for a multivalued place feature already point in the direction of more modern nonlinear representations of place.

Before I turn to recent feature theories and the status of coronals within these approaches in chapter 4, I would like to give a brief outline of the physiological basis for the production of coronals, which has become increasingly relevant to modern phonological theories. For this reason I have decided not to discuss phonetic matters and phonological matters separately, but instead I have chosen to refer to phonetics where it is necessary for phonological description.

3.2 THE CORONAL ARTICULATORS

In view of the numerous possible coronal articulations, it seems sensible not to speak of *the* coronal articulator, but instead to describe the range of different locations in the oral cavity at which the articulatory events involved in coronal production take place. It is further necessary to distinguish movable articulators on the one hand from articulatory targets on the other. For English, the interplay of active and passive articulators results in 13 consonantal segments commonly defined as coronals.

(3.7) Inventory of English coronals

CORONALS	θ ð t d s z ʃ ʒ ʧ ʤ n l r

The multitude of distinct sounds produced with the coronal articulators at different places and in different manners is doubtlessly grounded in human physiology:

> The tongue is the most versatile and mobile of all the organs of speech, and gains its speech capacity for agile and precise positioning from its primary biological role in moving food around the mouth and pharynx during chewing and swallowing. (Laver 1994: 121)

In a study on the phonetic universals of sound systems, Lindblom (1986) not only attributes a great versatility to the coronal articulators, but he also states that sounds produced in the anterior part of the oral cavity are auditorily more salient. Therefore, the articulatory space in the front can be exploited more efficiently than in the back part of the oral cavity, since smaller contrasts are easily perceived, which then accounts for the large number of contrastive coronal (and labial) segments.

> (1) Articulators have greater mobility at the front of the mouth (e.g. lips, tongue tip). (2) There appears to be a richer supply of structures for sensory control at anterior vocal tract locations […]. (3) Acoustic-perceptual effects are greater at the front than at the back, given geometrically comparable articulatory perturbations and conditions typical of, for example, voiceless consonants […]. (1986: 39)

I will specify below what is traditionally dealt with under the heading 'place of articulation of coronals' in the literature. More precisely, it is important to determine what moves and in which direction it moves.

3.2.1 Movable articulators

In phonetics, there is a long tradition of describing the vocal organs responsible for speech production. Interestingly as early as 1899, Soames identifies three distinct articulators in the oral cavity, which differs from modern classifications only with respect to the terminology used: *dental* and *guttural* for coronal and dorsal.

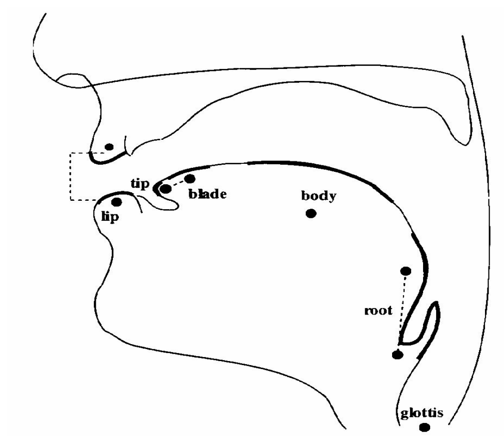
> It is by the movements of the lips, tongue and soft palate that the various vowels and consonants are formed, […]. (1899: 12)
> In the lip stops **p** and **b** [the breath] is stopped by closing the lips, in the point stops **t** and **d**, by the point of the tongue touching the upper gums, and in the

> back stops **k** and **g**, by the back of the tongue touching the soft palate. These three classes of consonants are sometimes called *labial*, *dental* and *guttural*. (1899: 35)

For phonetic description, the active movable articulator *tongue* in its entirety is commonly divided into the four areas (sometimes five areas, see Pike below) tongue tip (*apex*), tongue blade (*lamina*), tongue back (*dorsum*), and tongue root (*radix*), although there seems to be no anatomical justification for having this subdivision rather than some other (Hardcastle 1976: 92). Similarly, Pike states that

> the tongue […] is divided into several articulators; since various parts may cause strictures separately, tip, blade, middle, back and root are convenient arbitrary points of reference for these positions […]. (1943: 121)

(3.8) The five groups of moveable structures forming the active articulators in the vocal tract corresponding to the following five gestures: LABIAL, CORONAL, DORSAL, RADICAL, and GLOTTAL (after Ladefoged & Maddieson 1996: 12).

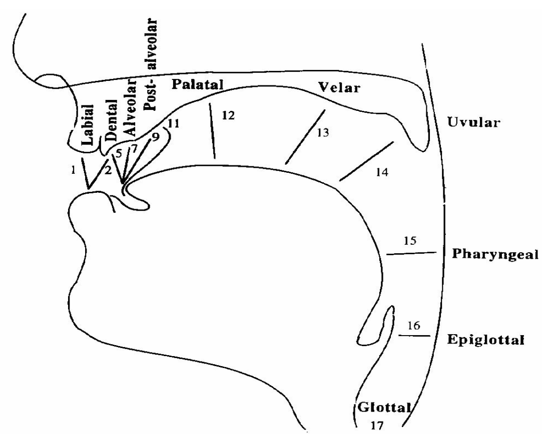


The tip and blade of the tongue can in principle be moved separately, but for reasons of physical proximity, they are regarded as a functional unit and together form the coronal gesture. "These two parts of the tongue […] are so closely connected only one or the other can be the primary articulator in any given speech sound." (Ladefoged & Maddieson 1996: 10) The same is true of the tongue body (cf. figure 3.8 above), the middle and back part of the tongue in traditional terms, which merge to yield the dorsal gesture (1996: 11).

3.2.2 Articulatory targets

Connected speech is a dynamic process with the active articulators in continuous motion from one targeted position to the next in order to produce phonetic contrasts. The starting point of the movement for the articulation of a consonant not only depends on the prescribed articulator setting for the specific sound but also on the vocal tract configuration needed for the previous and sometimes for the following sound or sounds. For the purpose of description however, the movements for specific gestures are thought to be performed from a neutral position of the vocal tract towards an articulatory target. Targets are to be viewed as abstract goals which are aimed at by the movable articulators, but not necessarily hit (Ladefoged & Maddieson 1996: 6), one crucial factor being certainly the speech rate.

(3.9) The nine regions of the vocal tract considered as target areas for the moveable articulators. The numbered lines show some of the 17 named articulatory gestures, including those in the glottal region (after Ladefoged & Maddieson 1996: 12).



The reference points must of course be arbitrary because the tongue is a continuous body without relevant anatomical landmarks. The tongue blade including the tongue tip is taken to be the part of the tongue which is not attached to the floor of the mouth (Ladefoged & Maddieson 1996: 10) and varies in measurement from 1.5 cm (for anterior coronals) up to 4 cm (for nonanterior coronals) due to the extraordinary elasticity and extensibility of the tongue muscles (Keating 1991: 31).[26]

[26] It must be noted, however, that Keating is one of the linguists who include palatals in the class of coronals. Therefore her estimate of the blade will necessarily turn out to be more than the usual conservative measurements of about 2 cm by other linguists (e.g. Catford 1977: 143; Catford 1988: 82).

Targets for the tip and blade of the tongue are places in the oral cavity which are located in the area of the upper teeth and the roof of the mouth. The consonant sounds produced in this anterior part of the vocal tract are frequently classified in terms of the following passive articulators: dental, alveolar, post-alveolar, and retroflex. Thus, when the traditional term *alveolar*, for example, is used, what is actually implied is a movement *of* the tongue tip or blade *to* the alveolar ridge.

In post-*SPE* phonology, the palatal sounds (relevant for English is merely the glide /j/) are a matter of controversy: some linguists (e.g. Kenstowicz 1994: 31; Hammond 1999: 5) have classified palatals as a species of coronals. Keating for instance does this on phonetic grounds with reference to their production:

> Halle and Stevens (1979) proposed a redefinition of coronal to mean the blade or front of the tongue so as to include the palatals. However, this move seems unnecessary, as palatals generally do involve the blade proper, in addition to the front of the tongue. (1991: 37)

Many other linguists (e.g. Jones 1960: 209; Catford 1977: 162; Lass 1984: 94-6; Catford 1988: 121; Carr 1993: 306; Laver 1994: 297; Ladefoged 1997: 597) regard them as dorsals, produced with the tongue body.

> When places of articulation are grouped according to the active articulator used, palatal articulations, which use the body of the tongue rather than the blade, fall outside the Coronal class of articulations. […] We use the term Dorsal for this group. (Ladefoged & Maddieson 1996: 31)

Hall even devotes a whole section of his book on coronals to this widely debated issue and he comes to yet another conclusion. According to Hall, true palatals like [ç], for instance, should not belong in the class of coronals, but some sounds (among these the stop [c], the nasal [ɲ], and the glide [j]), which likewise have traditionally been described as palatals (cf. the IPA chart 1999: ix), are reclassified as "alveolopalatals" by Hall (1997: 6-22). Further, since alveolopalatals are coronals by definition, reclassification causes these sounds to resurface as coronal sounds. I find this to be an ad hoc conclusion since Hall presents no particularly compelling phonological evidence for the alleged coronality of [j]. He argues from the data of only two languages (Feʔfeʔ-Bamileke and Ewe) "that [j] is a [+coronal] sound because it patterns with alveolars" (Hall 1997: 21). With regard to phonetics, he concedes that "phonetic evidence for this classification is unfortunately lacking" (ibid.). In fact, palatograms present clear evidence for classifying the glide [j] as a palatal sound (e.g. Jones 1960: 67, 209). I therefore follow the IPA description of the sound (for American English IPA 1999: 41) as a palatal and accept Ladefoged & Maddieson's inclusion of the sound in the class of dorsals. Considering the targets for [j] and [ç], I assume the places to be either identical or, especially in the case when [j] is followed by a [+low, +back] vowel as in <yard>, the target for [j] may be located even farther back in the mouth than for [ç], which Hall classifies as a true palatal sound (1997: 19).

I will show in chapter 4.3 below that from a phonological point of view, it is indeed useful to group some of the sounds produced at distinct places in one class because of the linguistic behavior these sounds exhibit. Thereby, the number of linguistically relevant targets can be greatly reduced. Prior to this, I will point to some evidence that consonants in which the tongue tip and blade are involved as active articulators, as well as consonants in which dental, alveolar, post-alveolar, and retroflex targets are involved as passive articulators, should be classed as coronal consonants. This is also justified from a phonetic point of view.

3.3 PHONETIC EVIDENCE FOR CORONALS

With respect to coronals, it is indisputable that there are universals to be observed. Due to the fact that language is spoken, it does not seem unnatural to assume that languages evolve sound systems which can be explained, at least in part, phonetically. Spoken language uses the vocal-auditory channel and should therefore be closely linked to the developmental and adult mechanisms of speech production and speech perception. This is commonly discussed under the notion 'functional load' of an explanation (e. g. Martinet 1968: 478-9; cf. also chapter 5.4 for a more in depth discussion of functional approaches). It is impossible in this study to provide an encyclopedic overview of phonetic phenomena that are of relevance for the study of coronals. For this reason, I have selected three broad areas of material evidence which, in my view, contribute to the general issue of creating a solid physical basis for phonological categories. One major object for phonetics as the study of speech is certainly its production by the human articulatory system. This is dealt with in section 3.3.3 on the theory of Articulatory Phonology developed by Browman & Goldstein. In the two following sections, I will present typological evidence for the class of coronals (3.3.1) as well as the arguments of the theory of Acoustic Invariance by Blumstein and Stevens (3.3.2), who have made some successful attempts at establishing a possible isomorphism between distinctive features like for instance [coronal] and certain physical properties of speech.

3.3.1 Evidence from typology

Similar to the notion that tongue tip and blade function as a unit to form the coronal gesture (cf. section 3.2.1 above), there is crosslinguistic phonetic evidence gathered by Maddieson that consonants produced in the traditional dental and alveolar regions should be grouped in one superordinate class of sounds "because of the frequent uncertainty as to whether they are dental or alveolar" (1984: 10). Maddieson evaluated the size and structure of the phonological inventories of 317 languages—a corpus known as the *UCLA Phonological Segment Inventory Database (UPSID)*. Specifically with respect to the segments /t/, /d/, /s/, /n/, /l/, /r/ (all dental or alveolar[27]) Maddieson suggests that "a certain amount of 'pooling' of similar segments" is inevitable due to the difficulties of determining the exact place of articulation (1984: 12). An observation made by Daniel Jones with respect to cardinal consonants may shed some light on Maddieson's dilemma.

Jones is well-known for establishing the principle of cardinal sounds in the eight cardinal vowels (1960: 31-41), which serve as reference points for the description of language-specific vowel inventories. For the description of consonants, he states, the principle is commonly not needed because they can be acquired and described without reference to cardinal positions; instead it is sufficient to know the required articulator configurations for a specific consonant in order to produce it correctly (1960: 43). However, there are some exceptions:

> It [the principle] is […] applicable in cases where it is possible to pass from one sound to another through a series of sounds each of which is hardly distinguishable from the preceding, and where it is therefore necessary to fix arbitrarily some points to which any given sound in the series may be referred.
> (Jones 1960: 42)

Interestingly, Jones finds the principle of cardinality applicable to all sounds that are articulated by the tongue against the roof of the mouth, and his examples are a stop

[27] It should be noted that the term *coronal* is not yet used in Maddieson (1984).

series ranging from dental [t̪] to alveolar [t] to retroflex [ʈ], and a fricative series ranging from dental [θ] to alveolar [s] to post-alveolar [ʃ]. What these sounds have in common is of course that they are all coronals.

If it is possible to demonstrate that not only dental and alveolar but all three (or four) separate targets for the active coronal articulators (cf. figure 3.9 above) function as a unit, then these targets can be analyzed as coronal targets which, in coordination with the active coronal articulators, produce coronal sound segments. Maddieson's "pooling" and the establishment of a class of coronals would then be justified on phonetic grounds.[28]

Evidence in support of a class of coronals was supplied by Ladefoged & Maddieson (1988; 1996). They argue on the basis of vast typological data drawn from nearly 400 languages that the 17 places of articulation commonly identified as individual articulatory gestures (cf. figure 3.9 above) should not be regarded as discrete categories. Instead they should be combined to form five major gestures (cf. figure 3.10 below), represented as nonterminal nodes in a feature geometry (cf. chapter 4.3.2 below). They corroborate their claim with the observation that the languages investigated tend to make use of only the major articulatory categories—more precisely the first two columns of figure 3.10 below—for the purpose of segmental contrast. The exact specification of articulation in terms of the individual places is then generally handled by language-specific, lower-level phonetic rules (Ladefoged & Maddieson 1988: 59).

In addition, Ladefoged & Maddieson substantiate their claim with the observation that a very high proportion of the world's languages place a restriction on consonant segments produced in the same manner: they may not be drawn from the same active articulator class. A frequently occurring inventory of voiceless plosives is the one found in English, /p/, /t/, and /k/, while a series like /c/, /k/, and /q/ is unlikely to occur in any language (1996: 43).

> The five distinct types of articulatory gestures based on independence of articulators can be regarded as establishing a set of major place features. Each of the larger number of individual places we have been discussing can be grouped under one of these major place features, […]. (ibid.)

[28] Stevens & Keyser, who use Maddieson's (1984) database for an empirical study on the enhancement of primary features in consonants—among these [coronal]—arrive at a similar way of classifying coronal sounds (1989: 81).

(3.10) Relationship between the 5 major gestures and the 17 individual places of articulation (after Ladefoged & Maddieson 1996: 44).

LABIAL		1.	bilabial
		2.	labiodental
CORONAL	laminal	3.	linguo-labial
		4.	interdental
		5.	laminal dental
		6.	laminal alveolar
		7.	laminal post-alveolar (palato-alveolar)
	apical	8.	apical dental
		9.	apical alveolar
		10.	apical post-alveolar
	sub-apical	11.	sub-apical palatal (retroflex)
DORSAL		12.	palatal
		13.	velar
		14.	uvular
RADICAL		15.	pharyngeal
		16.	epiglottal
LARYNGEAL		17.	glottal

3.3.2 Evidence from acoustics and perception

More evidence for phonetically real coronals comes from the theory of Acoustic Invariance as presented in Stevens & Blumstein (1981). The theory is guided by the claims that there are invariant acoustic correlates of phonetic features of natural language and that the human perceptual system is sensitive to these acoustic properties. Ideally the theory serves as an interface between models of speech processing and issues of phonological theory in that it supplies physical data which support the categories necessary for phonological investigation. In this way, the latent danger of drifting into ever more abstract representations of phonological units, which much of generative phonology was faced with, can be evaded by grounding phonology in phonetics, i.e. in physical facts.

With respect to place of articulation in stop consonants, Blumstein (1986) presents research conducted in the Brown University phonetics laboratory. For these consonants, a so-called "dynamic relative invariance" for place is postulated (1986:

182). Labials and coronals are explicitly contrasted, showing that the distinctive property in these two classes of sounds is indeed the traditional acoustic feature [grave] as first established in Jakobson's feature theory (cf. section 3.1.1 above). Blumstein confirms Jakobson's analysis that [+grave] labials exhibit a sustained spectral energy in the low frequencies around 1500 Hz, while [–grave] coronals are characterized by a sustained energy in the high frequencies around 3500 Hz.

> In essence, there are two invariant properties characteristic of the class of labial and coronal consonants. These two classes of sounds are distinguished by a feature of gravity; i.e. the relative distribution of energy in high and low frequencies. (1986: 182)

In his recent work on acoustic phonetics, Stevens identifies three active articulators responsible for producing phonetic contrasts in the oral cavity: lips (LABIAL), tongue blade (CORONAL), and tongue body (DORSAL). He states that

> one of these articulators is always active in forming the constriction in the oral cavity for [+consonantal] segments, and it is the movement of this articulator that creates the consonantal landmark. (1998: 249-50)

Consonantal landmarks are defined as the approximations or actual constrictions or release movements of a particular articulator with respect to its targeted goal. Acoustically, they are characterized by a decrease in the first-formant frequency, which is reported to be perceptually salient. When the coronal articulator, for example, is moved to the alveolar ridge to produce a constriction, this adjustment will give rise to an acoustic output which is distinguishable when compared to the output of the labial and dorsal articulators (Stevens 1994: 242-57; Stevens 1997: 490-505; Stevens 1998: 243-55).

Ladefoged on the other hand, has supported an opposing view. He has been critical of the idea that the relationship between phonologically motivated distinctive features and phonetic parameters could be one-to-one. A distinctive feature represents a parameter which has corresponding properties in acoustic, articulatory, and auditory domains, but Ladefoged (1980: 498-9) argues that these correlates are by no means invariant across languages, nor even within a single language. The distinctive feature, then, can be regarded as a "linguistic prime" (1980: 490), an archetypal contrast between elements of sound on an abstract level, which can be realized in a variety of ways.[29] For the feature [coronal] for example, he claims two parameters to be responsible—both the raising *and* the fronting of the tongue tip—and consequently "there is no way that Coronal (or Alveolar) can be interpreted in terms of a single physical scale" (1980: 491).

Lindau & Ladefoged explicitly refer to the work done by Blumstein and Stevens and call into question the view that the human perceptional system is fitted with special feature detectors. They deny the possibility that there may be a single invariant physical property that constitutes a phonetic correlate of a particular phonological feature:

> [X]-ray studies show that there is a continuum of possible places of articulation covering much of the front of the oral cavity. Different languages use different points within this continuum, much as they use different points in the vowel space. It is not possible to find discrete places of articulation such as dental, alveolar, and retroflex. (Lindau & Ladefoged 1986: 468-9)

Lindau & Ladefoged's view is thus more in line with one major position of American stucturalism in the tradition of Sapir[30] that out of an incalculable mass of possible speech sounds, individual languages arbitrarily select their phoneme inventories. Any apparent invariance of acoustic characteristics in speech sounds can only be due to perceptual abstraction on the part of the native speakers of the language, who are moreover, totally oblivious of the nondistinctive qualities of their native language (Gleason 1961: 257-70).

Promising as the results of Blumstein and Stevens may seem, for the time being the identification of places of articulation of obstruents with their method is restricted to the onset position of a CV syllable (Stevens 1997: 505). With Lindau and Ladefoged's reservations in mind, it remains doubtful whether further research in acoustics and further development of the accompanying technology will eventually lead to the ambitious goal of identifying the physical correlates of all relevant phonological categories, and more specifically, whether it will lead to the discovery

---

[29] In some way this distinction resembles the distinction between phoneme and allophone: one might use the term *allofeature*.

[30] Sapir speaks of an "ideal" phonetic system in the sense of having a mental reality which corresponds only indirectly to actual physical facts (1921: 55-6). An opposing view can be found in Bloomfield's definition of phonological units (phonemes) as "minimum unit of distinctive sound-feature". On this view, phonemes correspond to invariant acoustic properties which can be identified in the phonetic signal (1933: 79-80). Twaddell in turn refers to Sapir and Bloomfield and attacks both positions: He suggests that phonemes are a simple fiction (1935: 67-75).

of one distinctive feature set which then can be identified with an independently established set of phonetic capabilities of the human organism.

3.3.3 Evidence from articulation

From the late 1980s on, Keating (1988: 284) notes a shift of interest from acoustic and perceptual matters back to articulation, as phoneticians focus their attention on articulatory gestures as potentially basic units of phonetic *and* phonological description. One important approach in this area of research is the theory of Articulatory Phonology (AP), which was developed by Browman & Goldstein at Haskins Laboratories (1986; 1989; 1990; 1992) and addresses the recurring issue of bridging the gap between the physical and linguistic structure of speech from an articulatory perspective. Browman & Goldstein's model of speech production is appealing for several reasons, as I will show in the following paragraphs.

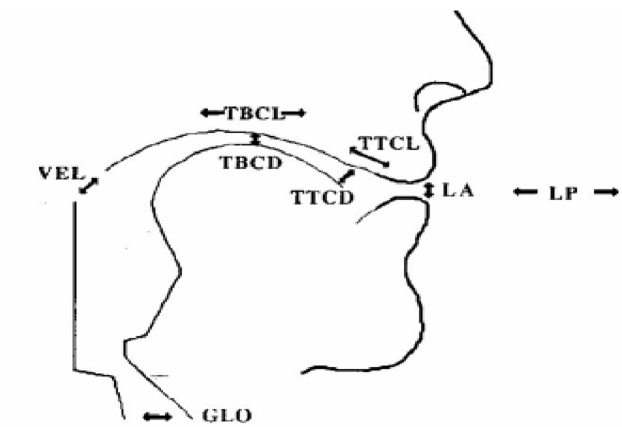
AP takes gestures to be "the basic atoms of phonological structures" (Browman & Goldstein 1989: 201) and conceptualizes them as dynamically modeled movements of the vocal tract articulators towards a range of specified locations whereby constrictions are formed and released. In the oral cavity, constrictions can be formed by the lips (L), the tongue tip or blade (TT), and the tongue body (TB). The articulators which are of relevance according to AP clearly correspond to the familiar classification of labial, coronal, and dorsal articulators respectively. Thus further evidence for the salience of coronal articulations is supplied by AP.

Oral gestures—including the TT articulator, which is of special interest here—involve two tract variables: a horizontal dimension concerning the location where the constriction is formed and released and a vertical dimension concerning the degree of the constriction (cf. figure 3.11 below). AP is a truly dynamic model of articulation, which meets the requirements mentioned in section 3.3.1, in that it subsumes both the active articulator, e.g. the coronal articulator tongue tip, and the targets that can be reached by it under one gesture, which is at the same time the basic unit of phonological investigation.

Browman & Goldstein point out that gestures neither correspond to traditional *SPE*-type segments nor to the traditional phonological features, since there is no one-to-one correspondence between gestures and either unit (1986: 225).

(3.11) Tract variables and contributing articulators of gestural model (after Browman & Goldstein 1989: 207).

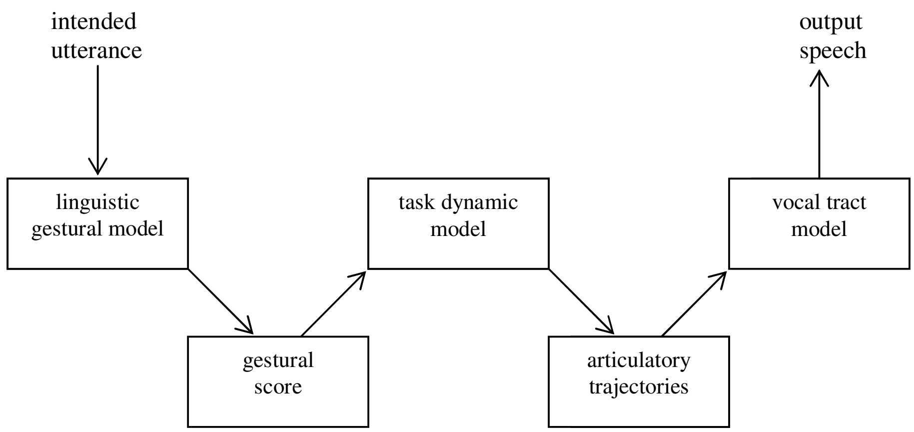
tract variables		articulators
LP	lip protrusion	upper & lower lips, jaw
LA	lip aperture	upper & lower lips, jaw
TTCL	tongue tip constriction location	tongue tip, body, jaw
TTCD	tongue tip constriction degree	tongue tip, body, jaw
TBCL	tongue body constriction location	tongue body, jaw
TBCD	tongue body constriction degree	tongue body, jaw
VEL	velic aperture	velum
GLO	glottal aperture	glottis



A segment /p/ for instance contains at least two gestures, a bilabial closure gesture and a glottal opening gesture, whereas a bilabial closure gesture is traditionally a bundle of several features ([–sonorant], [–coronal], [+anterior], [–continuant], etc.). However, returning to the gestures formed by the TT articulator sets (TTCL and TTCD), my claim is that it is these gestures which produce the class of sounds known as coronals.

When an utterance, e.g. /tæn/, is intended, gestures are organized in the *gestural model* to form a *gestural score* (cf. figure 3.12 below). For the initial oral closure gesture, two paired tract variables need to be specified, one for the constriction location and one controlling the constriction degree (in this case of the TT articulator sets); these two then result in the functional goal for the particular gesture. When such a dynamic system is to be imposed on the vocal tract, the *task dynamic model* calculates the *articulatory trajectories* of the individual articulators involved in this coordinative structure. Articulatory trajectories serve as the input to the *vocal tract model*, which then is responsible for the vocal tract shape required for the production of the actual output speech.

(3.12) AP gestural model (after Browman & Goldstein 1992: 160)



A further substantial advantage over older models which posit nonoverlapping discrete segments (e.g. *SPE*) is the fact that AP gestures may be subject to overlap. It is assumed that each gesture possesses an inherent spatial aspect as well as an intrinsic temporal aspect, which allows for gestural overlap in time and space. *SPE*-type segments which are viewed as linearly ordered feature bundles (cf. section 3.1.2 above) have therefore been attacked by linguists of very different theoretical orientation (e.g. Anderson 1974; Goldsmith 1976; Hooper 1976; Kahn 1976; Clements & Keyser 1983) and also by Browman & Goldstein (1986). AP challenges both the linearity and discreteness assumptions in that it accounts for the continuous nature of speech, where steady states are absent, with gestures that are allowed to overlap along imaginary vertical lines and that are ordered on different horizontal levels. Rather than assuming P-rules (*SPE*) that convert phonological representations into phonetic representations, Browman & Goldstein propose to ground phonological representations in explicit descriptions of articulatory actions in space and over time. As a phonological representation, an approach based on gestures is in line with nonlinear representations of phonological structure where overlapping gestures are similar to the autosegments in autosegmental phonology. This type of representation will be presented in chapter 4 below.

The basic units of AP provide an articulatory foundation for natural classes. Moreover, changes in the patterns of overlap between neighboring gestural units can account for different phonological alternations, e.g. the kinds of acoustic variation that have traditionally been described as allophonic variation, fluent speech alternations where the amount of overlap increases as the speech tempo increases, coarticulation, speech errors as well as phonological development in language acquisition. In the following chapters, I will demonstrate the usefulness of AP to OT.

CHAPTER 4

Theoretical foundations of OT

4.1 THE COMPONENTS OF AN OT GRAMMAR

Optimality Theory (OT) is a recent theory of grammar (Prince & Smolensky 1993; McCarthy & Prince 1993) introduced in response to major problems with preceding models of generative phonology (GP). Within a decade, OT has gained the attention of linguists all over the world. The widespread dissemination of OT ideas is in part due to the Rutgers Optimality Archive (ROA), an electronic archive which collects works on OT (papers and dissertations) and is freely accessible on the World Wide Web.[31]

Although new, OT is still a theory of generative linguistics (Archangeli 1997: 1) in that it shares with its predecessor models the assumption of an innate Universal Grammar (UG) and a quest for universal principles which are common to all languages. From the set of universal principles expressing universal properties of the subsystems of sounds, words, phrases, and meanings, the grammars of individual languages draw their specific options which then account for the variation among languages.

4.1.1 Basic assumptions

A crucial difference from other models of generative theory (GT) is, however, that OT is formulated as a theory of constraints on the well-formedness of phonological, morphological, and recently also syntactic representations, rather than as a theory of

[31] ROA started collecting material on OT in 1993 on an ftp-server, which accumulated to over 600 papers by the fall of 2003. By ROA's own definition, electronic archiving is not a form of publication; it solely serves as a distribution point for research in OT. When I cite papers from ROA, their exact location in the archive will be given in my references list.

rules for the construction of representations. OT thus shifts the focus from the input (formerly underlying representation) to the output (formerly surface representation) of the grammar, and shows an almost total disregard for the basis[32] of these representations.

A second important difference is that OT's principles (the universal set of constraints) can be violated. The constraints are assumed to be literally universal, in the sense that the total set is present in every grammar. A violation of some principle (constraint) is seen as the result of the interaction of the constraints, i.e. of their hierarchical ordering or ranking. Importantly, a given input-output pair is well-formed *not* because it satisfies every constraint (without violating any), but because no other pair does better on the constraints (minimal violation): this pair is then the *optimal* of all potential input-output mappings (Prince & Smolensky 1993: 2). The differences in the ranking of the universal constraints produce the grammars of individual languages and potentially also varieties of a single language, i.e. accents and dialects (Golston & Wiese 1995; Cassimjee & Kisseberth 1998; 1999; Moren 1999).[33] This means that interlinguistic variation does not lie in the constraints (principles) themselves but in their overall interaction potential. Conversely, other models of GT are characterized by a set of inviolable principles which can take a range of different forms in different languages, i.e. parametric variation. It must be noted that parameters are either set or nonset (e.g. Haegeman 1994: 12-16). On this view, interlinguistic variation lies in the principles themselves.

For phonology, the postulation of very general inviolable principles has crucial consequences: whenever one of these principles appears to be violated in some output form, an intermediate level of representation—intermediate between underlying and surface representation—is set up at which the principle is satisfied. A typical example of such a model is Lexical Phonology (Kiparsky 1982; Kaisse & Shaw 1985; Mohanan 1986). In this theory, the lexicon is the site of a large number of consecutive derivations, effected by phonological rules interacting with morphological rules. Various intermediate levels of representation constitute the output of various lexical strata. For example, the suffixation of <-ity> to an adjective

[32] The input has an extremely vague definition under OT (cf. section 4.1.2 below).

[33] To my knowledge, Halle (1962) originally offered an explanation for different dialects of one language in terms of rule ordering. Instead of examining the data directly, he proposed to focus on the grammars of the dialects as ordered statements. He suggested that the grammars of American and Canadian varieties, for example, share the same rules, which are merely applied in a different order (342-4). It seems sensible to appropriate Halle's approach for a constraint-based model.

like <serene> (cf. section 4.2.2 below) is described as a serial derivation, where phonological rules apply in tandem with morphological rules, i.e. stress assignment, affixation, trisyllabic laxing, and finally flapping (only American English) as a postlexical rule (Mohanan 1986: 10).

(4.1) A derivational grammar as an input-output mechanism

underlying representation		input to rule 1
	rule 1	
intermediate representation		output of rule 1, input to rule 2
	rule 2	
intermediate representation		output of rule 2, input to rule 3
	rule 3	
…………		…………
	rule n	
surface representation		output of rule n

Phonological processes are conceptualized as rewrite rules which take lexical representations as their inputs, and rule interaction takes place in linear order. These rules were soon discovered to be too powerful, because there was no limit as to what possible rules could be or do. In order to avoid impossible outputs, new notions and formal devices were introduced into GP: morpheme structure conditions (MSC) appeared, comparable to *SPE*'s redundancy rules (Chomsky & Halle 1968: 171), which functioned to predict redundant phonological information in the morphemes of a specific language (Stanley 1967: 393- 436); a large number of redundancy rules and MSCs were soon replaced by universal marking conventions intended to apply to all languages and therefore established once and for all (Chomsky & Halle 1968; Kean 1975; cf. section 4.2.1 below); as phonetic counterpart to the morphophonemic MSCs, explicit surface phonetic constraints (SPCs) were called for which are formulated as implicational statements and state distributional constraints of features and segments at the phonetic level (Shibatani 1973): "[A]ny successful approach to phonological evaluation needs to proceed from the evaluation of phonetic representations" (1973: 104)—an idea that OT proponents readily ascribe to today; notions of naturalness were invoked (Schachter 1969; Schane 1972; cf. section 4.2.2

below)[34]; rules were ordered extrinsically with respect to each other (in feeding, counterfeeding, bleeding, counterbleeding rule orders; Kiparsky 1968), and the unrestricted application of rules was constrained (Kisseberth 1970). The debates about rule application and rule ordering raised the question of functionality in phonology (an issue which is once more revived in recent models of OT; cf. chapter 5 below). The more functionally oriented approaches within generative phonology assumed rule application and rule order to be governed by natural or functional principles, while the more formal approaches favored an intrinsic ordering principle which is characterized by the logical properties of the rule system determining the sequence of rule application.

Kisseberth, for example, made an important contribution to phonological theory, especially from a more functionally oriented OT perspective. He realized that different, formally unrelated phonological rules may have the same output goal. The three rules below all share one common function: a ban on the configuration XAY.

(4.2) Three rules conspiring against the output XAY

$$A \rightarrow B \,/\, X \_\_\_ Y$$
$$A \rightarrow C \,/\, X \_\_\_ Y$$
$$A \rightarrow \emptyset \,/\, X \_\_\_ Y$$

Kisseberth calls this effect rule conspiracy (1970: 294) and recognizes that a grammar which values economy and simplicity of description very highly should handle this situation with one simple constraint on the output form: *XAY. Rules should not be classified according to their structural description but to their output-oriented function. The attention is shifted away from rules that apply to a certain input class of segments to rules that conspire to produce a specific output form. To the best of my knowledge, Kisseberth was the first to incorporate the notion of a derivational constraint into phonological theory, realizing that a functional unity of rules remains unexplained if the focus lies on inputs while outputs are disregarded (1970: 304). Other efforts to constrain GP were made in one instance by increasing the number of conditions on inputs—MSCs as mentioned above (Stanley 1967) or

[34] Schachter's and Schane's notions of naturalness are to be distinguished from other uses of the term, e.g. from natural classes of segments (for a definition of natural classes cf. section 4.3.1 below).

constraints on syllable structure (Hooper 1972; Kahn 1976)—and in another by representations that encode considerable restrictions in themselves, e.g. underspecification of feature values and nonlinear representations of autosegmental spreading (cf. section 4.3 below). Advocates of OT see an unnecessary duplication of grammatical devices in having static conditions on lexical representations and dynamic rules responsible for a subsequent structural change. Instead, OT is completely surface-based in that only surface forms are evaluated.

4.1.2 OT architecture

Before evaluation takes place, output forms (candidates) are generated in the GENERATOR (GEN), one of the three main components of the OT grammar. GEN creates any conceivable output candidate for some input selected from the LEXICON. In doing so, GEN may modify the given input by generating for it a huge amount of output candidates with changed, deleted, inserted, and permuted features or segments; it may also provide or even alter prosodic structure (McCarthy & Prince 1993: 86). Consequently, the extremely creative and productive GEN assumes the task of the former language-specific phonological rules which perform alterations[35] (deletions, insertions, assimilations, etc.) on underlying representations so that the grammar can account for specific output structures. GEN, however, creates a highly varied candidate set in one single operation with all possible forms present at once. GEN is subject to one restriction only: it must produce linguistic objects, i.e. objects which belong to the universal vocabulary of featural, segmental, prosodic, morphological, and syntactic structure, or "representational primitives" in the words of Prince & Smolensky (1993: 4). Since one of OT's maxims is its emphasis on surface evaluation of language structure, further limitations on the generation of candidate sets are neither possible nor desirable. This makes the output set virtually infinite, a characteristic of GEN called Freedom of Analysis:

> Freedom of Analysis is absolutely essential. Because of it, the basic principles of representational form supply a range of candidates so inclusive that no specific rules or repair strategies need be posited. There is, for example, no rule

[35] This type of rule application is commonly known as the structure changing mode of a P-rule in GP; for a discussion of structure building rules and structure changing rules in connection with Radical Underspecification (cf. section 4.3.1 below) see Mohanan (1991).

> 'add syllable' (that is, a rule of epenthesis), because GEN already, as it were, adds syllables. (McCarthy & Prince 1993: 88)[36]

The LEXICON, OT's second major component, contains all underlying representations of morphemes including their phonological, morphological, syntactic, and semantic information. A crucial property of the OT LEXICON in comparison to older generative models of the lexicon is that lexical representations are not subject to any type of constraint. Rather, output forms must be determined by the particular constraint ranking of the language in question. This feature of OT is known as Richness of the Base[37] (Prince & Smolensky 1993: 191).

> *Richness of the Base*: The source of all systematic cross-linguistic variation is constraint reranking. In particular, the set of *inputs* to the grammars of all languages is the same. The grammatical inventories of a language are the *outputs* which emerge from the grammar when it is fed the universal set of all possible inputs. (Smolensky 1996a: 3)

Richness of the Base requires that different inventories arise from different constraint rankings and not from differences in inputs, i.e. systematic lexical properties are the result of the language-specific grammar (ranking). OT thus bans *SPE*-type lexical redundancy rules, i.e the type of P-rules which generally apply in a structure building mode (cf. chapter 3.1.2 above), and MSCs on phonological inputs both of which, apart from filling in redundant feature specifications into possibly underspecified matrices (cf. section 4.3.1 below), formulate language-particular restrictions on underlying representations of lexical entries. An example is the rule restricting initial consonant clusters in English (Chomsky & Halle 1968: 171):

$$[+\text{consonantal}] \rightarrow \text{s} \ / \ + \_\_ \ [+\text{consonantal}]$$[38]

This lexical redundancy rule states that the first segment in a cluster of formative-initial true consonants (except liquids and glides) must be /s/ in English, a statement that holds at the level of underlying representation for a specific language. Constraints of this type on inputs are utterly inconceivable in an OT grammar with its sole focus on output structures.

After the generation of all potential output candidates, OT's third and central component—the EVALUATOR (EVAL)—selects one optimal candidate from a

[36] This specific feature of OT poses serious problems to an implementation of OT within linguistic models of production and processing (cf. chapter 9 below).

[37] Problems connected to a lexicon of this type will be addressed in chapter 9 below.

[38] A '+' stands for formative-initial position within the *SPE* framework.

virtually infinite list of output candidates on the basis of a language-particular ranking of the total set of universal constraints. Differences in grammaticality among the languages of the world are explained as differences in the ranking of these universal and violable constraints.

In an OT framework, UG includes a set of linguistic units (features, segments, syllables, lexical categories such as N, V, A, etc.), a set of universal constraints (CON), and the two functions GEN and EVAL, whose interaction is shown in figure 4.3 below:

(4.3) OT grammar as an input–output mechanism (after McCarthy & Prince 1993: 86)

GEN (input) → {cand_1, cand_2…cand_n}
EVAL {cand_1, cand_2…cand_n} → output

The grammars of individual languages include a LEXICON with the basic forms of morphemes listed, from which inputs are constructed, and a language-specific ranking for the constraints in CON. GEN works alike in all languages, whereas EVAL differs from language to language, since it is dependent on a language-specific constraint hierarchy. EVAL must rate the well-formedness of each member of a candidate set and come up with the candidate that best satisfies the hierarchy of ranked constraints. Rarely is it possible that a candidate can satisfy all constraints. Since the constraints relate to one another by way of strict domination, best satisfaction is accomplished by violation of a lower ranked constraint in order to avoid violation of a higher ranked constraint. This situation is illustrated schematically in the following tableau:

(4.4) Constraint tableau: Constraint_1 dominates Constraint_2 (C_1 >> C_2)

input	C_1	C_2
a. ☞ candidate a		*
b. candidate b	*!	

In this tableau, the constraints C_1 and C_2 are listed in ranked order from left to right across the top row. Candidate outputs are listed in the left column, with the sign '☞' marking the most harmonic candidate as optimal, while the second candidate is judged as suboptimal by EVAL. An asterisk in a cell indicates a violation of a constraint by the candidate in the same row, while an asterisk plus exclamation mark indicates a fatal violation of some high ranked constraint which excludes the candidate in this row from further consideration. Shaded cells in a tableau mark violation or satisfaction of the constraint in this column as irrelevant to the outcome of EVAL because of a fatal violation of a higher ranked constraint.

The computation of the optimal output candidate takes place across the whole set of ranked constraints; all candidates are assessed in one operation in parallel. This type of procedure marks OT's total rejection of serial derivations by means of rules and rule orderings, which had since been the center of attention of GP.

(4.5) Basic OT architecture (after McCarthy 2002: 10)

$\boxed{\text{LEXICON}} \rightarrow \text{input} \rightarrow \boxed{\text{GEN}} \rightarrow \text{candidates} \rightarrow \boxed{\text{EVAL}} \rightarrow \text{output}$

By positing CON, the set of universal constraints, OT formalizes the universal observation that all grammars are systems of conflicting forces. While structuralist phonological theories had placed more emphasis on the notion of contrast, both at the level of the phoneme and at the level of the feature (Trubetzkoy 1939 and Jakobson, Fant & Halle 1951 respectively, cf. chapter 3.1.1 above), later generative models realized the necessity of resolving conflicts. In derivational models (e.g. *SPE*, Lexical Phonology), this was accomplished by means of rule ordering, whereas representational models (e.g. underspecification theory, nonlinear phonology) formulated restrictions on the type of phonological representations licensed by UG. In OT, resolution of conflict is *the* central concern. Constraints are typically in conflict, since some of them will demand identity between underlying and surface forms while others will prohibit some surface form on the basis of phonetic or phonological plausibility considerations. The two main families of constraints are FAITHFULNESS (F) and MARKEDNESS (M)[39], the two making contradictory requirements on output forms.

[39] F and M are notations for constraint families.

FAITHFULNESS aims at the preservation of contrast in inputs which encode meaning. With respect to sound segments, FAITHFULNESS strives to retain lexical specifications by all means. For example, a feature [+voiced] in obstruents is preserved in all contexts, even in word-final or syllable-final positions, and the opposite feature specification [–voiced] in obstruents is also preserved in all contexts, even between vowels. Such outputs would be completely faithful to their inputs. It seems as if OT's notion of FAITHFULNESS has its roots in conditions or principles of older generative models: one example is Structure Preservation, which is part of the framework of Lexical Phonology (Kiparsky 1985: 92-4); it guarantees that constraints on the URs of segments hold throughout the derivation within the lexical domain of the phonology. However, traces of faithfulness can be detected even earlier in Chomsky & Halle's Invariance Condition (1968: 166). When the Invariance Condition is met, there is no divergence between the input to the phonological component (the phonological matrix) and the output of the phonological component (the phonetic matrix) (cf. chapter 3.1.2 above); the P-rules operate in their structure building mode. Obvious violations of the Invariance Condition led Chomsky & Halle to assume structure changing P-rules: "[P]honological rules not only fill in redundant entries of matrices but also may change inherent features marked in the lexical entry" (1968: 167). The two concepts, invariance and faithfulness, are at least closely related to each other. FAITHFULNESS works hand in hand with the requirements of speech perception in keeping outputs maximally distinct.

MARKEDNESS constraints are inherently in conflict with FAITHFULNESS. MARKEDNESS aims at the suspension of contrast in inputs in favor of the unmarked member in an opposition: $\mathbb{M} >> \mathbb{F}$. On the other hand, whenever some lexical contrast is preserved, some MARKEDNESS statement will be violated ($\mathbb{F} >> \mathbb{M}$), because in every opposition one of the two members is marked. MARKEDNESS may work hand in hand with the needs of speech production, but not necessarily so: a severe shortage of lexical contrast, with lexicons using only the most unmarked segment types, would ultimately lead to individual words being extremely long and would fail to meet the lexical requirements of languages, which need to encode an average of 100,000 meanings (Kager 1999: 7).

An elaboration of markedness theories, research in language typology and language acquisition, and a better comprehension of the universal aspects of the grammar's different domains are the main interests of linguists working within an OT framework, which as a general theory of grammar aims at finding explanations and making predictions in all linguistic domains (for a critical analysis of markedness conceptions cf. chapter 9.6.1 below).

(4.6) Summary of the five basic tenets of OT (after McCarthy & Prince 1994: 3):

1. **Universality**
   UG provides a set CON of constraints that are universal and universally present in all grammars.
2. **Violability**
   Constraints are violable; but violation is minimal.
3. **Ranking**
   Constraints are ranked on a language-particular basis; the notion of minimal violation is defined in terms of this ranking.
4. **Inclusiveness**
   The constraint hierarchy evaluates a set of candidate analyses that are admitted by very general considerations of structural well-formedness.
5. **Parallelism**
   Best satisfaction of the constraint hierarchy is computed over the whole hierarchy and the whole candidate set. There is no serial derivation.

I will present an overview of the development of markedness concepts in GT for two reasons: First, it seems to be *the* crucial notion upon which explanatory adequacy in OT is based; second, I have some reservations in mind about the explanatory value of markedness statements, which I will discuss in more detail in chapter 9 below.

4.2 UNIVERSALITY AND MARKEDNESS

The notions of universality and markedness have a long and controversial history in linguistics in attempts to deal with regularities found repeatedly, but not necessarily without exception across languages. Generalizations which have as their scope all languages, e.g. all languages have consonants and vowels in their inventories, are desirable when searching for universals in typological studies of language. The logical format of absolute and inviolable universals can be stated as: $\forall x\ [Px \rightarrow Rx]$; that is, for all x it holds, if x is a language (i.e. has the property P), then x also has the property R (i.e. has consonants and vowels).

Unfortunately, however, unrestricted and unconditional universals of this type are rare (Greenberg 1966: 10). Most generalizations must admit exceptions in that they describe universal tendencies, properties about languages that usually hold true but not always, e.g. all but one language have coronal plosives in their inventories—the notable exception being Hawai'ian (Maddieson 1984: 32). Nonabsolute and thus violable universals take the form: of all Px, *almost* all have the property R, or put differently, in the ideal, *unmarked* case they have the property R.

These logically weaker statements are the basis for OT's violable MARKEDNESS constraints. Before turning to markedness theory in *SPE*'s more restricted sense, I will summarize very briefly what is commonly meant by the notion of markedness in linguistics. The idea of marked and unmarked categories originates from the Prague school phonologist Trubetzkoy (1939) and refers to a difference between two or more elements of a class of otherwise equivalent elements. Depending on the specific definition of the *unmarked* member, it is more neutral, more natural, or more basic than its *marked* counterpart(s). Greenberg gives a set of five criteria for defining markedness in phonology:

(4.7) Criteria for markedness in phonology (after Greenberg 1966: 58-9)

a. the unmarked element is the element that contrasts neutralize to
b. the unmarked element is more frequent than the marked element
c. the unmarked element has greater allophonic variability than the marked element
d. the number of segments with a marked feature is always smaller in an inventory than the number of segments with an unmarked feature
e. the basic allophone (of a phoneme) is the one with the unmarked feature

The criteria b. and d. both refer to frequency, more precisely to the "typological inventory frequency" and to the "inventory frequency" of segments in a specific inventory, as discussed by Paradis & Prunet (1991: 11). Both criteria apply to the set of coronals in SAE (cf. section 4.3.1 below). Following a thorough discussion of the five criteria depicted in figure 4.7 above, Greenberg adds a sixth criterion for phonological markedness independent of distributional and frequency phenomena, namely articulatory complexity:

> A particular articulation is to be considered more complex [and more marked] than some other if it includes an additional articulation defined in terms of departure of an organ from the position it normally has in the absence of speech. (1966: 70-1)

Articulatory complexity as a means of distinguishing marked and unmarked categories is important in connection with the first criterion (a.). Neutralization and assimilation are both characterized by a certain degree of loss of contrast. However, while neutralization has as its result a less complex and thus unmarked segment, assimilation frequently gives rise to more complex and thus marked segments (e.g. place assimilation discussed in chapter 8.2.1.1 below).

As pointed out in the previous section, one of the first serious problems generative phonology had to deal with was the unconstrainedness of its rules, whose major function was to transform an underlying representation stepwise into its output form. Such a purely formal calculus, based solely on rule complexity in terms of feature specifications, is absolutely neutral to the phonetic content of the representations, e.g. to articulatory complexity mentioned above. This means that outputs can be derived which are not part of any existing phonological system. The inadequacy of this theory was realized almost immediately by the authors of *SPE* and led them to supplement the book with a kind of remedy theory: markedness theory (MT), which is *SPE*'s famous chapter 9.

4.2.1 Markedness in *SPE*

In *SPE*-phonology, P-rules of the shape A → B / X ___ Y convert fairly abstract lexical representations of the surface structure of sentences, which constitute the output of the syntactic component of generative grammar, into phonetic representations. This sound continuum is analyzed as a sequence of discrete sound segments, which consist of complexes of distinctive features drawn from the fixed set of universal features. The problem of rule unconstrainedness emerges in part from Chomsky & Halle's abolition of a phonemic level, intermediate in abstractness between lexical and phonetic representation (1968: 11).[40] The intended result is a considerable increase in abstractness and an extension of the power of phonology. To account for alternations in the surface structure of roots[41], e.g. criti[sɪzm]–criti[kəl]–criti[sajz], successive transformations on one common underlying representation are invoked. Among these operations are velar softening of /k/ to [s] and

[40] The terms *phoneme* and *phonemic analysis* are also banned from *SPE* vocabulary (ibid).

[41] A morphology module did not exist as part of the grammar of Standard Theory: while the phonological component (P-rules) was responsible for allomorphic variation, word formation was handled by the rules of syntax. This, of course, added to the general power of rules in the generative framework.

diphthongization of the word-final vowel, which in the end lead Chomsky & Halle to postulate the form criti[kijz] as UR (1968: 54). The format of this derivational approach forces phonologists to choose underlying forms, some of which are so abstract that they are never found in the pronunciation of actual English words.

The P-rules then are the devices which perform the operations on lexical representations—e.g. change feature values (+ to –, or vice versa), add nondistinctive features, delete, alter, or insert segments—and hereby change the underlying to phonetic representations. An assimilation rule, for example, is a P-rule by which "two segments are made to agree in the value assigned to one or more features" (Chomsky & Halle 1968: 350).

However, a problem arises when comparing two *SPE* rules (Chomsky & Halle 1968: 178): one assimilation rule and one dissimilation rule. The assimilation rule occurs quite frequently (voicing agreement in consonant clusters) while the dissimilation rule states the rare instance of diphthongs taking only low back vowels before the nonback glide [j] and only nonback vowels before the back glide [w]. Yet the theory makes the more unlikely dissimilation rule appear simpler because one feature specification less is required for its formulation; it is thus ranked higher on the naturalness[42] scale than the assimilation rule, a prediction that proves the theory to be deficient.

Among the early critics of the *SPE* framework is Schachter, who also discusses a pair of assimilation rules. He assumes vowels in some hypothetical language to be underlyingly voiced and consonants underlyingly voiceless, e.g. in the sound sequence /atat/. A rule voicing intervocalic consonants to yield [adat], which is widely attested (it is in fact part of the system of American English), is just as simple as a rule devoicing interconsonantal vowels to yield [atạt], which Schachter doubts is part of any phonological system (1969: 343-4). The framework of standard generative phonology is not sufficiently restrictive concerning rules which are either extremely uncommon or never occur in the languages of the world. An evaluation metric rating the naturalness of a rule in terms of the number of feature alterations cannot predict that one of two rules may in fact be more 'natural' than the other. Schachter summarizes: "[T]he current model falls short of explanatory adequacy in this case" (1969: 344). Furthermore, even Kenstowicz & Kisseberth's standard

[42] The use of the term *naturalness* in *SPE* does not correspond to the definitions of naturalness within Natural Phonology and Natural Generative Phonology (cf. section 4.2.2 below).

textbook for the *SPE* theory of the 1970s has some criticism which points in a similar direction:

> Any adequate theory of phonology must contain postulates that will define natural sound changes. Although many of these can be expressed by appeal to the notion of assimilation defined over the features of a feature system, it is clear that not all natural sound changes fit into this mold. For example, many languages have a rule converting consonants to *ʔ* or *h* in preconsonantal and final position. Such a process is clearly not assimilatory in nature. Nevertheless phonological theory must have some apparatus for expressing the fact that neutralization to a glottal stop in these positions is a natural rule as opposed to, say, neutralization to *l*. (1979: 251)

In order to constrain the theory and account for the intrinsic content of the features, Chomsky & Halle incorporate the Praguian notion of *markedness* into the *SPE* model. They introduce a set of universal marking conventions in order to distinguish "natural" and "expected" feature conjunctions from "unexpected" and "unnatural" ones within a given segment (1968: 402). If a specification (either "+" or "–") of a feature is the likely specification, it will be the unmarked (*u*) specification; and conversely, if the specification is unlikely, it will receive a mark for it and carry the value marked (*m*) in a particular intrasegmental context. Underlying representations of segments are characterized in terms of *u* and *m* feature specifications, with only *m* specifications contributing to segmental complexity. When the universal marking conventions cannot specify the unmarked or marked feature value in a segmental context, it will be specified "+" or "–" directly in the lexical representation. Such a specified feature will increase the complexity of the segment to the same extent as a feature with the coefficient *m*. I repeat the example /kæt/ from chapter 3 in figure 4.8 below, this time showing a feature matrix after the application of the marking conventions. Interestingly, the segment /t/ is not unmarked for coronality, as might be expected; instead it is specified [+coronal] underlyingly and thus has the (low) complexity 1. The reason for this is *SPE*'s marking convention XXIII (1968: 406), which defines segments [u coronal] that are either [–coronal] and [–anterior] or [+coronal], [+anterior], [+nasal], and [m continuant]. Contrary to the representation for vowels, where /a/ is the unmarked vowel (0 complexity), there is no unmarked consonant defined by *SPE*'s marking conventions. There are an unmarked nasal /n/, an unmarked continuant /s/, and "the unmarked plosive—if there were to be one—would have to be /t/" (1968: 413); all three are coronal consonants.

(4.8) *SPE*-type lexical representation of <cat> indicating only marked and specified features.

	k	æ	t
consonantal			
nasal			
low		m	
high			
back		m	
round			
anterior	m		
coronal			+
continuant			
complexity	1	2	1

Kean (1975) presents a revision of MT with a different ordering of the markedness conventions. Her system has an unmarked vowel, which is also /a/, and an unmarked consonant with 0 complexity—the voiceless coronal plosive /t/ (1975: 32). This is due to Kean's marking convention IX, which defines [u coronal] quite differently from *SPE*: a segment unmarked for coronality is [+coronal], [–back], and [–labial] (1975: 14). According to Kean the "optimal" segment inventory would be the system /i,a,u; p,t,k,ʔ,n,h,w,j/ (1975: 41); it is regarded as a formally simple inventory because of its large number of unmarked feature values and its low overall complexity.

Both approaches work with context-free marking conventions, in the sense that only individual segments and segment inventories are evaluated for their simplicity and naturalness (unmarkedness)[43]. Context-sensitive conventions referring to the interaction of segments in processes like assimilation, for example, are not employed. It is further doubtful, whether speakers really internalize almost empty matrices of the type above, since these representations would require speakers to process an enormous number of rules to finally arrive at the phonetic representation (cf. also the discussion of underspecification in section 4.3.1 below).

[43] Whether these two terms can actually be used synonymously will be discussed in section 4.2.2.

However, the general idea that markedness is basically a relative concept can be grasped even from these early versions of MT: A marked linguistic element, here a feature, is not well-formed or ill-formed in itself, but only in relation to other linguistic elements. Marking conventions are presented as universally valid rather than language-specific generalizations. They will enforce unmarked types of linguistic structure in individual languages, but they also allow for the possibility of particular languages to violate the conventions in a soft sense and produce marked structures if a particular grammar explicitly states this.

Certainly, the theory of markedness is an important addition to the original *SPE* model. It allows a distinction between more and less likely segments, and in later versions markedness ranks the P-rules along a scale of complexity in a more realistic fashion. However, a latent danger of circularity seems to be connected with MT (cf. also the elaboration on this problem in section 4.4.3 below). Lass describes the inherent problem as follows:

(4.9) Logical progression of markedness generalizations (after Lass 1975: 486-7)

1. Some configurations, segment-types, etc., are relatively common, and others are not.
2. Therefore 'common' = 'unmarked' and (relatively) 'uncommon' = 'marked', i.e. the unexpected is noticeable.
3. Since what is commonest must clearly be the most 'natural' state of things, 'unmarked' = 'natural' and 'marked' = (relatively) 'unnatural'.
4. Since the viability shown by something 'common' and 'natural' is a sign of its (relative) optimality, and since our metatheory tells us that 'optimization' is to be defined in terms of criteria of 'simplicity' for the person speaking or acquiring language, then 'common' = 'natural' = 'optimal' = 'simple'.
5. This shows that 'common' or 'natural' = 'simple' (and conversely, 'uncommon' or 'unnatural' = 'complex'), because things would not, given the assumptions in 4., be so 'common' and 'natural' if they were not also 'simple'.

4.2.2 Naturalness and OT

Lass' critical description of MT is interesting in one sense because the notion of optimality is addressed in a pretheoretical sense and in another because naturalness and unmarkedness seem to be almost interchangeable concepts for him. How intimately the two are in fact related and in which form naturalness resurfaces in OT will be discussed in this section.

Considerations of naturalness became an important issue after the publication of *SPE*, as the limitations of an approach became evident which tends to reduce language to pure form and linguistics to the study of the internal properties of particular formal systems. Suddenly, there was an increased tendency to focus on the substance of language, to look for a basis for language facts in extralinguistic domains, e.g. phonetics, and thus to show that language is 'natural' (cf. also section 4.3.3 on Grounded Phonology). Naturalness is a notion that was deliberately introduced to reduce excessive abstractness. Schachter's motivation for the proposal of natural and unnatural processes (cf. section 4.2.1 above) was the observation that many assimilatory processes are asymmetric in nature. One such example is the nasalization of vowels before nasal consonants:

(4.10) Two formally symmetric rules, the 1st describing a natural process, the 2nd an unnatural one.

Rule 1: $V \rightarrow [+nasal] / \_\_\_ [+nasal\ C]$

Rule 2: $C \rightarrow [-nasal] / \_\_\_ [-nasal\ V]$

While vowels are quite frequently nasalized in the neighborhood of nasal consonants (rule 1), denasalization of a consonant in the vicinity of an oral vowel is not frequently attested and is thus rather unnatural (rule 2). Schachter proposes equating natural rules with 'no cost' rules in terms of counting the unmarked and marked feature values involved (1969: 346) thereby extending the notion of markedness from segments and segment inventories to rules.

However, all early pronouncements on naturalness suffer from one major inadequacy, namely the lack of a coherent definition of what exactly is meant by 'naturalness'. A common practice was to make purely subjective judgments about the naturalness of rules and processes, giving a phonetic 'explanation'[44] of the dubious type 'ease of articulation', e.g. for consonantal place assimilation. Rules of assimilation were seen as prototypically natural because they meet the requirements of articulation (Schane 1972: 213). A process occurring in a variety of languages, even if the phonetic cause is unknown, is also regarded as indicative of naturalness. Schane, for example, makes no secret of his procedure:

> Why are natural rules 'natural'? So far I have been relying heavily on intuitive feelings for determining which rules are natural. It goes without saying that

[44] For a discussion of what is an explanation in linguistics cf. chapter 9.6.1 below.

> such intuitions exist. […] Natural rules, then, have a universality—we expect them as phonological phenomena in languages of diverse type. Why should such rules recur in language after language while other rules—that is, the unnatural ones—may have an extremely limited distribution? (1972: 206)

This approach to naturalness of course turns out to be circular: natural rules are natural because they frequently occur, and they occur frequently because they are natural. Natural, then, is basically anything that occurs in nature. Criticism of this kind is indeed more than justified (e.g. Lass 1980; Anderson 1981); but Schane's paper contains another interesting feature, which anticipates a basic tenet of Natural Phonology (see below). He claims natural rules to be universal and context-free, i.e. all environments in which they apply are predictable while deviations from the natural rules are unnatural, context-sensitive, and language-specific. Unnatural rules are thus part of individual grammars (1972: 228).

For some phonologists, efforts like for example Schane's (1972) more or less intuitive account of naturalness or Kiparsky's (1968) approach to rule ordering are not carried far enough in the direction of constraining generative power. Hooper (1976) introduces a coherent phonological framework, known as Natural Generative Phonology (NGP). According to Hooper, the basic tenet of NGP is the following:

> The major claim of natural generative phonology is that speakers construct only generalizations that are surface-true and transparent […]. An important property of surface-true generalizations is that they are falsifiable in a way that the more abstract generalizations of generative phonology are not. (1979: 106)

Speakers will thus prefer a transparent morphological analysis over an abstract phonological analysis. Three constraints are formulated to limit the excessive abstractness of *SPE*-type phonology, the first of which is a condition on rules: The True Generalization Condition (1976: 13) claims that speakers and language learners alike, who are confronted with surface data, will only formulate hypotheses that are consistent with surface forms rather than constructing abstract and opaque rules (cf. OT's principle of Lexicon Optimization below). The natural rules of NGP, then, relate one surface form to another instead of relating an UR to a surface representation (SR) (cf. OT's IDENT-OO constraints, which require uniform output paradigms; Kager 1999: 414-7). Next is the No Ordering Condition (Hooper 1976: 18), a constraint on rule application that explicitly prohibits extrinsic ordering: extrinsic rule ordering has the effect that each rule has an arbitrary position within a fixed rule order and applies only once to the output of a preceding rule. In NGP however, rules apply whenever their structural descriptions are met, a quality which in a way likens these rules to OT's constraints applying all in parallel. And finally, the Strong Naturalness Condition (1976: 116) is a constraint on URs requiring lexical representations to be identical to their surface (phonetic) forms as though the P-rules had already applied to derive a SR (cf. figure 4.8 above). The consequence of such a highly constrained theory is a greatly restricted scope of phonology. Since every rule must represent a transparent generalization about the surface of the language, P-rules must operate without exception, and only these surface-true systematic relations count as phonology proper, and hence as natural. Three rule-types are assumed in NGP: these include the P-rules, morphophonemic rules (MP-rules), and via rules.

NGP's P-rules are specified in purely phonetic terms, i.e. in phonological features and phonological boundaries (syllable boundaries, pause boundaries). P-rules are automatic, e.g. the formation of the regular English plural or the rule aspirating syllable-initial voiceless plosives in English. These rules correspond to the natural processes of Natural Phonology (see below). Morphophonemic rules (MP-rules) refer to morphological or syntactic categories, to arbitrary lexical categories, and to word or morpheme boundaries. The rule voicing fricatives in certain English plurals (e.g. <life> vs. <lives>) is an example of an MP-rule, which contains both lexical and morphological information: it applies to only a small class of nouns (<knife, house>, but not <safe, face>), and it applies only in the plural. Via rules are used in NGP to relate pairs of words without deriving them from the same underlying form. An appropriate account of the synchronic residue of the Great Vowel Shift of English, for example, is a via rule. Accordingly, ser[ij]ne and ser[ɛ]nity would not be derived from the same underlying form as proposed in *SPE*; rather, the underlying forms would be essentially identical to the phonetic forms, and the lexical entries for both items would be marked to indicate that they are related to each other by a via rule of the following format: [ij] ↔ [ɛ]. Therefore, it would be possible for individual speakers to perceive a relationship between <serene> and <serenity>, but they would fail to see a relationship between <obscene> and <obscenity>. In this case, the former pair would be lexically marked as related by the via rule, whereas the latter pair would not.

NGP's proposals did not receive much appreciation from phonologists, and work within this framework stopped after the 1970-s, but looking at NGP from an

OT perspective, it is evident that some issues which were first addressed in this context, such as NGP's insistence on surface-true generalizations, resurface in OT. Moreover, repercussions are not restricted to the area of phonology. It is interesting to note that Jackendoff, focusing on morphology and the nature of the lexicon, as early as 1975 postulates lexical redundancy rules, which—as far as I can judge—have exactly the same function as NGP's via rules. Jackendoff's aim is to formulate a theory of the lexicon in line with the Lexicalist Hypothesis expressed in Chomsky's famous paper 'Remarks on Nominalization' (1970). Whereas *SPE*-type markedness theory would claim one basic member and one derived member in a pair like <decide> vs. <decision> (on account of naturalness considerations, the shorter one wins), Jackendoff postulates two distinct but related lexical entries (1975: 640-1). In Jackendoff's formalization, separate but related entries are linked by lexical redundancy rules that appear to be the mirror image of NGP's via rules: lexical redundancy rules function as descriptions of the partial relations and idiosyncrasies typical of the lexicon.

In many ways different from NGP, Natural Phonology (NP) as presented by Stampe (1979) also distances itself explicitly from earlier generative models, especially from *SPE*. Stampe models his theory on Jakobson (1941), who showed that feature values are asymmetrically distributed and acquired, and that children's speech is characterized by the use of unmarked feature values.

> [T]he child possesses in the beginning only those sounds which are common to all languages of the world, while those phonemes which distinguish the mother tongue from the other languages of the world appear only later. (1941: 50)

Guided by these observations, Jakobson formulated implicational universals ("laws of solidarity") of the following kind: If a language admits a marked feature value (e.g. nasal vowels), it must also admit the unmarked value (oral vowels); in language acquisition, "secondary vowels" arise only after the corresponding "primary vowels" (1941: 51-7). NP accounts for the obvious differences between the speech of very young children and that of adults speaking the same language by assuming that children's speech is governed by a large number of natural processes. Part of the acquisition process consists of suppressing these natural processes in favor of language-specific acquired rules. In each language, mature speakers have learned to suspend certain processes, while others are left unaffected. Since the set of unaffected processes varies from one language to the next, NP takes this variation to be the cause for the differences among languages. Phonological processes are termed natural, because NP conceives of these processes as rooted in the articulatory and perceptual capabilities of humans, displaying, as it were, the sum total of what is possible in human speech. These processes are said to be universal, innate, and inviolable in the sense that they represent the phonological capacity children start from when learning any language.

> [T]he phonological system of a language is largely the residue of an innate system of phonological processes, revised in certain ways by linguistic experience. (Stampe 1979: vii)

Stampe clearly anticipates ideas that later appear in OT. All children have at their disposal a common basis in the shape of very general natural processes. These processes, which appear to be very much like OT's universal set of constraints (CON), are subject to suppression depending on the language to be acquired (cf. OT's language-specific ranking of the constraints in CON). The outcome are language-specific acquired rules, examples of which are given by Stampe: the alternation of [k] with [s] in English words of Romance origin, e.g. in word pairs like <electric> vs. <electricity> (1979: 45) and <critic> vs. <criticism> (cf. *SPE*'s analysis of velar softening presented in section 4.2.1 above) and the effects of trisyllabic laxing in <serene> vs. <serenity> (cf. NGP's analysis as a via rule above); all of these are not phonetically motivated and are thus not natural processes. Rules typically operate in a selective fashion (not all /k/ segments become [s] when followed by written <i>), and are sensitive to grammatical considerations (suffixation of <-ity>). Processes, on the other hand, operate across the board without a single exception. Rules must be learned while processes are, at least partially, unlearned.

> Implicit in each process are various subtle and strict hierarchies, ranging from the greatest generality which is phonetically motivated, to the complete suppression of the process. (1979: viii)

In contrast to *SPE*, which highly values the formalization of language-internal factors as a method of arriving at its explanations, Natural Phonologists emphasize data drawn from children's speech, careful and casual speech, speech errors, and other sources of what is commonly referred to as external evidence. These domains receive so much attention from NP because it is claimed that there is a known base variety or

UR (e.g. an adult informal standard pronunciation, cf. chapter 2.2.2 above) from which the observed varieties systematically differ in ways that are not learned (1979: xix).[45] The universal processes are inferred from exactly these nonconventionalized phenomena.

An example of such a natural process is the denasalization of vowels, a process which is differently distributed across different languages. In English, nasal vowels are allophones of nonnasal vowels, that is, nasal vowels are not present at the level of UR. Within the framework of NP, the nonoccurrence of nasal vowels is 'explained' (cf. chapter 9.6 below for the status of linguistic explanations) by the universal process of vowel denasalization. Allophones are thus the product of a language-specific, context-sensitive rule, sounds of a type which have already been eliminated by a prior natural process (1979: 27). Only in a language that completely lacks nasal vowels does the natural innate process remain totally unconstrained, whereas French, for example, has a very strong constraint suppressing the denasalization process completely allowing for an underlying contrast between nasal and oral vowels to be upheld.

Another natural process is the devoicing of syllable-final and word-final obstruents, an example of which is the German 'Auslautverhärtung'. Alternations as in [tɑ:k] vs. [tɑ:gə] are not imitated by children acquiring German but are developed independently because they are the result of a natural process (Donegan & Stampe 1979: 134). Children acquiring English, for example, will have to learn a language-specific constraint partially undoing this innate predisposition. While *SPE* has taken over the structuralist concept of neutralization of phonetic contrast in certain environments and has postulated an underlying archi-segment[46], /G/ in this case, NP 'explains' this process with the notion of phonetic naturalness. In the context mentioned above, obstruents are naturally voiceless because during their production the airflow necessary for voicing is obstructed, whereas in voiced environments assimilation naturally causes obstruents to be voiced. Stampe observes that phonological processes often have such contradictory outputs because they reflect conflicting phonetic requirements (1979: vii). This notion is also mirrored in OT's

[45] In Stampe's view, different dialects of one language (cf. chapter 2.2 above) are the result of the suppression of distinct natural processes (1979: xix).

[46] Archi-segments are not fully specified underlying segments in phonological matrices as opposed to the fully specified segments in phonetic matrices (cf. chapter 3.1.2 above).

constraints which are inherently in conflict. A task of NP is to explain the conflicts in terms of phonetic plausibility of vocal tract configurations and offer solutions that resolve these conflicts in the grammar. Specific phonologies balance competing needs and capacities, such as making oneself understood vs. speaking at a reasonable rate without undue effort by fully or partially suppressing natural processes.

The postulation of language-specific rules restricting general processes in place of powerful rules generating whatever appears at the phonetic surface and an emphasis on conflicting forces and their resolution in individual grammars are important theoretical assumptions reappearing in OT. However, some of the conceptions formulated in NP liken this model more to other models of generative theory than to OT. An NP grammar consists of inviolable natural processes which are universal; it is this universality that has rightly been criticized. Anderson (1981) shows in his paper "Why Phonology isn't *Natural*" that phonological processes may indeed originate in phonetically natural processes, but once they become an integral part of a particular language system, they are fossilized and behave arbitrarily. NP still needs rules interacting with the natural processes, and these rules are ordered (feeding, counterfeeding, bleeding, counterbleeding orders; Kiparsky 1968) to derive the phonetic surface from a lexical representation (Donegan & Stampe 1979: 145-58).

Where OT is concerned, Prince & Smolensky (1993) explicitly mention one highly relevant notion as an elaboration of NP: it is the demand that underlying representations should match surface representations as closely as possible in the absence of contrary evidence. This resembles Stampe's claims about representations:

> A corollary of the theory of underlying representation presented here is that underlying segments are ontologically of the same status as any segment in surface representation; they are mental representations of sounds which are, at least in principle, pronounceable. They are not, in particular, semi-abstractions like the 'archisegments' of structural and generative phonology. (1979: 35)

As mentioned in section 4.1 above, the surface form is the one that receives the highest evaluation in OT. However, a problem with underlying lexical specifications may occur where MARKEDNESS constraints are ranked above FAITHFULNESS constraints. In such a situation, an IDENTITY-INPTUT-OUTPUT constraint (IDENT-IO), which requires the feature specifications of an input segment to be preserved in its output correspondent, is rendered completely inactive. No matter whether the input is a nasal vowel or an oral vowel, if a MARKEDNESS constraint of the type $*V_{NASAL}$ is ranked above IDENT-IO, the output will be an oral vowel irrespective of the input. Although all inputs are possible in principle (cf. Richness of the Base in section 4.1 above), Prince & Smolensky do not think this to be a desirable side-effect of ranking, especially in view of language acquisition. Children acquiring a language should be able to construct a lexical form (input), e.g. /kæt/ from a surface form (output) [kæt]. They should do so by assuming an input identical to its output unless evidence from surface forms tells them to postulate a nonidentical input form, e.g. a nasal vowel. This strategy, called Lexicon Optimization (1993: 192) in OT (cf. critical discussion of Lexicon Optimization in chapter 9.2 below), clearly mirrors NP:

> [S]egments are taken at face value unless this is prevented by their alternants or by a general process that bars them from underlying representation. This means that all underlying segments are fully specified as to their underlying phonetic features. (Stampe 1979: 38)

Both NP and NGP make claims with respect to phonetic substance in order to counterbalance the overly formal approaches of former generative phonology. However, NGP allows even less abstractness in URs than NP by formulating very strong constraints on phonological representations and rules, all of which refer exclusively to the surface forms of the language. But whereas NGP is still rule-based, even if the rules are highly constrained, NP is a process-and-constraint (=suppression)-based approach. OT thus acquires from NP the salient notion of a universal base that needs to be constrained in language-specific ways and incorporates from NGP the strong requirement on surface-true statements.

Ultimately, it seems to me that an interesting question remains as to whether naturalness and markedness really are different concepts notwithstanding the fact that the former is superficially associated with articulatory or perceptual substance while the latter commonly stands for a formal evaluation metric assigning degrees of complexity to language structures. Stampe compares *SPE*-type markedness theory with his concept of naturalness and comes to the conclusion that natural processes indeed resemble implicational laws and marking conventions, but that the latter merely have the status of appearances, at the base of which is an innate system of natural processes (1973: 52). Turning to the example of nasality in vowels again, NP's alleged theoretical superiority is based on its explanation of nasal vowels in a language: nasal vowels are considered to be effected by the suppression of a natural process denasalizing vowels. In markedness theory, on the other hand, nasal vowels (falsely) appear to be the result of the acquisition of a marked feature value, an exception to the law or convention (1973: 45). Abstracting from the question of whether a constraint-based approach is to be favored over a rule-based approach, it seems that there is no substantial difference between the two concepts. Considering again Lass' (1975) logical progression of markedness generalizations (cf. figure 4.9 above), naturalness is a precondition for unmarked structures and vice versa. I agree with Lass that what is expressed by these concepts is

> the blinding tautology that nature tends toward the natural; […] markedness theory in its formal dress is simply an alternative (actually pseudo-mathematical) representation of this intuitive judgment. There is therefore no difference (except the trivial formal one) between the two notions: they both make the same predictions, have the same conceptual content, and may be regarded as the same thing in two different languages […]. (1980: 43-4)

4.3 REPRESENTATIONS AND LEVELS OF ANALYSIS

Different notions of markedness and naturalness still prevail in the discussions of phonological phenomena. In this section, I will turn back to features and their representations within current phonological theory. I have presented arguments showing that linear phonology (*SPE*-type) does not provide adequate representations because within this model almost any change can be operated irrespective of its context. An open question is whether underspecification theory (UT), which was developed well before the advent of OT and which is regarded by some phonologists as a recent version of theories of naturalness or markedness (Paradis & Prunet 1991: 21), has suitable representations to offer to OT. This issue is discussed in section 4.3.4 below. Representations that are used as a matter of course in OT are commonly nonlinear representations. The main innovation of nonlinear phonology is perhaps the elimination of arbitrariness in phonotactics and phonological alternations through the employment of elaborated and explicit representations. Nonlinear phonology relates phonological processes to representations, which by their very structure reveal well-formedness conditions and possible process types, rather than to rule typologies or to nonstructural, intuitive notions of naturalness (cf. section 4.3.2 below).

I also want to examine which and how many levels of analysis are necessary for an investigation of the behavior of coronals under OT, since *SPE*'s levels (mainly the lexical and the phonetic levels) can no longer be sustained. It was not least the postulation of these two very abstract levels that resulted in the disturbing rule unconstrainedness and the mushrooming of intermediate representations which could no longer be located on designated phonological levels (cf. section 4.2.1 above). In stressing the importance of evaluating relevant representations and levels, I refer to Goldsmith's book *The Last Phonological Rule* (1993), which in my view, marks the transition from derivational models to nonderivational models of phonology. Inspired by the results of recent studies in language processing, Goldsmith claims that phonological models must be modified in order to account for the fact that information is processed simultaneously and not sequentially (1993: 6); in computer terminology one could say processing must be carried out digitally and not analogously. According to the findings in the fields of connectionism and neurocomputation, the task of processing is supposedly not carried out by one single central processing unit (comparable to the central processing unit of a personal computer), but by a large number of relatively simple units, which work cooperatively and competitively in parallel (1993: 7).[47] For obvious reasons a model with serial derivations, with virtually no limit on intermediate representations, and a dependence of later steps on earlier ones (rule ordering) is not compatible with the processing of phonological information in parallel. In the framework of Harmonic Phonology Goldsmith (1993: 28-33) renounces the validity of a rule-based approach to phonology and limits levels to exactly three: a morphophonemic level (M-level), a level at which the well-formedness of syllables and words is checked (W-level), and a level of broad phonetic description (P-level). He claims that

> [a]ll theories of phonology—and, more generally, of formal linguistics—can be usefully divided into theories of representations, of levels, and of rules. (1993: 22)

Goldsmith relates the major tools which can be used to address salient phonological issues—phonotactics, alternation, and contrast—to this tripartite division (refined in Goldsmith 1995). An appropriation of this classification for OT purposes would require the substitution of a theory of rules for a theory of constraints, since the universal set of constraints (CON) without doubt is OT's core component.

---

[47] Before OT was formulated as a coherent framework, concepts like the ones put forward by Goldsmith (1993) within the framework of Harmonic Phonology, which now form part of the theoretical core of OT (parallel processing by EVAL), were being developed by linguists.

(4.11) Grid with the three fundamental questions and the three basic tools of phonological theory (after Goldsmith 1995: 3)

	phonotactics	alternations	contrasts
representations	autosegmental and metrical structure; prosodic hierarchy	feature geometry, limits on kinds of assimilation; prosodic morphology	underspecification theory
levels	licensing; abstractness; structure-preservation	issues of stratal organization	organization of the lexicon
(rules) **constraints**	metrical theory; harmonic rule application; **OT**	equation of phonotactics and alternations: strict cyclicity	structure building vs. structure changing operations

4.3.1 Underspecification theories

In section 4.2.2, the issue of specification of segments was touched upon. Segments are specified (or underspecified) with respect to their featural content. Thus, the change of /t,d,s,z/ to [ʧ,ʤ,ʃ,ʒ] in SAE[48] phrases like <bet you>, <fed you>, <miss you>, and <please you> should not be analyzed as simply involving four unrelated segments. Instead, the data are best accounted for by analyzing what these segments have in common. They all share the feature [coronal]. A basis for describing such 'naturalness' in formal terms is to recognize these segments as forming a natural class in Halle's definition: "a set of speech sounds forms a *natural class* if fewer features are required to designate the class than to designate any individual sound in the class" (1961: 90). For example, the set /p,t,k/ is simpler in featural terms ([–voiced, –sonorant, –continuant]) than one single segment, e.g. /t/ ([–voiced, –sonorant, –continuant, +coronal, +anterior]). The set /ʊ,p,z/, on the other hand, cannot at all be described in terms of common feature specifications nor is it likely to appear together in any phonological process of any language; consequently, this set does not form a natural class. Segments that do occur together in a phonological process have one or more features in common, and the process is best expressed as involving these features rather than a list of segments. The two groups of segments participating in the assimilatory process above can be distinguished with the feature [anterior]: the unassimilated segments are characterized by the feature [+anterior], while the

[48] From now on, the term SAE is used in the sense of Wolfram and Schilling-Estes (1998: 282) to reflect the fact that a national variety (US) is meant and in the sense of Chafe et al. (1991: 69) that SAE encompasses formal and informal styles, and a host of regional, class, gender and ethnic varieties (cf. discussion in chapter 2 above).

assimilated segments share the opposite value [–anterior] under the influence of [–anterior] /j/. Formally, this process can be written:

$$[\text{–sonorant, +coronal, +anterior}] \rightarrow [\text{–anterior}] \;/\; \underline{\quad}\; [\text{–anterior}]$$

A further case in point that all segments which have the feature specification [+coronal] indeed form a natural class with respect to their phonotactic behavior. In SAE all 13 coronals /θ,ð,t,d,s,z,ʃ,ʒ,ʧ,ʤ,n,l,r/ (cf. also figure 4.13 below) are banned from occurring syllable-initially before the sequence /juw/, e.g. in <tune, dune, assume, new, lure> while noncoronals can cooccur with this sequence in onsets, e.g. in <beauty, music, fuse, view, cube>. Natural classes of sounds are potential candidates for underspecification.

For the sake of representational simplicity, underspecification theories see it as desirable that redundant information be omitted from URs and be filled in by structure building rules in the course of the derivation to the SRs of segments; however, it is debatable which values of which features are specified as being underlying and when in the derivation they are supplied. There are two distinct feature theories currently in use, each with its own merits: place of articulation theory (PT), which employs *SPE*-type binary features, so that [coronal], for example, defines exactly two classes, coronals and noncoronals (Clements 1985); and articulator theory (AT), which establishes a connection between the six active articulators (Labial, Coronal, Dorsal, Soft Palate, Tongue Root or Radical, and Laryngeal or Glottal) and unary or monovalent features (McCarthy 1988; Kenstowicz 1994: 145-50). It must be noted that the majority of linguists now favor some variant of the AT model (Sagey 1986; McCarthy 1988; Yip 1989; Halle 1992; Halle 1995) over the older, predominantly binary feature representation of PT.[49] According to AT, the unary features are claimed to control the movable articulators as the basic actors in speech production (cf. chapter 3.2.1 above) and more specific features, which then may be binary (e.g. [anterior], [distributed]), are dominated by particular articulator nodes. Such a representation plausibly accounts for the fact that not all features can randomly combine to produce phonological contrast nor to form natural classes, which the representations of linear phonology as unstructured feature bundles indeed suggest (cf. chapter 3.1.2 above).

[49] Cho (1991) presents evidence that for some languages (e.g. Korean) PT works better than AT, but for English I think there is no doubt that AT is preferable to PT (cf. sections 4.3.2 and 4.3.4 below).

Archangeli (1988) very thoroughly reviews three different approaches to underspecification: Inherent Underspecification, Contrastive Specification, and Radical Underspecification (RU), the latter being the one she advocates. Inherent Underspecification refers to inherent properties of the features themselves, e.g. the fact that some features are inherently privative[50] or monovalent and thus cannot be specified for binary feature values due to their very nature (e.g. Sagey 1986; cf. section 4.3.2 below). In this sense, [coronal] is regarded as a monovalent feature in most post-*SPE* nonlinear phonology. Monovalent features are either present ([+F]) or absent ([blank]). It is claimed (Mohanan 1991: 315) that a process like assimilation, which spreads e.g. the feature [coronal], must refer to the plus specification of the feature, since missing underlying specifications would leave segments unspecified for the monovalent features. These segments would thus be unsusceptible to change. I think it is not favorable to have a binary feature [coronal] and to speak of a class of noncoronals ([–F]), since it seems unlikely that labials and dorsals, possibly together with radicals and glottals, could form a single class. Apart from the articulator features, however, I noted above that there is evidence from phonological processes that some more specific features are essentially binary. Processes like Turkish vowel harmony, for example, treat vowels as if they belonged to one of exactly two classes, either [+back] or [–back] (Halle 1992: 208). For English it can be safely claimed that segments which have a place feature other than [coronal] hardly share a single common phonetic property nor do they seem to engage in similar alternations (cf. chapter 8 below). Monovalent features can thus be regarded as inherently underspecified (Archangeli 1988: 190).

Contrastive Specification (e.g. Mester & Itō 1989) is an approach which examines the contents of URs of segments with respect to their contrastive or noncontrastive potential in language-specific inventories. Only redundant features remain unspecified. These are filled in by redundancy rules (structure building rules) of the format: [+sonorant] → [+voiced]. In English the feature [voiced] does not serve to contrast sonorants, consequently, they are unspecified for voicing, whereas obstruents must be fully specified for it due to the apparent voicing contrast among these segments. Contrastive Specification's dependence on inventories is its most

[50] The notion of privativity originates from Prague School Phonology; within this framework it refers to an opposition, e.g. /d/ vs. /t/, where one member is marked by the presence of some feature (voicing in obstruents) which the unmarked member lacks (Trubetzkoy 1958: 67).

significant weakness. As Archangeli points out, by relying on language-particular segment inventories, the focus of this approach clearly lies on the segment and not on the feature, which is however desirable. Features cannot be underspecified by reference to crosslinguistic aspects of UG, but only by virtue of their intralinguistic function (1988: 193-7). Basically both theories, Inherent Underspecification and Contrastive Specification, are theories of redundancy in that only redundant information is left blank (Paradis & Prunet 1991: 21).

RU (Kiparsky 1985; Archangeli 1988; Kiparsky 1995), on the other hand, as the term implies, is a more radical approach to underspecification. It can be regarded as a form of MT because all universally unmarked, predictable feature values are removed from URs, and only universally marked values are allowed in URs. Universal rules provide the default specifications for the unmarked values. Feature specifications are present in URs only if a default rule assigns the 'wrong' value. As already mentioned, the feature [voiced] distinguishes obstruents but not sonorants, since all sonorants are redundantly voiced. Sonorants are thus crosslinguistically unmarked for voicing in the lexicon. This is expressed in formal terms as a context-free markedness rule which prohibits sonorants from being specified for voicing in the lexicon: *[+sonorant, αvoiced] (Kiparsky 1985: 92). The default rule supplying the predictable feature value has the format: [ ] → [αF], where α is either '+' or '–', in this case: [ ] → [+voiced]. Markedness in terms of formal properties which account for the crosslinguistic distribution of elements (cf. section 4.2.1 above) is incorporated directly into the model of RU via UG. Specifications are of two types: universal markedness statements and language-particular statements. A segment system is more highly valued to the extent that it employs feature specifications that are universally given, i.e. left blank in the language-specific grammar, because these need not be specified in the individual grammar.

Paradis & Prunet (1991) and the contributors to their book on coronals present an overwhelming amount of empirical evidence regarding the phonotactics of coronals, their involvement in alternations, and their overall frequency. On the basis of these data they argue that of all articulators [coronal] is the unmarked or predictable or default articulator, i.e. that coronal segments should be represented as unspecified for place or placeless: [øplace] → [coronal]. Before I move on to the question of whether underspecification is a suitable form of representation for OT, I would briefly like to comment on the general frequency of coronals. Their properties in distribution and alternations will be discussed in chapters 6 and 8 below.

Paradis & Prunet distinguish between "inventory frequency", "typological inventory frequency", and "occurrence frequency" (1991: 11). Considering the English consonant inventory[51], of 24 consonants more than half (13) are coronal consonants. Coronals are not only the most frequent consonant type in English but also across languages: Maddieson shows that of 317 investigated languages 316 (sole exception: Hawai'ian) have at least one coronal plosive (1984: 32), in 266 languages the coronal fricative /s/ is attested (1984: 52), and of 307 languages with at least one nasal 304 have the coronal nasal /n/ (1984: 69). It has been long known that coronals are also the most frequently occurring consonant type in English:

> [A]mong the consonants, dental and alveolar articulations account for about two-thirds of all consonant occurrences […]. Most of the frequently occurring consonants have an alveolar place of articulation and are distinguished by their manner of articulation. […]. The frequency-of-occurrence tables also show that […] six consonants, /t,n,s,d,l,m/, [are responsible] for half of all consonant occurrences. (Denes 1963: 894)

Wang & Crawford (1960) arrived at an almost identical result in another frequency study of consonants in American and British varieties of English.[52] They found that /t,n,r,s,d,l,ð/ are the seven most frequent consonants, which amounts to about 60% of all consonant occurrences in English (1960: 136; for details cf. frequency tables in both studies). Newman et al. (2000: 300) concluded in their study that "/t/ is the most common syllable-initial consonant in English". The results of these frequency studies and the data on inventories doubtlessly support an interpretation of coronals as underspecified for place in URs. I have already mentioned the widespread albeit not uncontroversial belief among linguists that coronal unmarkedness follows from underspecification of the feature [coronal] (Kiparsky 1985; Avery & Rice 1989; Paradis & Prunet 1989). After a discussion of nonlinear representations, which should clarify the location of the coronal place node within a geometry of features, I will turn to the issue of whether UT is compatible with OT in section 4.3.4.

[51] As discussed in chapter 2, it is mostly the vowel systems which vary from one accent of English to the next.

[52] The consonant frequencies vary only slightly with respect to /r/ among American and British accents (Wang & Crawford 1960: 138).

4.3.2 Nonlinear representations

Nonlinear phonology is a cover term for any phonological model which avoids linear, segment-sized representations of phonological units. In this sense Prosodic Analysis, an approach to phonology developed by the London School linguists Firth (1948) and Robins (1957), can be regarded as an early nonlinear model. Prosodic Analysis focuses on the types of phonological phenomena which extend beyond discrete segments. These are not only 'traditional' prosodies, such as pitch, tone, stress, and intonation (cf. e.g. Murray's classification "accent", "quantity", "emphasis", and "cadence" (1795: 146) in his *English Grammar*), but also process types like vowel harmony, nasalization of vowels, and assimilation in NC-clusters (Robins 1957: 267). Robins also claims, in a way anticipating OT's now familiar renunciation of a derivational approach to language:

> It is an unsuitable metaphor to say that one sound operates at a distance over intervening sounds to exert a force on another sound, and change it from something which in fact it never was (in the words concerned) into something else.[53] (1957: 270)

Among later nonlinear developments are Metrical Phonology (Liberman & Prince 1977), Dependency Phonology (Anderson, Ewen & Staun 1985), and Autosegmental Phonology (Goldsmith 1976).

To the best of my knowledge, the demand for a multi-linear phonetic representation in terms of a "geometry" of features was first formulated by Goldsmith in his dissertation on autosegmental phonology, which deals mainly with tonal phenomena (1976: 16). Feature geometry adds, as it were, a third spatial dimension to the former two-dimensional representations (feature bundles) of standard GP, which it was intended to replace. Goldsmith adopts an articulatory perspective, from which speech is viewed as a sequence of continuously overlaping gestures, performed by the active articulators lips, tongue, soft palate, and larynx (cf. chapter 3.2 above). Within the model of standard GP, however, speech is represented as a linear succession of feature bundles which stand for discrete "orthographic" (e.g. <p-i-n>) or "psychologically real" sound segments (/pɪn/) (1976: 16). This

[53] Usually, when some linguistic theory is superseded by another, "unsuitable metaphors" are substituted for by new, more or less suitable metaphorical expressions. OT is a superb example hereof, using a metaphorical terminology which in my opinion is reminiscent of TV quiz-shows (e.g. *candidate*, *competitor*, *winner*, *loser*, *race*, *fatal violation*).

assumption is what Goldsmith calls the "Absolute Slicing Hypothesis" (1976: 17). He goes on to demonstrate the inadequacy of such representations showing that an "orchestral score" of the articulator activity (e.g. of <pin>) must include phonologically relevant phenomena, such as pitch, which are not part of the representation of individual segments[54] (1976: 17-18). Overlap and (re-)organization of features lead Goldsmith to postulate a multi-linear approach instead, in which the different features are hierarchically represented on different tiers[55] and linked to each other by association lines (1976: 23). Well-formedness conditions, such as "association lines do not cross" (1976: 27), are introduced as constraints on representations in order to avoid ill-formed outputs. The separate tiers are intended to reflect the overlap of articulatory movements and the hierarchical organization of phonological phenomena on separate levels. It follows that the units of speech, among these the features, are indeed 'autosegments', i.e. functioning as units autonomous of the other units.[56]

Goldsmith (1976) is also one of those linguists who realized that very general phonological processes may be blocked by constraints on the output. For example, the Obligatory Contour Principle[57] (OCP), which prohibits any identical elements if these are adjacent, is known to block epenthetic identical segments or vowel deletion between identical consonants (McCarthy 1988: 88; Yip 1988: 66).

Although Goldsmith (1976) laid the foundation for nonlinear representations, the first elaborated developments in feature theory appear in Clements (1985) and Sagey (1986). The geometric organization of features as essentially independent properties of the segment (autosegments) radically differs from the *SPE* concept of segments as homogeneous, unstructured feature matrices. It is curious to note in this context that Chomsky & Halle had in fact anticipated a hierarchical structure in line with the subdivisions of features shown in chapter 3.1.2 above (figure 3.1):

[54] Cf. also *SPE*'s laborious and in the end inadequate analysis of English stress phenomena, for example the numerous stress assignment rules (1968: 69-162).

[55] Goldsmith indicates that his notions of tiers and of hierarchical structure are in fact much closer to the concepts of Prosodic Analysis than to the representations of standard GP (1979: 203-4). Prosodic Analysis postulates larger-than-segment "phonematic units" (Firth 1948: 253), e.g. syllables (CV-units).

[56] The concept of tiers also appears in Kahn's (1976) work of about the same time. Together with Hooper's (1972) proposal to reintroduce the syllable as a legitimate unit into GP, Kahn's postulation of a CV-tier was also a major step in this direction.

[57] For an in depth discussion of the OCP cf. chapters 6.3.1 and 8.3.2 below.

> It seems likely […] that ultimately the features themselves will be seen to be organized in a hierarchical structure which may resemble the structure that we have imposed on them for purely expository reasons. (1968: 300)

Embedded in a geometry, features are now represented separately from segments within a hierarchical tree structure comprising multiple levels of domination (tiers).

(4.12) Partial feature geometry: the Articulator Node Hierarchy (after Sagey 1986: 4)

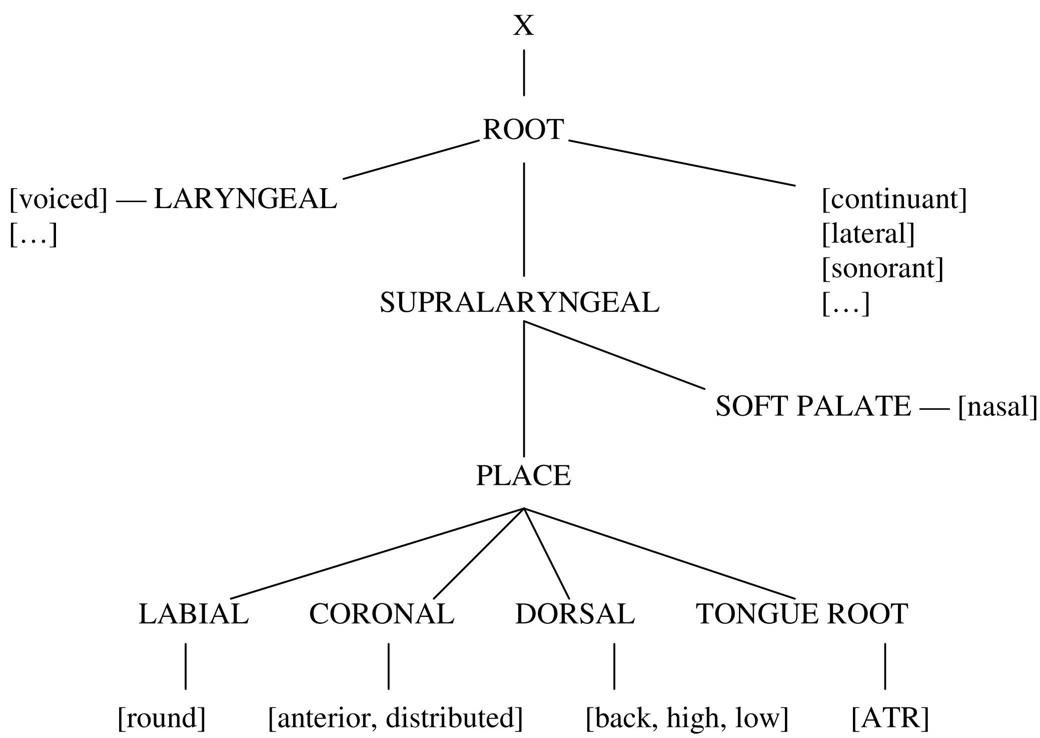


Figure 4.12 above shows a selection of terminal features, which are binary, at the lowest level of the feature tree. Most of these features are grouped into six unary class nodes: LABIAL, CORONAL, DORSAL, TONGUE ROOT, SOFT PALATE[58], and LARYNGEAL, depending on which of the six vocal tract articulators is involved in the execution of the particular feature (Sagey 1986: 205-6), i.e. these features are articulator-bound (Halle 1992: 208; Halle 1995: 16). For example, [anterior] and [distributed] are represented as dependents of the CORONAL articulator node, because the active articulator tongue blade can either reach targets in front of or behind the

[58] In addition to the five active articulators discussed in chapter 3.2.1 above (cf. also Ladefoged & Maddieson 1996: 12), Sagey has the soft palate as a sixth articulator.

alveolar ridge [±anterior], or it can produce a long stricture at the target, as opposed to a short one with just the tip [±distributed].[59] Four of the six active articulators form the PLACE constituent, which together with the SOFT PALATE node forms the SUPRALARYNGEAL constituent. The SUPRALARYNGEAL node in turn merges with the LARYNGEAL node and the articulator-free (Halle 1992: 208) features [continuant], [lateral], [sonorant], etc. to form the ROOT tier, which represents the whole segment. The top node X finally is part of the skeleton and directly embedded in the CV-timing tier.

The feature hierarchy is designed to show that each node in the tree dominates a group of features which function as a unit. Assimilation and other process types provide evidence that the features combining to form sound segments are not simple lists, but are in fact hierarchically organized. In such a representation assimilation, for example, is accounted for as the spreading of specific nodes. Clements identifies three distinct types of assimilation which directly relate to the feature hierarchy: "Single-feature assimilation" is found if one of the terminal features is involved (e.g. assimilation of the type mentioned in section 4.3.1 above in <this year> (sbc0003.wav: 07'57"): /s/ [+anterior] → /ʃ/ [–anterior]). "Partial assimilation" is the case if a class node is the spreading element (e.g. PLACE). An assimilation as in <that guy> /tg/ → [kg] (sbc0002.wav: 08'49"), for example, is represented as the spreading of an articulator node. The CORONAL articulator is delinked from the PLACE node, and the adjacent DORSAL articulator is linked in its place to the formerly coronal segment. "Total assimilation" or gemination is involved if the whole ROOT node spreads (1985: 231). Considering the nodes in the medium plane of the hierarchy, phonological processes such as voicing assimilation and aspiration involve LARYNGEAL features while leaving SUPRALARYNGEAL features unaffected.

Not only does feature geometry determine which processes are possible and which are not, since processes are now "necessary consequences of particular conditions prevailing in phonological representations" (Harris 1994: 13), but by grouping features hierarchically under class nodes, it also makes evident the fact that certain feature combinations cannot cooccur in a segment. Thus if a segment is produced with the CORONAL articulator activated, features which are dependents of the LABIAL or DORSAL articulators (e.g. [round] or [back]) cannot be present in the

[59] The contrast [±distributed] parallels the distinction between laminal vs. apical made by Ladefoged & Maddieson (1996: 10-1).

representation of the coronal segment. In addition, Sagey claims that the articulator nodes in her model reflect the different acoustic effects of the features they govern. The supralaryngeal features are said to affect the formant structure, while the laryngeal features supposedly do not; and among the supralaryngeal features the oral place features are said to influence the formant structure to a greater degree than does the nasal feature, which only adds a second resonator (1986: 5).

For the purposes of this investigation, I consider Sagey's (1986) model preferable to that of Clements (1985) because Sagey recognizes six articulator nodes mentioned above as basic, for which there seems to be abundant phonetic (cf. chapter 3.2 above) and phonological (cf. chapter 8 below) evidence. The distinction between the two feature models parallels the distinction between AT and PT mentioned briefly in section 4.3.1 above. The great merit of Clements' seminal paper "The Geometry of Phonological Features" is that the features are grouped according to their properties and arranged hierarchically on different tiers.

> If we find that certain sets of features consistently behave as a unit with respect to certain types of rules of assimilation or resequencing, we have good reason to suppose that they constitute a unit in phonological representation, independently of the actual operation of the rules themselves. (1985: 226)

However, Clements still uses the *SPE* feature set. It has a binary coronal feature, thereby classing segments in either coronals or noncoronals (PT), but it lacks the articulator/place features labial and dorsal altogether. In his feature model, the place node dominates the rather random set of terminal features: [±coronal], [±anterior], and [±distributed]. In my opinion, the grouping still reflects the more or less arbitrary fashion in which these binary features were employed in *SPE*-type phonological rules. Clements makes very explicit that the ultimate justification for a feature model must be derived from phonology, and "not from *a priori* considerations of vocal tract anatomy or the like" (1985: 230).

Sagey's AT model, on the other, hand is construed to account for the physical characteristics of vocal tract anatomy and acoustics. The model is claimed to be grounded in phonetics, reflecting the fact that the active articulators produce sound segments which form natural classes. Moreover, it is claimed that the articulator nodes dominate exactly those features which cause variation among the segments where one particular articulator is involved. Considering the English consonant inventory, only one[60] of three articulators dominated by the PLACE node is activated during the production of a consonant segment (LABIAL, *or* CORONAL, *or* DORSAL; cf. figure 4.13 below); the role of the TONGUE ROOT is restricted to the production of vowels in English.

(4.13) Distribution of the English consonants over the PLACE node

LABIAL	**CORONAL**	DORSAL
p b f v m w	θ ð t d s z ʃ ʒ ʧ ʤ n l r	k g ŋ j

As an articulator-based feature, CORONAL refers to a monovalent class node involving the tongue blade as active articulator. As such, it dominates the terminal feature [±anterior] in order to distinguish between /θ,ð/ ([+anterior]) and /ʃ,ʒ/ ([–anterior]). The articulators or articulatory gestures are dominated by the unary major class nodes PLACE and SUPRALARYNGEAL as well as by the ROOT node at the very top. An arrangement of features in a hierarchy of independent tiers linked to segmental timing slots is justified on the grounds that features behave as classes in phonological processes such as assimilation. In nonlinear models, assimilation is described as the spreading of elements mainly on one tier. There seems to be strong evidence for the fact that assimilation works exactly in this straightforward way supporting a hierarchical feature representation. Feature geometry thus provides a nonarbitrary description of assimilatory and other process types.

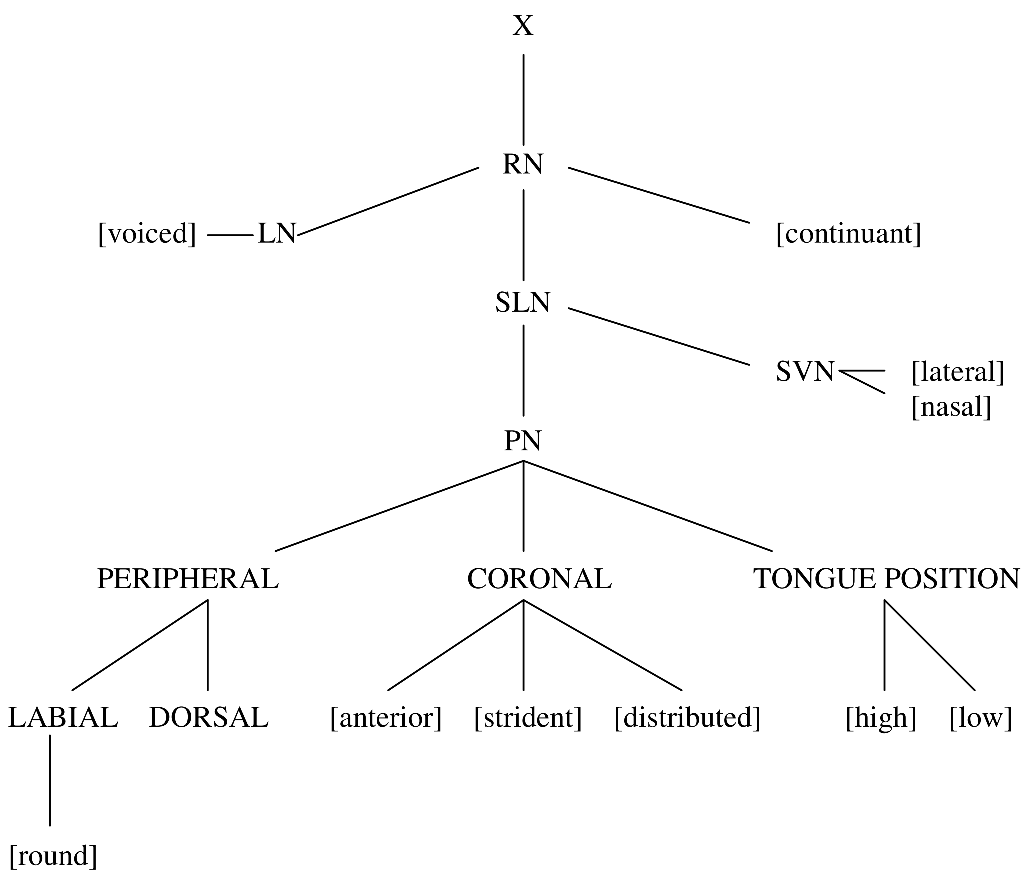
However, a closer inspection of different types of alternations reveals that establishing an absolutely adequate representation of features and nodes within the geometry is indeed a difficult task. Ever since the first proposals of geometric models (Clements 1985; Sagey 1986; McCarthy 1988), a subject of continued controversy has been the crucial issue as to which features form a constituent within the hierarchy (Paradis & Prunet 1991: 1-28). Even for a single language, it is not entirely clear whether, for instance, vowel and consonant features should be represented separately in the geometry, or whether a unified representation is appropriate for accounting for

[60] There are of course other languages in which more than one articulator is involved in the production of a single consonant. Examples are the 'palatalized' consonants in Russian, where DORSAL functions as a second minor articulator in addition to the respective major articulator, or complex consonant segments, such as Igbo labiovelar /kp/ with the LABIAL and DORSAL articulators activated simultaneously, which Sagey uses as evidence for her feature model (1986: 69-92).

their interaction in some assimilatory phenomena, e.g. the affinity between high front vowels or glides and coronals.

An interesting alternative analysis to Sagey's (1986) model is the proposal presented in Rice & Avery (1991; cf. figure 4.14 below) which is based on Avery & Rice (1989: 195). It is of significance that they introduce a new node, SPONTANEOUS VOICE, which is a sister of PLACE and a daughter of SUPRALARYNGEAL, and which dominates the two terminal features [nasal] and [lateral].

(4.14) Feature geometry (after Rice & Avery 1991: 103)



Rice & Avery distinguish two main types of node: "organizing nodes", largely equivalent to Clements' class nodes (1985: 231; cf. discussion just above), e.g. SUPRALARYNGEAL and PLACE, and "content nodes", e.g. CORONAL (1991: 104). They argue that sonorants, which are characterized by a specific *manner* of articulation, need an organizing node (SPONTANEOUS VOICE) in much the same way as obstruents have their organizing node (PLACE), which refers to the *place* of articulation. In line with Kean (1975) and Maddieson (1984), Rice & Avery claim that the coronals /n/ and /l/ are the unmarked nasal and lateral segments respectively, and specify that, compared with its neighbor [lateral], [nasal] is the unmarked SVN dependent (1991: 105).

A second point here is that the coronal articulator is seen as the default value for PLACE, while LABIAL and DORSAL together form the PERIPHERAL constituent. This representation is justified in so far as it supports the claim that coronals are unspecified for place. It has the disadvantage, however, that phonetic grounding is lost by grouping the labial and dorsal articulators somewhat arbitrarily under a PERIPHERAL node and reintroducing a binary feature [±coronal] with a class of noncoronals.However, as recent nonlinear investigations of the internal organization and possible representation of speech sounds show (e.g. Clements & Hume 1995), not only OT but also 'OT-free' phonological theory of the recent past is unified by endeavors to bridge the gap between phonological structure and phonetic interpretation, thereby eventually removing one of the principal motivations for a separate phonetic component in the grammar.

4.3.3 The Grounding Hypothesis

Another step in this direction is Archangeli & Pulleyblank's book *Grounded Phonology* (1994), which makes several important proposals concerning the structure of segments and the cooccurrence of features (F-elements) which are viewed as phonological primitives. Among their proposals is the Grounding Hypothesis, which expresses the idea that statements governing the cooccurrence of features in segments must be phonetically motivated.[61] Archangeli & Pulleyblank (1994: 19-21) use Sagey's (1986) representational model (cf. figure 4.12 above), which, as I argued in the preceding section, is preferable to that of Clements (1985), because as an AT representation it takes into account both the articulatory implementation of the features and their behavior in phonological processes. Archangeli & Pulleyblank

[61] Support for the idea that phonetic substance is of eminent importance for an adequate phonological description also comes from the distant discipline of Cognitive Linguistics. Bybee views lexical representations as cognitive entities rooted in the realities of speech perception and production. Instead of features and phonemes, Bybee regards Browman & Goldstein's (1986; 1990) complex articulatory gestures (cf. chapter 3.3.3 above) as the ideal units for the description of—as she claims—purely phonetic, connected speech phenomena like e.g. SAE flapping. In this view, such phonetic processes are due to hiding, overlap, or deformation of contiguous gestures (1994: 285-305).

confirm that Sagey's model makes "both a formal and a substantive claim about the sets of features that constitute natural classes for the purpose of phonological rules in natural languages" (1994: 23).

In formalizing this notion, they define the domain of cooccurrence statements as a *path*. Informally, a path is a "set of associated nodes, features, or prosodic categories (cf. section 4.3.5 below) such that no more than one token of any node, feature, or prosodic category is included in the set" (1994: 21). Cooccurrence conditions are defined as *path conditions*. A path condition is formulated as an implicational statement which determines whether paths between features and nodes are well-formed or ill-formed (1994: 22, 169). These implicational statements are of two basic formats (cf figure 4.15 below). Both positive and negative formulations of conditions are necessary in order to account for the substantive content of phonological representations (1994: 170).

(4.15) Formats of path conditions as implicational statements (after Archangeli & Pulleyblank 1994: 169)

a. positive: if *p* then *q*
b. negative: if *p* then not *q*

Considering again Sagey's representational model with monovalent articulator nodes (cf. figure 4.12 above), implications for English, which distinguishes single articulated consonants, have to be formulated negatively, for example: if CORONAL then not LABIAL and if CORONAL then not DORSAL. On the other hand, a language such as Russian, which has double-articulated consonants, can additionally have positive implicational statements of the type: if LABIAL then DORSAL. Path conditions that express phonetically motivated feature cooccurrence statements are *grounding conditions*. The grounding conditions which are explored by Archangeli & Pulleyblank (cf. figure 4.16 below) express not only a positive requirement but also a prohibition on the cooccurrence of features. For the exposition of their model they present grounded path conditions on the features [ATR] vs. [RTR] (position of the tongue root) and [high] vs. [low] (vowel height) (1994: 165-282). It is argued that the tongue root advancement responsible for [+ATR] is antagonistic to the low tongue body position required for [+low], and similarly, that the tongue root retraction which implements [–ATR] is antagonistic to the high tongue body position needed for [+high].

(4.16) Grounding conditions (after Archangeli & Pulleyblank 1994: 177)

I. Path conditions invoked by languages must be phonetically motivated.

II. The stronger the phonetic motivation for a path condition Φ,

a. the greater the likelihood of invoking Φ,

b. the greater the likelihood of assigning a wide scope to Φ within a grammar,

and vice versa.

While a combination of the features [+ATR] and [+low] is not completely incompatible, the path condition does, however, predict that it will be disfavored because these two features are physiologically antagonistic. Grounding conditions may thus prevent the spreading of [+ATR] onto a segment specified for [+low] or may even cause the deletion of a segment. Conversely, the features [+ATR] and [+high] make a perfect match because the articulation of a [+high] sound with an advanced tongue root is effortless from a physiological point of view.

(4.17) Grounded path conditions: ATR/RTR and vowel height (after Archangeli & Pulleyblank 1994: 176)

a. ATR/HI condition: If [+ATR] then [+high]; if [+ATR] then not [–high].
b. ATR/LO condition: If [+ATR] then [–low]; if [+ATR] then not [+low].
c. RTR/HI condition: If [–ATR] then [–high]; if [–ATR] then not [+high].
d. RTR/LO condition: If [–ATR] then [+low]; if [–ATR] then not [–low].

Such sympathetic feature relationships, Archangeli & Pulleyblank claim, may be involved in spreading and epenthesis. It must be noted that their claim seems to be accurate for some types of alternation in SAE, namely epenthesis and neutralization as well: here the segments that are inserted and that contrasts are neutralized to indeed have grounded path conditions. For spreading, however, the claim that grounded path conditions may prevent the spreading of an antagonistic feature cannot be upheld uncontroversially because in SAE place assimilation, for example, it is precisely the sympathetic feature relationships, i.e. the unmarked coronal stops—both oral and nasal—which are made to disappear in spreading, in that coronals undergo these alternations and turn into somewhat disfavored feature combinations (cf. the detailed analysis in chapter 8 below).

I think that grounded path conditions (similar to the formulations presented in figure 4.18 below) are well-suited for my analysis of the distributional and alternational patterns of coronal consonants in the SAE data taken from the *SBCSAE* (cf. chapter 2.1).[62] I also think that the basic concepts of the Grounding Hypothesis are consistent with an OT analysis for the following reasons: First, condition I (cf. figure 4.16 above) roots feature cooccurrences in phonetic substance by making an absolute claim about the physical relationship between two F-elements:

> The phonetic relation may be sympathetic or antagonistic, it may be physiologically or acoustically motivated—but it must exist. A path condition not expressing such a phonetically motivated relation may not be invoked in any natural language. (Archangeli & Pulleyblank 1994: 177)

This is in line with the claim in favor of an articulatory and auditory implementation of the features and their behavior, which I argued for in chapter 3. It is also the upshot of the model of Articulatory Phonology (Browman & Goldstein 1986; 1989; 1990; 1992; cf. chapters 3.3.3 above and 7.3.2 below), and it seems to coincide with recent developments in OT (Lombardi 2001a; cf. section 4.3.4 below). Second, condition II (cf. figure 4.16 above)—contrasting with the absoluteness of condition I—expresses general tendencies (crosslinguistic and intralinguistic) via implicational statements which Archangeli & Pulleyblank classify according to their strength value (cf. figure 4.17 above) into "very strong", "strong", "medium", and "weak" path conditions (1994: 179). In order to illustrate the scalar nature of such implicational statements, I will briefly discuss the positive formulations of potential path conditions shown in figure 4.18 below. Condition a. is an absolute universal as there are no nonsonorant nasals in any language inventory. In OT terms, the constraint NAS/SON ('a nasal must be a sonorant') would be undominated in all language-particular rankings of CON.

[62] What exactly constitutes an adequate representation for OT analyses is still an open question in current work in OT (cf. section 4.3.4 below).

(4.18) Decreasing strength value for grounded path conditions

a. If [+nasal] then [+sonorant].
b. If [+sonorant] then [+voice].
c. If [+obstruent] then [–voice].
d. If [+consonantal] then [+coronal].

Condition b. is no doubt very strong but the constraint SON/VOI ('a sonorant must be voiced') may be violated by individual languages that have voiceless sonorants in their inventories, though certainly not by English (cf. section 4.3.1 above). Thus condition b. is a constraint that is undominated in most grammars. And finally, the medium to weak conditions c. and d. express the general tendencies, e.g. based on frequency measurements, that voiceless obstruents and coronal consonants are both more natural and less marked than voiced obstruents and noncoronal consonants, although violations of these two conditions are quite common. Within an OT framework condition c. can be formalized as a MARKEDNESS constraint on the cooccurrence of features OBS/VOI ('an obstruent must be voiceless') while condition d. can be represented as a constraint on feature cooccurrences: *PL/DOR , *PL/LAB >> *PL/COR (Prince & Smolensky 1993: 181; cf. section 4.3.4 below), which is violated in language after language (Maddieson 1984).

The Grounding Hypothesis is thus largely a formalization of markedness considerations focusing on the interaction of F-elements in hierarchies of strong and weak path conditions. Languages will prefer strong (unmarked) conditions to have the widest distribution possible. However, it is perfectly conceivable that a weak, but grounded, condition will surface in a language as a marked case (Archangeli & Pulleyblank 1994: 179). This obviously leaves room for scalar phenomena and thus provides an ideal, theoretically motivated frame for the establishment of certain constraints under OT.[63] It has been observed in OT literature (Prince & Smolensky 1993: 67-82) that a particular constraint may always work hand in hand with another constraint, that their energetic effects are, as it were, conjoined to achieve a specific

[63] In the concluding chapter of their book, Archangeli & Pulleyblank (1994: 405-27) actually carry out some of their analyses concerning the feature [ATR] within a still preliminary OT framework. What they tentatively call the "optimization hypothesis" is thus defined: "[G]rammars prefer relations between representations that are defined in *formal* and *grounded* terms as optimal: synchronic and diachronic processes should gravitate towards optimal relations." (emphasis added; 1994: 406)

output goal. Grounding conditions require certain features to cooccur on the grounds of phonetic motivation. The same is true of constraint conjunctions, which have a synergistic effect and capture scalar phenomena (McCarthy 1997: 238). My investigation incorporates the Grounding Hypothesis to isolate what is actually found in SAE from patterns that are unattested.

4.3.4 Representations in OT

In the years immediately before the rise of OT, there were various proposals for phonological representations, all of which share a common goal. This is the attempt to resolve the controversial claims which specific types of representation make. In the sections above, I have presented different proposals for underspecified, nonlinear, and phonetically grounded representations. Although UT and nonlinear representations brought about an improvement of phonological theory by encoding restrictions in the representations themselves, they were criticized for the nonuniversality of their predictions. The different versions of UT especially all reached a similar theoretical impasse. The problem lies in constraints on representations that are regarded as absolute and inviolable well-formedness conditions, only relativized in terms of specific grammars (parameter setting; cf. section 4.1.1 above). Such a theory has only one option for the resolution of conflicts which inviolable antagonistic principles create: to organize these on different levels of the derivation (early vs. late) as their respective domains (e.g. lexical phonology; cf. also section 4.3.5 below). Representations thus have redundant values in some contexts but not in others. Underspecification paradoxes of this type can be demonstrated extremely well with the example of coronals, as will be shown below. It seems only too plausible that a theoretical framework like OT, which holds that all phonological constraints are universally present at all levels and violable in principle, holds the promise of showing a way out of this dead end. Before I comment on OT's attitude (or rather *non*attitude) towards phonological representations, I will present some recent criticism of RU and nonlinear representational models.

RU is a theory which is aimed at expressing the asymmetries of feature values in distribution and alternation by devising URs of segments that may not contain any predictable information (cf. section 4.3.1 above). I would like to pick out two important pre-OT critiques of RU (Mohanan 1991; Steriade 1995) which reveal the formalism of RU to be unsuited for an account of actual language data.[64] Both authors substantiate their claim by drawing on data from English phonotactics, assimilation (nonlinear spreading), neutralization, epenthesis, deletion and lenition. In English, coronals commonly assimilate to noncoronals without noncoronals assimilating to coronals. Within the framework of RU this fact is captured by postulating the feature [coronal] as unmarked (or nondominant) and by representing coronals as unspecified for place in URs: /n/, for example, then has the underlying specification [+nasal], while /m/ has [+nasal, –coronal]. Assimilation is seen as a structure-building process in that marked[65] (or dominant) feature values spread to neighboring segments which have not yet received a specification for place. Yet, as Mohanan shows (1991: 310-9), lexical regularities (MSCs)—such as the illicitness of the English onsets */tl, dl/ (versus /pl, bl/ and /kl, gl/) or quite contrarily, the illicitness of *non*coronal homorganic nasal plus voiced plosive coda (/nd/ versus */mb, ŋg/), and the diphthongs /aw/ and /oj/ which can be followed only by a coronal coda (<mouth, mouse, point, coin> versus */awp, ojk/)—require coronals to be specified for place in URs. A further case in point is the example mentioned in section 4.3.1 above, which shows that all 13 coronal consonants in SAE are disallowed before /juw/ in onsets. This phonotactic restriction treats the whole class of coronals, marked ones (/ʧ,ʃ/) and unmarked ones (/t,s/) alike[66], as a natural class. Reference to this class is impossible unless the feature [coronal] is specified in URs. If, however, coronals are specified for place, assimilation can no longer be regarded as a structure-building process, as RU assumes, but must be seen as a structure-changing process (Mohanan 1991: 315). This way much of the desired economy of description of underspecified segments is lost. Moreover, SAE lenition of /t/ (/t/-

[64] Mohanan also shows that cooccurrence restrictions on features can be expressed either as rules, or as filters, or as implicational constraints (1991: 290). Considering again an example from section 4.3.1, that in English sonorants are redundantly voiced, three notations are possible:

a. [+sonorant] → [+voiced]	(*SPE*-type structure-building rule)
b. *[+sonorant, αvoiced]	(RU-type filter)
c. If a segment is [+sonorant], then it is [+voiced]	(implicational statement)

By presenting an alternative to the common (obsolete) rule notation, the formulation in c. in a way anticipates the feature cooccurrence conditions (path conditions) introduced by Archangeli & Pulleyblank (1994) which are also formulated as implicational statements (cf. figure 4.16 above).

[65] Even outside the framework of RU, this peculiarity in the behavior of coronals constitutes a constant difficulty when trying to present a unified account. They are unmarked and frequently occurring segments for one thing and for another assimilation makes them disappear.

[66] It is generally assumed that the default value for anteriority in coronal consonants is [+anterior] (and not [–anterior]), which results in /t,d,n/, traditionally called alveolars, being unmarked coronals (Kean 1975: 23; Paradis & Prunet 1989: 321; Paradis & Prunet 1991: 6; McCarthy & Taub 1992: 364).

voicing) and the frequent deletion of coronals cannot be explained with segments that are not specified for place. Mohanan concludes that some unified mechanism must be found which is capable of accounting for the special behavior of coronals in all mentioned processes (1991: 319); this is a mechanism unlikely to be of a derivational nature (Steriade 1995: 166).

The targets of McCarthy & Taub's (1992) criticism are the nonlinear representational models brought forth in Paradis & Prunet's (1991) collection of articles on coronals. On the grounds of their exceptional distributional properties and their behavior in alternations, coronals are frequently represented as placeless within nonlinear frameworks, i.e. they either lack the CORONAL articulator node (Avery & Rice 1989: 179-200) or even the superordinate PLACE node (Paradis & Prunet 1989: 319; Paradis & Prunet 1991: 6); CORONAL is then supplied by default:

ØPLACE → CORONAL

Incorporating ideas from UT, Paradis & Prunet (1991: 1-28) present an enormous amount of empirical evidence for the claim that coronals are unmarked and should be represented accordingly. Coronals have, for example, an exceptionally high typological, inventory, and occurrence frequency (cf. also section 4.3.1 above); they are frequently inserted as epenthetic segments and are more susceptible to place assimilation than other consonants. Their special behavior allegedly follows from their lower featural complexity as compared to other segments. This supposedly accounts for their high overall frequency and their proneness to epenthesis and assimilation. However, as McCarthy & Taub point out, the notion of placeless coronals is problematic when dealing with coronals which need a specification for one of the terminal features dependent on CORONAL (1992: 364).[67] Specifically the high inventory frequency of coronals (e.g. 13 coronal consonants in the English inventory) necessitates at least some CORONAL dependent features for the distinction of the different coronal segment types. Considering the feature geometries in figures 4.12 and 4.14 above, [anterior] and [distributed] (Sagey 1986; Avery & Rice 1989; Yip 1989) and possibly also [strident] (Paradis & Prunet 1991) are needed as CORONAL dependents. Returning to the example given above, it is obvious that, in

[67] This problem has previously been noted by Avery & Rice (1989: 179-196), Paradis & Prunet (1989: 321), and Yip (1989: 355) who state that coronal segments can only be unmarked for place features if there is no contrast in the class of e.g. coronal obstruents, i.e. if no CORONAL dependent feature is needed to distinguish two coronals. The solution offered by them is that only the unmarked dentals and alveolars, i.e. [+anterior] coronals, may be represented lacking a place specification.

order to distinguish between marked and unmarked coronals, at least /ʧ,ʃ/ must be specified [–anterior] while /t,s/ share the default value [+anterior]. If, however, either the PLACE node or the CORONAL articulator is completely lacking in the representation of coronals, no feature dependent on CORONAL can be specified in the hierarchy of features, i.e. the segments cannot surface distinctly. But of course, if a segment is specified for either terminal feature, it must also be specified for the superordinate node in the feature geometry (McCarthy & Taub 1992: 364).[68]

As a result of the dissatisfaction with GP's inadequate rules (cf. Goldsmith's (1993) *The Last Phonological Rule*) and questionable representations (cf. Mohanan 1991; McCarthy & Taub 1992; Steriade 1995), the general criticism of generative phonology was expanded in Prince & Smolensky's (1993) seminal study and new framework (*Optimality Theory: Constraint Interaction in Generative Grammar*). OT does not acknowledge lacking ARTICULATOR and PLACE nodes or unspecified place features in URs, since URs (or inputs) are not subject to any type of restriction (cf. section 4.1.2 above and chapter 9.2 below for critical remarks). Moreover, rules that fill in features or supply nodes and serial derivations simply do not exist in OT. Quite the contrary is true, OT introduces two families of FAITHFULNESS constraints (PARSE and FILL) which function to ensure faithful mappings between input and output candidates under evaluation of CON:

(4.19) The two basic families of FAITHFULNESS constraints (after Prince & Smolensky 1993: 181)

PARSE$^{\text{FEAT}}$: An input feature must be parsed into a root node.
FILL$^{\text{PL}}$: A PLACE node must not be empty (unassociated to any features).

PARSE, for example, requires input features to be present in the surface form, thus militating against deletion of feature specifications, while FILL demands PLACE nodes to be filled in the input form, thus militating against unspecified objects in inputs. The phonological inactivity of predictable features, such as CORONAL, is thus attributed to rankings of a specific type of constraint, rather than to the absence of

[68] In addressing this problem, OT departs from feature geometry by postulating that the constraints which refer to feature classes are gradiently violable in the sense given within OT (Prince & Smolensky 1993). This then results in partial class behavior, where only a proper subset of the relevant features can participate in spreading, for example, for reasons of markedness (e.g. Padgett 1995b). Cf. chapter 5 below for an in depth discussion.

features from the representation at some stage of a derivation. Prince & Smolensky do, however, subscribe to the notion that CORONAL is the unmarked place of articulation. They propose to capture "coronal unmarkedness" in the generalization: "don't have a place of articulation other than CORONAL" (1993: 180). This naturally leads to the following constraint ranking of a selection of MARKEDNESS constraints concerning feature cooccurrences (1993: 181):

*PL/DOR , *PL/LAB >>[69] *PL/COR

The constraint prohibiting coronals is lowest in the dominance hierarchy indicating that coronals are unmarked for place, at least among the supraglottal places.[70]

Precisely the interaction of these two antagonistic types of constraint, one favoring richness of specification and one favoring the opposite, causes Itō, Mester & Padgett (1995) to attempt to salvage some of the concepts of UT from doom in OT. In line with OT doctrines, they argue for phonological underspecification of segments, not, of course, as an input property but as an output property enforced by feature licensing constraints (1995: 573). Licensing is understood in Goldsmith's sense such that all segments must be part of a higher-level unit, whereby certain units are licensers which license a set of phonological features (1990: 123-7). For example, Itō, Mester & Padgett's licensing approach is intended to establish a connection between feature redundancy and underspecification (i.e. between Contrastive Underspecification and Radical Underspecification; cf. section 4.3.1 above). If one feature implies another and is thus predictable or redundant, predictability excludes the implied feature from being licensed or unspecified (1995: 580). Returning to the example of redundantly voiced sonorants in English, it turns out that a sonorant does not license the feature [voiced], because a segment-internal condition demands that all segments which have the feature [sonorant] must also have the feature [voiced]. This is taken care of by the redundancy implication: [sonorant] > [voiced] (cf. also Mohanan 1991: 290 in note 34 above), or in another notation by the high-ranking constraint: SONVOI (1995: 581). The sonorant voicing constraint may be in conflict with a feature licensing constraint, in this case with the constraint:

---

[69] A comma between two constraints in the notation means that they are not ranked with respect to each other whereas ">>" symbolizes strict domination of the preceding constraint over the following.

[70] This place hierarchy is modified in Lombardi (2001b: 29) to include a lowest ranked constraint against the appearance of glottal stops, which in Lombardi's view are the truly unmarked consonants: *PL/DOR , *PL/LAB >> *PL/COR >> * PL/PHAR.

(4.20) Feature licensing constraint (after Itō, Mester & Padgett 1995: 581)

LICENSE(VOICE): the feature [voiced] must be licensed.

An obstruent, on the other hand, does license the feature [voiced]. Since voicing in obstruents is contrastive, no redundancy condition can lead to a cancelation of licensing. If one assumes the unmarked feature value for obstruents in English to be [–voiced], the RU-type filter that obtains in the language *[+obstruent, +voiced] translates into the following licensing statement: in a specific segment, the feature [–voiced] is licensed by the presence of the feature [+obstruent] in that segment.

As a matter of fact Itō, Mester & Padgett's proposal appears more like a formalization of markedness than of underspecification, which is then in accordance with OT assumptions. They claim that licensing constraints can account for degrees of underspecification, i.e. contrastive or radical. An example of this is said to be the behavior of coronals which is attributed to a conflict between FAITHFULNESS constraints, requiring the presence of feature specifications and licensing constraints. These constraints belong to the family of MARKEDNESS constraints, which prohibit the presence of redundant features (1995: 608). Depending on how FAITHFULNESS constraints and MARKEDNESS constraints interact, segments will be either specified or unspecified for certain features. Beyond these more or less general claims, no analysis is offered.

Leaving these issues for further discussion in chapters 8 and 9 below, one can say that the basic tenets of OT make no particular claims about the nature and form of phonological representations, since OT is a model of the interaction of constraints and not of the representation of segments (McCarthy 1997: 231; Oostendorp & Hermans 1999: 4; Rose 2001: 27). And indeed, if one considers that only two separate representations are involved in the evaluation of a phonological form (input and output representation; cf. section 4.1.2 above), it makes perfect sense that OT analysts readily tone down the issue of representation. Some authors go even further and argue that OT annuls the need for any URs, i.e. for a lexicon: under these assumptions, all the work is done by constraints (e.g. Flemming 1995; 1997; Hammond 1995; Russell 1995).

OT is said to work with any type of featural representation, whether these are binary feature bundles or monovalent geometries (Lombardi 2001a: 3). UT and nonlinear phonology on the other hand (cf. sections 4.3.1 and 4.3.2 above) have taken great pains to devise representations which account for actual language data as adequately as possible. Analysts working within these models usually stress the issue that representations are in fact theory-internal constructs. Considering autosegmental phonology for example, it is clear that a theory which assumes a hierarchical organization of features on different tiers needs representational devices which match these claims (cf. Goldsmith's example of an "orchestral score" (1976: 17) of articulator activity). In contrast to this, much early work in OT has tacitly used nonlinear representations and never even commented on the question of what sort of representation is suitable for OT purposes (Prince & Smolensky 1993; McCarthy & Prince 1993; McCarthy & Prince 1995b). In response to OT's original indifference to representations, Lombardi now argues in favor of "correct representations" (2001a: 3-4) that ought to include rich phonetic detail in order to gain additional insights into phonological patterns and alternations when working within an OT framework (cf. also Hayes 1995 for an early proposal of this claim). An obvious problem connected to using rich and detailed phonetic representations is that phonetic phenomena tend to be frequently optional, gradient, noncontrastive, and not grammatically conditioned while phonological phenomena commonly are obligatory, categorical, contrastive, and subject to grammatical conditioning. The danger of including too much phonetic detail in representations lurks at the point where it seems justified to question whether this is still the realm of phonology. Thus the real challenge is to ground phonology in phonetics without going so far as to represent merely phonetics. This is an issue which will recur as a subject of discussion in all subsequent chapters.

4.3.5 Phonological domains in OT

Another important question is which phonological domains are needed for OT analyses, and more specifically for an investigation of coronals using OT. Constraint subhierarchies like the one mentioned in the preceding section (*PL/DOR , *PL/LAB >> *PL/COR >> * PL/PHAR), which is a formulation of markedness statements regarding place features, express implicational universals. At the same time they refer to phonological domains, such as the feature, the segment, the syllable, etc. A problem frequently addressed in recent OT literature is the virtually unrestricted rise of the number of constraints, some of which make reference to very specific phenomena (Hammond 1999: 28; McMahon 2000: 78). This means the domain in which they are active is very local (cf. further critical remarks in chapter 8 below). What makes the constant accumulation of constraints problematic is the fact that the whole set of constraints provided by UG, i.e. CON, is assumed to be universal and universally present at all times by definition (cf. section 4.1.2 above). Consequently, ranking by strict domination becomes all the more difficult the more constraints there are to be ranked for any specific grammar.

A sensible organization of phonological domains is the proposal by Selkirk (1984a), who claims the following minimal set of categories to be of relevance for a linguistic description of English:

(4.21) Hierarchical structure of phonological domains (after Selkirk 1984a: 26)

a. intonational phrase (IP)
b. phonological phrase (PhP)
c. prosodic word (PW)
d. foot (Ft)
e. syllable (Syl)

For the purpose of this investigation, this organization is supplemented by LaCharité & Paradis' (1993) prosodic hierarchy, who include structure below the syllable as well. A similar proposal is made by McCarthy (1984: 301).

(4.22) The universal Phonological Level Hierarchy (after LaCharité & Paradis 1993: 142)

a. syllable (or: syllable tier)
b. syllable constituents: onset, rhyme, peak, coda (or: onset-rhyme tier)
c. timing slot (or: CV-template tier)
d. segment (or: melody tier)
e. nonterminal feature (or: feature tier)
f. terminal feature (or: feature tier)

As will be shown in the following chapters, a Phonological Level Hierarchy (PLH) is of prime importance for OT accounts of distributional and alternational phenomena. Not only do the constraints themselves refer to phonological domains, but OT subtheories (e.g. Generalized Alignment, McCarthy & Prince 1993; Positional Faithfulness, Beckman 1999; cf. chapter 6.3.1.1.1 below) also deal—at least implicitly—with the issue of levels. For example, there is evidence that syllable structure is in fact inherently present, i.e. in the linguistic objects generated by GEN and subsequently evaluated by EVAL. Golston & van der Hulst (1999) claim that unsyllabified strings of segments are not valid phonological representations, just as representations in which a syllable node dominates a prosodic word node can never be valid linguistic objects (cf. the discussion of the Strict Layer Hypothesis in chapter 6.2.2.1). To qualify as linguistic objects, the so-called "representational primitives" (Prince & Smolensky 1993: 4) produced by GEN must be syllabified (Golston & van der Hulst 1999: 153-4). The intermediate level of the syllable constituents (onset, rhyme, peak, and coda), which I added to the list, also play a crucial role in OT analyses, as will be shown in chapter 6.3 below.

The prosodic hierarchy proposed by Selkirk has certainly been adopted for OT research (e.g. McCarthy & Prince 1995a: 320). There is, however, to the best of my knowledge at least, no generally accepted hierarchy of categories or levels binding for OT analyses.

4.4 SUMMARY

After having presented these major difficulties with rules and representations, I think it should be apparent that in the early 1990s the time was ripe for a new, optimized framework—OT. Nevertheless, it can be safely stated, as I have done, that OT is not quite as new as it is often claimed to be.

CHAPTER 5

Theoretical premises for an analysis of the behavior of coronals under OT

5.1 INTRODUCTION

On the basis of a crosslinguistic comparison, coronals turn out to be exceptional consonants in many ways, both with respect to their distributional properties and their special behavior in alternations. In chapter 4.3.1 above, I presented data and analyses confirming that coronals are the most frequent consonant type with respect to three established frequency types (Paradis & Prunet 1991: 11). Their inventory frequency is usually higher than of any other consonant type in a given language; the same is true of their typological inventory frequency since coronals attested in the universal segment inventory outnumber all other consonant types (cf. Maddieson 1984); and lastly, the occurrence frequency in language-specific inventories also shows an overwhelming predominance of coronals, e.g. for English half of all consonant occurrences (Wang & Crawford 1960; Denes 1963), with /t/ being the most common syllable-initial consonant in English (Newman et al. 2000: 300). This unequivocal evidence has been the factual basis for regarding coronals as default consonants, as unmarked for place, and even as unspecified for place (cf. chapters 4.3.1 and 4.3.2 above).

5.2 THE PROBLEMATIC NOTION OF PROCESS UNDER OT

That coronals are essentially different from other consonants is also obvious when looking at phonological alternations or processes. In this context, the question immediately arises whether it is legitimate to speak of a *phonological process* at all when working within an OT framework since the theory explicitly bans all serial procedures and derivational processes whatsoever (Prince & Smolensky 1993: 205). The term *process* is closely linked in conceptual content to the term phonological rule, and is thus readily associated with older, derivational models of GP. A more or less abstract input is altered step by step to derive a phonetic output and, by doing so, one passes through a series of intermediate representational levels.[71] It is obvious that an operational process notion like the one just described cannot be appropriated for an investigation conducted in an OT framework.

It does, however, seem justified to use the term if it is understood in the sense of OT's *one* process of evaluation of a given candidate set which leads to the selection of the optimal output candidate. The evaluation process is controlled by the function EVAL, which implements the language particular ranking of constraints (CON) for its task. In this definition, process actually refers to the one single step of parallel processing of CON by EVAL[72] (cf. figure 4.2 in chapter 4.1.2 above). More precisely, the workings of EVAL which give rise to phonological processes can be described as a specific constraint ranking, one in which some MARKEDNESS constraint (M) crucially dominates some FAITHFULNESS constraint (F): M >> F. Taken together, the family of MARKEDNESS constraints, each of which is universal, constitutes the universal theory of markedness in OT (for a critical discussion cf. chapter 9.6.2 below). The structural constraint M militates against some output structures, for example against distinct articulators in NC clusters (homorganic place assimilation) or voiceless obstruents between vowels (/t/-voicing in SAE), whereas the opposite ranking F >> M leaves, as it were, everything as it is. The input candidate is selected as the optimal output candidate, and no process takes place (McCarthy 1997: 234). Process thus implies that some change to the input form has occurred. Consequently, phonological phenomena such as assimilation, epenthesis,

[71] It may well be the case that OT's reluctance about determining certain representations as OT-internal representations (cf. chapter 4.3.4 above) has something to do with their conceptual closeness to rules. Cf. Anderson (1985: 9-10): "[T]heories of rules and theories of representations deal with intimately interrelated and indissoluble aspects of the same linguistic structure."

[72] This notion of process is sometimes critically described as having a "derivational residue" or "derivational effect", a condition which is inappropriate for OT, because the parallel processing of all relevant data must take place in one step without resorting to (serial) derivation (Oostendorp & Hermans 1999: 3; LaCharité & Paradis 2000).

and deletion will retain the label *phonological process* in the following chapters, however in the sense of OT's strict definition.[73]

In the absence of phonological rules which apply to very specific targets in the phonology of a language (e.g. /t/-voicing in SAE), the question has been raised as to how specific phenomena, which formerly were regarded as the output of a phonological rule, could be handled in OT (Davis 1995). McCarthy's paper (1997) offers an answer by proposing process-specific constraints. What he actually means is not constraints that apply to specific processes, i.e. specific input-output mappings, but a ranking for process-specific constraint interaction. In order for a phonological process to take place, some structural constraint(s) $\mathbb{M}$ must be ranked above the relevant $\mathbb{F}$ constraint(s). McCarthy assumes a further structural constraint $\mathbb{C}$ to intervene between two MARKEDNESS constraints $\mathbb{M}_i$ and $\mathbb{M}_j$. This ranking results in a mitigation of the activity of $\mathbb{M}_j$ by $\mathbb{C}$, while $\mathbb{M}_i$ is not affected by $\mathbb{C}$:

(5.1) Ranking for process-specific constraint interaction (after McCarthy 1997: 237)

$$\mathbb{M}_i >> \mathbb{C} >> \mathbb{M}_j >> \mathbb{F}$$

"The constraint $\mathbb{C}$ […] is 'process-specific' to the $\mathbb{M}_j >> \mathbb{F}$ interaction, but specificity is achieved through language-particular ranking rather than language-particular parameterization." (ibid.)

The crucial issues which I will address in the following are: What are the constraints that act as MARKEDNESS constraints and FAITHFULNESS constraints? Is it possible to show that these constraints are motivated in such a way that they are grounded in physical facts of language production and perception? Do the diverse constraint rankings result in any explanatory gain for the observed phenomena of Standard American English, for the class of coronals, or for the theory?

[73] The notions assimilation, epenthesis, deletion, etc. of course carry a processual meaning element. An exclusion of the term *process* from OT terminology would thus imply that terms like assimilation, etc., too, could no longer be used.

5.3 THE SIGNIFICANCE OF THE PERCEPTUAL DOMAIN FOR PHONOLOGY

Much of phonetic and phonological research is characterized by a bias in favor of the production side of speech sounds, often to the exclusion of descriptions of perceptual impressions produced by speech sounds. The general focus on articulation and vocal anatomy can be traced all the way through the history of linguistics, from the Sanskrit grammarians to 19th century linguists like for example Sievers (1881; cf. also chapter 3.1.1). This tradition, which focuses on articulation, constitutes the basis for the structuralists' approaches to phonetics and phonology as well (e.g. Pike 1943; 1947; cf. also chapter 3.2.1). So-called 'explanatory' notions referring to vocal tract anatomy, such as the 'principle of economy', 'ease of articulation', and 'ease theory', also have a long tradition in phonetics and phonology (Whitney 1868: 28; Sweet 1888: 49; Passy 1890: 227; Sapir 1921: 183; Jespersen 1922: 261-4; Zipf 1949: 56-133; Martinet 1955: 94-152). In chapter 3.1.2, I showed that most features—including [coronal]—which are used in contemporary phonological models are defined in articulatory terms.

However, I also mentioned that a closer look at the *SPE* feature set, for example, reveals its hybrid nature since some of the features, e.g. [sonorant] and [strident], in fact suggest a perceptual basis and some, e.g. [continuant] and [voiced], can be defined equally well both in articulatory and perceptual terms. I showed that there is an abundance of evidence for the class of coronals from articulation-based research (cf. chapter 3.3.3), while there is not nearly as much evidence from perception-based studies (cf. chapter 3.3.2). Although the results of Stevens' work (1994; 1997; 1998) point in this direction, invariant acoustic correlates of the class of sounds that all share the feature [coronal] are yet to be discovered. Moreover, it has yet to be shown that the human perceptual system is indeed sensitive to such physical cues—as transmitted between mouth and ear—which in their sum would sufficiently characterize the class of coronals thereby excluding any other sound which is not a member of this class.

It is quite obvious that an investigation of the articulatory movements, e.g. aided by electropalatography (EPG), is much easier to conduct and will yield more robust results than an investigation of the perceptual responses to speech sounds, as mediated by ear, auditory nerve, and brain. Auditory phonetics deals with the processes of receiving and decoding the speech signal. Hearers need to process acoustic cues present in the speech sounds and make use of their knowledge of the sound patterns of their language. This knowledge enables them to make the necessary abstractions from all phonetic detail contained in the signal and to—ideally at least—arrive at what was encoded by the speaker. Jakobson, Fant & Halle, whose universal set of acoustic features I presented in chapter 3.1.1, convincingly argue that the perceptual domain is in fact the one which is most relevant to phonology:

> The closer we are in our investigation to the destination of the message (i.e. its perception by the receiver), the more accurately can we gage the information conveyed by its sound shape. […] Each of the consecutive stages, from articulation to perception, may be predicted from the preceding stage. Since with each subsequent stage the selectivity increases, this predictability is irreversible and some variables of any antecedent stage are irrelevant for the subsequent stage. (1951: 12)

Difficulties are encountered with the identification and measurement of neurological and psychological responses to speech sounds and few efforts have been made in phonology to construct models which represent the psychological and perceptual properties of sound patterns (Ohala 1997: 682; but cf. Stevens 1998). This is the research area which deals with the attributes of sounds that enable abstraction and are thus responsible for the perception of distinct sound categories instead of a continuous flux of sound (Sapir 1921: 56; Taylor 1995: 222-38). That this perceptual effort must be made by the hearer in order to decode the signal and eventually arrive at the meaning of the message has been known to linguists for a considerable length of time (Bühler 1934; cf. other references below). To the best of my knowledge, however, no attempt has been made in the field of phonology to formalize perceptual phenomena as a perception-based grammar model or even only as part of a grammar model, comparable to, for example, the theory of Articulatory Phonology (cf. chapter 3.3.3) or the model of autosegmental spreading in Nonlinear Phonology (cf. chapter 4.3.2), with their unambiguous foundations in the articulatory domain of speech production.

The traditional notions of the phoneme (Baudouin de Courtenay 1895; Sapir 1921; Bloomfield 1933; Twaddell 1935; Trubetzkoy 1939; Jones 1950; Hill 1962; Schane 1971 for a proposal to integrate the phoneme into GP; and Taylor 1995 for a concept of the phoneme modeled on prototype theory) owe their existence to linguistic attempts at filtering out redundant phonetic detail from the sounds *heard* in speech and reduce them to distinct, contrastive units of sound—the phonemes.[74] Needless to say, all attempts to attribute any sort of physical reality to phonemes, apart from their obvious psychological reality[75], have invariably failed.

Despite the problems tied to the objectification of perceptual properties of sound patterns, I wish to stress the role which perceptual representations and functional principles connected to speech perception ought to play if both an adequate description and a sensible overall representation of speech phenomena are the goal. This can only be achieved by clearly separating the roles of speech perception and speech production and by trying to put audition on an equal footing with the role articulation plays in the phonology of a language.

Bühler, for example, stresses this point in his influential book *Sprachtheorie* (1934), which took decades to be rediscovered after the end of World War II. Linguistics, he says, should not content itself with the analysis of linguistic expressions on the side of the speaker, instead the analysis of linguistic appeals from the perspective of the hearer constitutes an equally important field of linguistic investigation because appeals do in fact have an effect on the hearer and may cause a certain reaction (1934: 32). Elaborating on the basic functions of the linguistic sign in communicative processes, Bühler adds a third function to the expressive and appellative functions, which is on equal terms with the other two. The referential function accounts for the fact that in a concrete act of communication speaker and hearer talk about *something*, they refer to concepts[76] by using linguistic signs. The three-way relationship that shows the linguistic sign as a *symbol* with respect to its indirect relationship with the concepts, as a *symptom* with respect to its direct relationship with the sender, and as a *signa*l with respect to its direct relationship with the hearer is captured in the shape of a triangle—the so-called Organon model of communication:

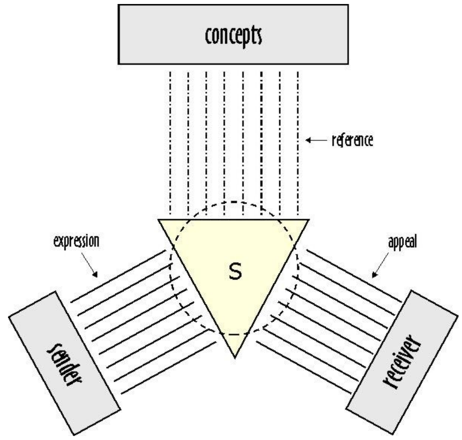
---

[74] The postulation of phonemes as units of phonology and specific phoneme inventories for all the different languages of the world in turn gave rise to the setting up of the *IPA* alphabet in 1888 and so-called 'phonemic transcription'.

[75] Evidence for the psychological reality of phonemes also comes from language games such as Pig Latin (e.g. <ig-pay atin-lay>) or speech errors (e.g. <blake fruid> for <brake fluid> Fromkin 1971: 31): "By far the largest percentage of speech errors of all kinds show substitution, transposition (metathesis), ommission, or addition of segments […]." (Fromkin: 1971: 30) This claim is confirmed by a recent study by Dell (1995), who also finds phonemes to be the units which are most frequently involved in speech errors (35%; followed by words with 33% of the cases investigated).

[76] Bühler himself uses the terms "Gegenstände und Sachverhalte" (1934: 24-33). Following insights of modern psycholinguistics (e.g. Aitchison 1994: 41), I prefer the notion that human beings, when they talk, refer to concepts of entities and not to the entities themselves.

(5.2) The three prime functions of the linguistic sign (S) shown in the Organon model of communication (after Bühler 1934: 28)



The expressive function of the linguistic sign corresponds to the productive processes of articulation, which have been the focus of phonetic and phonological investigations for some time. The other two, however, are equally important. Applied to phonology, it is the appellative function of the linguistic sign that figures in processes of speech perception. The third function links the linguistic sign to the meaning concept which must be encoded, for example in the form of a phonological representation (UR), and subsequently is decoded by UR. A UR then leads to the recognition and categorization of a given linguistic item (cf. figure 7.11 in chapter 7.3.1 below).

In the following, I briefly want to present some of the assumptions made by various linguists with respect to the function of speech perception. One of the linguists who also insisted on the importance of the role of speech perception and the psychological nature of the representation of speech sounds was Edward Sapir.

> In watching my Nootka interpreter write his language, I often had the curious feeling that he was transcribing an ideal flow of phonetic elements which he heard, inadequately from a purely objective standpoint, as the intention of the actual rumble of speech. (1921: 56)

He said that if one wanted to learn anything about the patterning of sound, it was not enough to look at just one side of the coin:

> Many linguistic students have made the fatal error of thinking of sound change as a quasi-physiological instead of as a strictly psychological phenomenon, or they have tried to dispose of the problem by bandying such catchwords as "the

> tendency to increased ease of articulation" or "the cumulative result of faulty perception". […] These easy explanations will not do. "Ease of articulation" may enter in as a factor, but it is a rather subjective concept at best. (1921: 183)

Sapir is referring to two conflicting forces in language, the functional notion 'ease of articulation' and its perceptual counterpart potential perceptual confusion. These inherently opposing tendencies were formulated well before Sapir in a very clear fashion by Paul Passy as two fundamental principles of language, the "principle of economy" and the "principle of emphasis" respectively:

> [D]es tendances phonétiques que nous avons constatées, se dégagent bien nettement deux principes fondamentaux:
> **1. Le langage tend constamment à se débarrasser de ce qui est superflu**.
> **2. Le langage tend constamment à mettre en relief ce qui est nécessaire**.
> [C]e principe d'économie est très actif et d'une application universelle. C'est de lui que dérivent la faible accentuation des syllabes peu importantes, […] l'assimilation plus ou moins complète de deux sons consécutifs; […].
> Le deuxième principe, que nous pouvons appeler *principe d'emphase*, est moins généralement reconnu, […]. Son action n'est pourtant ni moins certaine, ni moins universelle que celle du principe d'économie. […] C'est au principe d'emphase que sont dus l'aspiration et […] les dissimilations […].
> Au fond, les deux principes, d'économie et d'emphase, dérivent d'une seule et même cause, celle qui motive l'existence du langue lui-même: la nécessité de se faire comprendre. **On parle pour être compris** […].
> Bien entendu […] il n'est pas question de conservation volontaire des éléments importants, d'omission volontaire des éléments superflus. Si je néglige un élément important, on ne me comprend pas, je me corrige, et sans doute j'exagère; si je néglige un élément superflu, on me comprend bien, et je recommence. Voilà tout.
> Malgré leur origine commune, les deux principes sont naturellement en lutte constante […]. (1890: 227-9)

For principle number one, Passy names Henry Sweet's formulation of the "two principles of economy" as his source:

> (a) dropping of superfluous sounds, […];
> (b) ease of transition from one sound to another, which leads to convergence and assimilation, […]. (1888: 49)

Whitney had identified two similar tendencies in language even twenty years earlier:

> [F]irst, to make things easy to our organs of speech, to economize time and effort in the work of expression; second, to get rid of irregular and exceptional forms, by extending the prevailing analogies of the language. (1868: 28)

The innovational aspects of Passy's claims are, first of all, that he adds to the principle of economy (henceforth *minimization of articulatory effort*) a second opposing force called the principle of emphasis (henceforth *maximization of perceptual distinction* or *minimization of perceptual confusion*), secondly, that both principles can be deduced from the primary function of language, which is the need of human beings to communicate, and thirdly, that the two principles by their very nature are in constant conflict. Passy thereby anticipates OT's insistence on the conflicting forces in the phonology of a language. The realization of these two equally valid principles of language can of course be expressed extremely well in a grammar model which has a set of conflicting constraints as one of its core functions.

In more recent times, it is again the phonological model which appears to be a predecessor of OT in many respects—Natural Phonology (cf. chapter 4.2.2)—which stresses the role of perception for the analysis of speech processes:

> [Natural Phonology's] basic thesis is that the living sound patterns of languages, in their development in each individual as well as in their evolution over the centuries, are governed by forces implicit in human *vocalization* and *perception*. […] the implicit phonetic forces are manifested through processes, in the sense of Sapir—mental substitutions which systematically but subconsciously adapt our phonological intentions to our phonetic capacities, and which, conversely, enable us to perceive in others' speech the intentions underlying these superficial phonetic adaptations.
>
> (Donegan & Stampe 1979: 126; emphasis added)

5.4 FUNCTIONALISM VS. FORMALISM

Although OT is a self-declared descendant of generative phonology and shares with it the claim to being a model of the human innate knowlegde of language, it is evident that OT has closer theoretical affinities to some of the function-based hypotheses of Natural Phonology than to the overly formal mechanisms of *SPE* (cf. also chapter 4.2 where I presented a critique of *SPE*'s one-sided, formalistic approach to phonology and outlined some of the theorems NP and OT share). These affinities become even more apparent when looking at NP's criticism of standard generative phonology: NP focuses on the conflicting goals of speech perception and articulation (cf. also Dressler 1984: 31) and emphasizes the role of a functionalist basis for a linguistic theory aiming at explanatory adequacy:

> The discrepancy between the sound perceived and intended, and the sound produced, is simply phonology. This tension between clarity and ease is one of the most obvious, and oldest, explanatory principles in phonology. Modern theories, however, to the extent that they incorporate analoguous principles,

> tend to make them monolithic, like the principle of distinctiveness in structuralism or simplicity in generative phonology. […] In that framework, positing conflicting criteria would be like pitting Ockham's razor against an anti-Ockham who multiplies entities as fast as the razor can shave them off: it would defeat their purpose of evaluating alternative analyses. But an evaluation criterion, necessarily monolithic, cannot replicate conflicting explanatory principles. (Donegan & Stampe 1979: 130)

SPE is indeed focused on the formal simplicity of language-specific rules and on an evaluation metric which considers features independent of their "intrinsic content", that is, independent of their substantive, phonetic, or functional properties (Chomsky & Halle 1968: 296-9, 400). *SPE* admits only data which qualify as "internal" evidence (Kenstowicz & Kisseberth 1979: 140-54; Kenstowicz 1981: 131). In Zwicky's words, these data comprise "(1) variant shapes of morphemes and (2) distributional restrictions on phonological elements" (1972: 154). What these definitions of internal evidence refer to is basically equivalent to the methodology of the structuralists, i.e. the setting up of rules that account for alternations and the distribution of sounds with as few exceptions as possible. Additional data, which are commonly referred to as "external" evidence (Kenstowicz & Kisseberth 1979: 154-75; Kenstowicz 1981: 131), are not taken into account in the standard versions of generative phonology, since this methodology requires the adduction of independent theories (other than linguistic theories), e.g. of vocal tract physiology, of speech production and perception. Given the nonavailability of external explanations to standard GP, this lack was compensated for by the construction of numerous grammar-internal formal devices, which subsequently led to various paradoxes.

Precisely external data, however—data obtained by considering e.g. acoustic, articulatory, and auditory phonetics, typological surveys of inventories and processes, dialect variation, historical change, language acquisition, speech errors, and aphasia, to name just a few (cf. Zwicky 1972: 154-5 for a more exhaustive list)—are the ones favored by NP[77], because these data count as natural evidence—due to e.g. the universality of the human vocal tract shape—the evaluation of which supposedly leads to the creation of a truly explanatory theory (Donegan & Stampe 1979: 128). Phonological explanations should thus account for at least two opposing forces, "clarity"—the enhancement of perceptual distinctiveness aimed at the hearer,

[77] For a general discussion and defense of external evidence cf. Singh (1988).

and "ease"—the reduction of articulatory effort from the point-of-view of the speaker.

As mentioned in chapter 4.2 above, efforts to supplement *SPE*'s formalism with phonetic substance and function-based explanations for the phenomena in question were undertaken almost immediately (e.g. the theory of universal markedness in *SPE*'s chapter 9). However, in addition to markedness theory and NP, there have been quite a few other functionally oriented proposals. Kiparsky's paper "Explanation in Phonology" (1972) is one such noteworthy attempt at discussing the role of functional conditions in linguistic theory from within the framework of GP. He points out that certain controversial devices of a strictly formal generative model, e.g. variable rules, rule conspiracies, and rule ordering paradoxes, can be attributed at least in part to general functional conditions, such as the need for "distinctness of categories" and the "optimalization of phonotactic structure" (1972: 222).

> [I]t seems plausible to assume that category distinctness derives from perceptual requirements, and optimalizes language for the hearer by providing more clues to meaning in surface structure. Phonotactic conditions presumably optimalize language for the speaker by keeping down the need for executing complex articulatory movements. (ibid.)

An important question Kiparsky raises is whether such performance conditions should be stated as part of the grammar of a specific language (e.g. English), i.e. part of linguistic competence and thereby as grammar-internal properties, or whether they are the result of universal functional conditions on speech performance that have nothing to do with the peculiarities of English phonology. Kiparsky does not answer this question but he is clear on the subject that seeming paradoxes arise from the interaction of conflicting functional forces: the dominant functional factor in a specific speech situation wins.[78] His conclusion contains the following far-sighted remark:

> I have suggested a way in which the concept of a "tendency", which lends functionalist discussions their characteristic unsatisfactory fuzziness, can be made more precise in terms of hierarchies of *optimality*, which predict specific consequences for linguistic change, language acquisition, and universal grammar. Enormous areas of vagueness obviously remain. But there is enough to show that the project is a worthwhile one. (1972: 224; emphasis added)

[78] Kiparsky assumes that telephone talks are characterized by the importance of the distinctness factor due to limited audibility, while phonotactic factors gain the upper hand in face-to-face rapid speech (1972: 223).

Another proposal stressing the importance of a phonetic basis for phonological description, i.e. the grounding of phonology in phonetics is Lindblom (1986; 1992). He conceives of sound systems as dependent on the principles of maximal and sufficient contrast: when languages select their distinctive sound categories, these two principles have to be kept in equilibrium (cf. chapter 6.2 below).

Mohanan (1993) claims that the ultimate explanations for the principles of UG can only be found by appealing to substantive universals, to "nonlinguistic systems that shape the nature of UG, including the human articulatory and auditory systems, and the requirements of human communication" (1993: 110). After discussing a great variety of generative phonology's formal devices, such as structure building rules, structure changing rules, linking rules, constraints, repair strategies, blocking, and structure preservation, he arrives at the conclusion that the antagonistic patterns of distribution and alternation are better captured in terms of a metaphor that refers to the gravitational effects on physical objects moving through space.

> [T]he phonological patterns of natural languages can be explained as results of phonological fields of attraction [which] find an explanation in what I called phonetic fields of attraction of articulatory ease and perceptual discrimination. (Mohanan 1993: 110-1)

These *fields of attraction*, which interact dynamically in the language and give rise to regularities that are not without exceptions, are subdivided into phonological fields of attraction that account for language-specific patterns of distribution and alternation, e.g. assimilation, and into phonetic fields of attraction that are viewed as the physical motivation for the competing forces of minimization of articulatory cost and maximization of discriminability (1993: 98).

In chapter 4 above, I presented Archangeli & Pulleyblank's (1994) Grounding Hypothesis, which is yet another attempt to enrich phonological theory with phonetic substance. The Grounding Hypothesis is concerned with the structure of segments and claims that statements governing the cooccurrence of features in segments must be grounded in phonetics. Archangeli & Pulleyblank also state unequivocally that the phonetic relation between the features may be sympathetic or antagonistic, that it may be *physiologically* or *acoustically* motivated, but that it must definitely exist (1994: 177; cf. also chapter 4.3.3 above).

In their groundbreaking work of 1993, which marks the advent of OT, Prince & Smolensky envision an imaginary *SPE*-grammarian who criticizes OT for its loss of content in comparison with its abundantly formal generative counterpart *SPE*:

> Appeal to scalar constraints—degrees of well-formedness—leads inevitably to a *functionalizing* narrative mush of the 'better for this/better for that' sort. By means of such push-pull, any imaginable state-of-affairs can be comfortably (if hazily) placed in a best of all possible worlds. (1993: 197; emphasis added)

In answer to this, Prince & Smolensky fault *SPE* for being a grammar model with "a more-or-less arbitrary assortment of formal rules, where the principles that the rules subserve (the 'laws') are placed entirely outside the grammar, beyond the purview of formal or theoretical analysis, inert but admired" (1993: 198). It is indeed true that *SPE*'s much cited chapter 9 (markedness theory) has never really lost its status of an appendix to the grammar, since as a theory of universal markedness regarding the featural make-up of segments, it supplies the external causes for the grammar-internal rules. If the *SPE*-position of the preceding chapters 1 – 8 is applied in its strict sense, this is an illicit procedure. MT thus remains outside the grammar by necessity. OT, on the other hand, can pride itself on integrating markedness into the grammar as one of its core formal devices, the M (MARKEDNESS) family of constraints, which is strictly opposed to the F (FAITHFULNESS) family of constraints. Prince & Smolensky's reply to the criticism consequently demands a reassessment of an essentially formalist position and makes room for functionalist concerns:

> If phonology is separated from the principles of well-formedness (the 'laws') that drive it, the resulting loss of constraint and theoretical depth will mark as major defeat for the enterprise. (1993: 198)

These claims are underlined by contemporary developments within the field of Cognitive Linguistics. Bybee (1994), for example[79], advocates a view of phonology that arises from the application of functionalist principles to phonology, such as considering the phonetic substance of phonology rather than just its structural aspects (1994: 287). More precisely, this means that lexical representations are believed to consist of real phonetic substance in production as well as in perception. With respect to production, Bybee is thinking of phonetic alternations which are motivated by articulator activity, e.g. SAE flapping of [t] (1994: 287) and which can be stated

[79] Cf. also Werker (1995) for another cognitive study stressing the importance of investigating perceptual phenomena.

in the framework of Articulatory Phonology of Browman and Goldstein (1986; 1989; 1990; 1992; cf. also chapter 3.3.3) as reduction, overlap, or hiding of articulatory gestures. As far as perception is concerned, Bybee envisions auditory representations of linguistic units which are stored as actual tokens in much the same way as visual representations of natural objects, possibly also organized into schemas:

> Each new perceptual instance of an item—whether the input is from another person or from self-monitoring—is superimposed on existing representations and will gradually change the representation.[80] (Bybee 1994: 292)

In OT terms, one could say that "new perceptual instances" are in fact created by the GENERATOR when producing a candidate set for some input form. Among other possible candidates in the set, such a new candidate may eventually turn out to be the winner if it is the candidate that best satisfies CON, i.e. the language-specific ranking of constraints. Due to CON, a speaker should be able to evaluate any candidate, including the ones that are made up by some other speaker. Consequently, if anyone came up with a better candidate than the currently top ranking one, this candidate should be recognized by speakers as the optimal form and gradually diffuse throughout the speech community. This assumption of course raises questions about the functioning of GEN, and I will come back to this issue when discussing the controversial status of GEN in chapter 9.3 below.

As Prince & Smolensky (1993; cf. the quotatation just above) noted early on, OT is capable of integrating both formal *and* functional requirements in the grammar without falling short on either side. It is important to note that functional postulates, such as *maximization of perceptual distinction* and *minimization of articulatory effort*, merely express physiological tendencies, because there is no physiological *necessity* that compels either phonological systems or individual speakers to observe them:

---

[80] A study by Lehmann on the change of American English /t/ undertaken in the 1950s supports Bybee's view. He detects clear cases of hypercorrection in American school children, e.g. when reciting the following poem (1953: 275):

Kitty, my pretty white kitty,
Why do you wander away?
I've finished my work and my lessons,
And now I'm read[t]y for play.

Lehmann argues that hyperforms like <read[t]y> are caused by a "phonemic shift" taking place and working its way through the SAE lexicon using the mechnism of analogy until speakers of American English can no longer distinguish "the allophones of their *t* and *d* phonemes when these occur in voiced surroundings in noninitial position" (1953: 271-5). Ohala expresses a very similar view in his paper on a possible perceptual basis of sound patterns (1995: 90-1).

> Though the physical constraints shaping speech sound behavior are universal, their influence on languages is probabilistic, not absolute, because there are often ways that they can be overcome. (Ohala 1997: 687)

For example, there is a crosslinguistic bias against voiced obstruents (ibid.) or nasal plus voiceless obstruent clusters (Pater 1996) but nevertheless many languages, including English, have them in their inventories or syllable make-ups respectively. Similarly, a consonant may be inserted to enhance perceptibility, but not necessarily so; a consonant may be assimilated to simplify the task of production, but not necessarily so. A one-sided highlighting of functional principles is just as undesirable for phonology as an overemphasis on formal criteria of evaluation.

> Pure functionalism is to phonology what natural selection is to genetics: it can supply compelling post hoc explanations for why things ended up the way they did, but it cannot tell us how many peas from this particular plant are going to be yellow or green, round or shriveled. (McCarthy 2002: 221)

From this statement, one can deduce that a functionally oriented phonological theory can only achieve predictive power if an account of the substantive teleology is combined with the rigidity of description of a formal grammar. As conflicting tendencies, these forces of language are particularly well suited to be expressed as FAITHFULNESS constraints and MARKEDNESS constraints respectively, and with a universal set of violable constraints, ranked in a language-specific (or variety-specific or speaker-specific) way, OT meets the conditions for a formal grammar.

5.5 TREATMENTS OF PHONOLOGICAL CONTRAST

It is common ground in phonological theory that from the mass of phonetic detail contained in the speech signal, only a very small subset of the gradient phonetic properties is substracted to form the contrastive sound material of a given language, which is then relevant for phonological analysis.

5.5.1 Standard GP representation of contrast

In standard generative phonology, this situation is dealt with by excluding the nondistinctive phonetic properties from higher levels of representation, where the abstract *and* categorical representations of the contrastive sound units are placed. A consistently formal and categorical approach to contrast will run into difficulties, when elements that are contrastive in some cases (e.g. voiced plosives and voiceless plosives in English), turn out to be noncontrastive in other instances. Such a treatment of contrast is schematically shown in figure 5.3 below.

(5.3) The standard representational treatment of contrastiveness (after Kirchner 1998: 60)

underlying representation	(ideally) pure representation of contrast
↓	
PHONOLOGICAL COMPONENT	noncontrastive properties may be filled in, particularly if contrastive in other languages
↓	
PHONETIC COMPONENT	remaining noncontrastive phonetic properties, including gradient values, filled in
↓	
phonetic representation	representation of all speaker-controlled phonetic properties of the utterance

Since contrast is regarded as an either-or issue in all models of generative phonology prior to OT, scalar phenomena, such as the voicing of voiceless consonants (fully voiced or slightly voiced) or conversely the devoicing of voiced consonants (fully voiceless or slightly devoiced), which are prone to be represented as continua, cannot be sensibly treated within such frameworks. For example, contrast with respect to the feature [voiced] is frequently suspended in SAE coronal plosives in certain contexts. If they appear in the vicinity of other voiced segments and at the onset of an unstressed syllable (e.g. <writer, winter, shelter, sorted, little, get it>; Kahn 1976: 56-61), the plosives are typically flapped (i.e. voiced). The scalar nature of these voicing phenomena is underlined by Oswald in his paper "'Voiced T'—A Misnomer" decades before Hooper (1972) and Kahn (1976) reintroduced the syllable into GP as a unit relevant for the description of phonological alternations.

> [A]t least three perceptibly different voiced sounds occur to represent the stop in words like *latter* and *ladder*: a mere flip of the tongue, a brief tap, and a well articulated voiced stop. Obviously, however, these three 'sounds' represent mere points in a continuum […]. Now there is nothing about such a sound, phonetically speaking, which would lead an objective observer to conclude that it is a kind of [t]. (1943: 24-5)

In the standard approach, first a contrastive underlying representation for /t/ ([–voiced] as in <writer>) and for /d/ ([+voiced] as in <rider>) must be assumed (Kahn 1976: 56), and then a rule must be invoked requiring the neutralization of this contrast under certain circumstances (Kahn 1976: 60).

(5.4) A constraint-ranking treatment of contrastiveness (after Kirchner 1998: 63)

underlying representation	no restrictions on this level of representation
↓	
PHONOLOGICAL COMPONENT GEN + EVAL	contrastive, categorical behavior of particular features falls out from constraint system
↓	
phonetic representation	representation of all speaker-controlled phonetic properties of the utterance

In OT on the other hand (cf. figure 5.4 above), contrastiveness emerges directly from the interaction of violable constraints without any reference to serial derivations. As there are no restrictions on inputs (with the exception of the very general restriction that inputs must consist of linguistic objects), phonetic detail (and not exclusively contrastive elements; cf. discussion of underspecification theories in chapter 4.3.1 above) may be included in underlying representations ad libitum. An integration of constraints that refer directly to such phonetic properties, which doubtlessly play a role in both distribution and alternation, e.g. the conditions on production and perception, works in a straightforward way and yields a much better account of scalar phonetic phenomena.

5.5.2 Correspondence Theory and contrast

Correspondence Theory as a subtheory of OT provides a general framework for the statement of constraints which require faithfulness to linguistic objects in the relations between input and output strings: "Given two strings S_1 and S_2, correspondence is a relation $\mathscr{R}$ from the elements of S_1 to those of S_2." (McCarthy & Prince 1995: 14) It is assumed that each candidate pair S_1-S_2 comes from GEN[81] equipped with a record of input-output disparity which is then evaluated by CON for

[81] For a more in depth discussion of the role of GEN in OT cf. chapter 9.3 below.

the candidates' faithfulness to their inputs (McCarthy & Prince 1995: 14-5). Correspondence Theory expands on the older version of OT where the sound inventory of a language is "the set of segments found among the optimal *output parses* for all possible inputs" (Prince & Smolensky 1993: 179). This task is handled by two types of FAITHFULNESS constraints, PARSE and FILL (Prince & Smolensky 1993: 181; cf. also chapter 4.3.4 above). The former prohibits deletion by requiring each input element to be faithfully parsed into its respective output form, while the latter prohibits empty segments in inputs which might then surface as epenthetic segments.

PARSE and FILL are substituted for in Correspondence Theory by correspondence constraints, which assess correspondence and identity of correspondent elements between phonological strings. MAX–IO supplanting PARSE, militates against deletion in that it states that every element of some specific input form must stand in correspondence with its output form (S_1 is *maximally* present in S_2). DEP–IO, formerly FILL, prohibits epenthesis by requiring every output element to have a correspondent in its input (S_2 is *dependent* on S_1). McCarthy & Prince introduce a third type of input-output (IO) correspondence constraint which plays a crucial role in the grammar: IDENT–IO(F) demands identical specifications between input and output strings with respect to some feature F. This constraint type thus militates against all sorts of assimilation. Conversely, crucial domination of one or more IDENT–IO(F) constraints will lead to featural disparity between S_1-S_2, and to an alternation like assimilation (1995: 16-7). IO correspondence constraints, which secure an element's underlying specifications to surface faithfully in the output, have the crucial function of guaranteeing contrast in the sound system of a language. If all MARKEDNESS constraints were allowed to dominate all FAITHFULNESS constraints in CON of some language, all outputs would be something like /ta/ or /pa/ in that specific language, and thus not suited for communication.

A feature F is contrastive in a grammar if there is an IDENT–IO constraint on F which outranks some MARKEDNESS constraint on the surface distribution of a specific feature value. If on the other hand, IDENT–IO(F) is ranked below the MARKEDNESS constraint banning a particular value of a (binary) feature, then contrast with respect to this feature is neutralized. Consider as an example the neutralization of contrast in the pair <writer> vs. <rider> in SAE.[82]

(5.5) Low-ranking IDENT–IO(F) results in F [voiced] being not contrastive (SAE)

/t/	*[–voiced]	IDENT–IO(F)
☞ [+voiced]		*
[–voiced]	*!	

/d/	*[–voiced]	IDENT–IO(F)
☞ [+voiced]		
[–voiced]	*!	*

In both tableaux of figure 5.5 above, irrespective of voiceless or voiced inputs, the voiced candidate wins due to a high-ranking MARKEDNESS constraint banning the voiceless coronal stop from the surface. A violation of the low-ranking correspondence constraint is not fatal in this ranking.

The tableaux of figure 5.6 below exemplify the situation for Standard British varieties, where there is a contrast in this context in the pronunciation of the pair <writer> vs. <rider> with respect to the feature [voiced]. Both faithful candidates are selected as the optimal candidates, since the nonfaithful candidate in the first tableau incurs a fatal violation of the high-ranking FAITHFULNESS constraint.

(5.6) High-ranking IDENT–IO(F) results in F [voiced] being contrastive (RP)

/t/	IDENT–IO(F)	*[–voiced]
[+voiced]	*!	
☞ [–voiced]		*

/d/	IDENT–IO(F)	*[–voiced]
☞ [+voiced]		
[–voiced]	*!	*

The question is left open why it should be the case in SAE that the FAITHFULNESS constraint IDENT–IO(F) is ranked below a MARKEDNESS constraint milititating

[82] Since the tableaux in figures 5.5 and 5.6 are intended to offer merely a schematic description, the specific environment of the neutralization (e.g. syllable onset) has not been taken into account; for an analysis of actual assimilation data from SAE cf. chapter 8.2.1 below.

against voiceless coronals in an all-voiced environment, and why should the ranking for Standard British be just the other way around. OT's answer to these questions is that even within a single 'language' there are numerous, variety-specific rankings of the constraints in CON, and thus, even if only slightly, differing grammars (Golston & Wiese 1995; Cassimjee & Kisseberth 1998; 1999; Moren 1999).

5.5.3 Dispersion Theory and contrast

An alternative to Correspondence Theory is the Dispersion Theory of contrast introduced by Flemming (1995; 1997). Flemming takes as the basis for his assumptions Lindblom's notion of maximal and sufficient system contrast (1986; Lindblom & Maddieson 1988) and implements their analysis of the language-specific selection of contrastive sound categories into the constraint-based grammar model of OT. According to Flemming, the selection of sound inventories, which express phonological contrasts, is subject to three functional goals:

(5.7) The goals in the selection of phonological contrasts (after Flemming 1995: 15)

1. Maximize the number of contrasts.
2. Maximize the distinctiveness of contrasts.
3. Minimize articulatory effort.

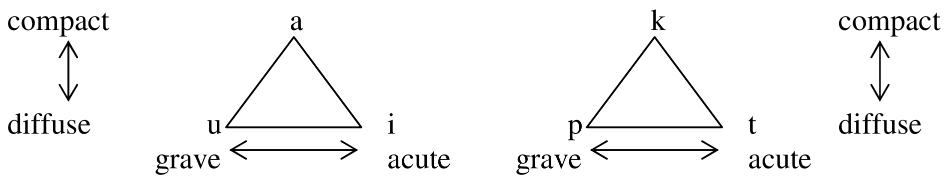
These goals can easily be traced all the way back to Passy's (1890) principles of *maximization of perceptual distinction* and *minimization of articulatory effort* (cf. chapter 5.3 above). Goals two and three were previously formulated by Jakobson as the "principle of maximal contrast" (1949: 300) and the "principle of least effort" (1949: 295), both with respect to the emergence of sound systems in child grammars and the distribution of sounds in the inventories of the world's languages. According to Jakobson, those consonants and vowels appear early in language acquisition and frequently in typological terms which—syntagmatically and paradigmatically—offer a maximum degree of perceptual distinction (cf. figure 5.8 below). This realization ultimately led to the development of Jakobson et al.'s (Jakobson, Fant & Halle 1951; Jakobson & Halle 1956) rigorously formulated feature theory and the decomposition of vowel and consonant phonemes into their fundamental oppositions: 12 binary, acoustically defined features were employed to account for the universal inventory of human speech sounds (Jakobson 1979: 292; cf. also chapter 3.1.1 above).

Jones' fixation of the cardinal areas for sounds (1960: 31-43; cf. also chapter 3.3.1 above) nicely illustrates the workings of goal number two at the expense of goal number three—a situation schematically shown in figure 5.9 b. below. Considering Jones' vowel diagram with eight cardinal vowels (1960: 36), these 'ideal' sounds display a dispersion in the available articulatory and perceptual spaces, where the distance between the individual vowels is as large as it potentially can be and where a relatively high degree of articulatory effort is involved. All eight vowels are produced with "tongue-positions as remote as possible from 'neutral' position" (1960: 37). This results in perceptually and acoustically maximally distinct reference points, which Jones uses for the description of the language-specific vowel inventories. Looking at the vowel systems of natural languages, and more specifically at the distribution of front vowels in a perceptual space containing a four-term height contrast, the best candidates with respect to categorical perception are /ɪ, e, ɛ, æ/ (Boersma 1998: 179). These candidates, which optimally exploit the perceptual space and yield four equally strong perceptual categories, are roughly the ones found in the SAE system of lax front vowels.

The system of contrasts is also believed to play an essential role in motivating sound change in cases of asymmetries in segment inventories, where the distribution of segments is such that the available articulatory and perceptual spaces are not optimally exploited (Martinet 1968: 475-85; cf. also figures 5.7 and 5.8 below). The dynamic forces in systems of contrast work in two principal ways: on the one hand, gaps in articulatory and perceptual spaces exert a certain gravitational pull that will eventually cause these gaps to be filled by segments which were formerly located otherwise; on the other hand, crowded areas in inventories, i.e. low distinctivity among neighboring segments, will lead to a dispersal of these segment clusters in favor of a more homogeneous distribution, which in turn will highten the contrastive properties of the elements in question. Martinet believes that in a three-vowel system consisting of /i, a, o/, for example, the /o/, in its effort to become more different from /a/ without at the same time getting closer to /i/, will eventually acquire an /u/-like quality (1968: 477). With respect to consonant systems, Lindblom & Maddieson (1988) state that expanding inventories will use more and more perceptual space, which is necessarily linked to an increase in articulatory complexity.

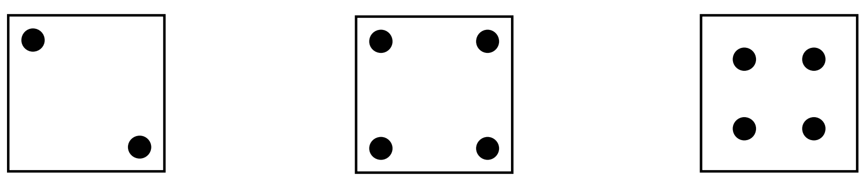
> Consonant inventories tend to evolve so as to achieve maximal perceptual distinctiveness at minimum articulatory cost. (1988: 72)

(5.8) Parallel analysis of vowels and consonants exploiting maximal perceptual contrast using the features compact vs. diffuse and grave vs. acute (after Jakobson 1949: 284)



Returning to Flemming's three goals, they are grounded in the prime function of language as a means of communication: first, phonological contrasts should be numerous enough to guarantee the formation of substantial vocabularies; second, the auditory distinctiveness of the sound segments in an inventory should be maximized in order to ensure perceptibility on the side of the hearer; and third, due to the factors governing human motor activity, effort[83] should be minimized.

(5.9) Selection of contrasts from a schematic auditory space (after Flemming 1997: 3)



a. two segments	b. four segments	c. four segments
most distinct	less distinct	least distinct
more effort	more effort	less effort

Goals one and two are inherently in conflict since contrasting sounds share both an articulatory and a perceptual space (cf. figure 5.9 above for a schematic depiction). Fitting four segments into the same space which was formerly inhabited by only two segments implies that the contrasts between those four cannot be as distinct as the ones between two segments. Following Flemming's proposal that only surface contrasts should be evaluated by an OT grammar, and not correspondences between

83 For an account of the neuromuscular basis of 'effort' and a biomechanical approach to 'effort' in terms of the mechanical notion of force see Kirchner (1998: 37-50).

input and output specifications of feature values, i.e. between underlying representations and surface representations—as Correspondence Theory does (cf. section 5.5.2 above)—the distribution of coronal plosives in unstressed syllable-initial position between voiced segments can now be shown in two single tableaux irrespective of any UR for SAE and Standard British English (SBE) respectively:

(5.10) Low-ranking MAXCONTRAST(F) results in neutralization of F contrast (SAE)

	*[–voiced]	MAXCONTRAST(F)
[–voiced] ≠ [+voiced]	*!	
☞ [+voiced]		*
[–voiced]	*!	*

When a MARKEDNESS constraint MINEFFORT(F), which embodies goal number three (minimize articulatory effort), outranks the FAITHFULNESS constraint MAXCONTRAST(F) (maximize featural contrast), which works to ensure a surface contrast in the values of the feature [voiced], neutralization is the natural consequence of this ranking.

(5.11) High-ranking MAXCONTRAST(F) requires F contrast directly in the output (SBE)

	MAXCONTRAST(F)	*[–voiced]
☞ [–voiced] ≠ [+voiced]		*
[+voiced]	*!	
[–voiced]	*!	*

Conversely, when MAXCONTRAST(F) outranks the MARKEDNESS constraint, the winning candidate is the one that maintains two distinct surface values for the feature [voiced] as is the case for SBE. Featural contrast is determined directly from surface forms without consideration of input-output correspondences which by their very nature require reference to input forms.

While Correspondence Theory accounts for contrast in features without imposing any restrictions on underlying representations (inputs), Dispersion Theory goes even one step farther by eliminating URs altogether. If it is assumed that inputs are totally irrelevant for outputs and that all relevant surface contrasts result directly from the constraint ranking, then indeed one can say that URs are erased from the grammar completely. Supposing further that the system of SAE coronals can be shown to confirm the claims of Dispersion Theory with an appropriate distribution over available spaces (articulatory, perceptual, and acoustic), even if this were the case, if applied to OT, Dispersion Theory runs into difficulties. McCarthy (2002: 225-7) has pointed out that unlike the phonological rules of standard GP, which are designed to account for alternations as well as for the gaps in the distribution of segments in a sound system (Kenstowicz 1981: 432), constraints evaluate candidate sets produced by GEN and not sound systems, i.e. phoneme inventories.

In this dissertation, I will not adopt a totally surface-oriented approach to phonological contrast with constraints that operate directly on surface representations. For the analysis of SAE alternations and the description of perceptual and articulatory demands, a version of Correspondence Theory appears to be more suitable.

5.6 FORMALIZATION OF FUNCTIONAL PRINCIPLES

In chapter 4.1, I pointed out that the formulation of an OT constraint in many ways parallels an expression of contextual markedness. For example, a segment's high or low sonority value does not by itself imply that the segment is either well-formed or ill-formed. Only in the context of other voiced segments is the more sonorous /d/ in onsets before an unstressed vowel more harmonic than its voiceless counterpart /t/ (this is the situation in SAE; cf. previous section). Similarly, the feature [+anterior] is not well-formed in itself, but the combination with [+coronal] results in a harmonic segment (in *SPE* terms (1968: 406): anterior coronals are unmarked coronals, while nonanterior coronals are marked). Such notions of markedness were formalized as *SPE*'s and Kean's (1975) marking conventions (cf. chapter 4.2.1), as well as Archangeli & Pulleyblank's (1994) grounding conditions (cf. chapter 4.3.3 for a discussion of the grounded path conditions on the features [ATR] vs. [RTR] and [high] vs. [low]).

5.6.1 Types of scales under OT

Standard OT views phonological theory as consisting of two main subtheories (Prince & Smolensky 1993: 67): a theory of substantive universals of phonological well-formedness (gradient phonetic phenomena, shown in the first two scales of figure 5.12 below) and a theory of formal universals of constraint ranking (categorical phonological phenomena, depicted in the bottom dominance hierarchy). These three types of scales or hierarchies are clearly distinguished in standard OT.

(5.12) Three different scales in Optimality Theory (after Prince & Smolensky 1993: 68)

scale or hierarchy	relates	symbol	example	meaning
phonetic scale	points along elementary representational dimensions	>	[a] > [l]	[a] is more sonorous than [l]
harmony scale	well-formedness of structural configurations built from elementary dimensions	$\succ$	[á] $\succ$ [ĺ]	a nucleus filled by [a] is more sonorous than a nucleus filled by [l]
dominance hierarchy	relative priority of well-formedness	>>	ONS >> HNUC	the constraint ONS strictly dominates the constraint HNUC

Phonetic scales show the importance of phonetic representational dimensions for OT, a good example of which is the sonority scale referred to in the top scale. With respect to feature combinations in sound segments there seems to be a phonetic basis for assuming CORONAL to be the unmarked place of articulation (Avery & Rice 1989; Paradis & Prunet 1991). Harmony scales refer to the well-formedness of units (e.g. features, segments, onsets, nuclei, codas) in specific structural contexts, that is, to the way these units function in certain domains. As mentioned in chapter 4.3.4 above, Prince & Smolensky propose to formalize coronal unmarkedness—"don't have a place of articulation other than CORONAL" (1993: 180)—in the following universal harmony scale: PL/COR $\succ$ PL/LAB, PL/DOR (1993: 180). The harmony scales are in fact directly linked to the third type of hierarchy in that they form the basis for the formulation of the constraints: dominance hierarchies regulate constraint interaction by strict domination. For example, the constraint which requires syllables to have an onset (ONS) strictly dominates the NUCLEAR HARMONY CONSTRAINT (HNUC)[84]. Thus the phonetic basis for setting up HNUC is: if [x] > [y] then nuc/x $\succ$ nuc/y (in other words: if x is more sonorous than y then a nucleus filled by x is more harmonic than a nucleus filled by y). The harmony scale capturing coronal unmarkedness readily

[84] HNUC states that a higher sonority nucleus is more harmonic than a lower sonority nucleus (Prince & Smolensky 1993: 16).

converts into the following universal dominance hierarchy of a selection of MARKEDNESS constraints concerning feature cooccurrences (1993: 181):

*PL/DOR , *PL/LAB >> *PL/COR

This constraint ranking expresses the fact that a violation of the ban on DORSAL and LABIAL, which is the case whenever DORSAL and LABIAL are associated to the PLACE node, constitutes a more serious offense than a violation of *PL/COR. The constraint against coronals is lowest in the dominance hierarchy indicating that parsing CORONAL is the least offensive violation, i.e. CORONAL is indeed the unmarked place of articulation. Such a universal ranking of MARKEDNESS constraints reflects the facts about the distribution of phonological contrasts found in the sound inventories of the world's languages (discussed in previous section), analyses of which are carried out as typological studies. However, the universal ranking says nothing about the distribution of sounds in a specific inventory of a particular language. Actual inventories are the result of the interaction of MARKEDNESS constraints with FAITHFULNESS constraints that strive to preserve input contrasts in the output.

I argue that the fact that even standard OT recognizes two scales (phonetic and harmony scales), which supply, as it were, the phonetic basis for the ranking of constraints (dominance hierarchy), provides the foundation for a unified analysis of phonetics and phonology within an OT framework. Phonetic phenomena must play a crucial role for the setting up of the constraint hierarchy.

5.6.2 Types of constraint families

The recent developments within OT allow for the statement and formalization of very general principles, e.g. articulatorily grounded constraints, to interact dynamically with other general principles, e.g. auditorily grounded constraints. As a result, there has been a revival of interest in exploring the role of perceptual phenomena, and relatively new efforts are made to integrate functional aspects of vocal tract organization into the OT model (Smolensky 1996b; Hayes 1996; Flemming 1997; Boersma 1998; Pater 1998; Hume & Johnson 2001; Flemming 2001; Flemming 2002). As shown in section 5.2 above, phonological inventories and phonological alternations are the result of the interaction of at least two conflicting constraint families: MARKEDNESS constraints (M) and FAITHFULNESS constraints (F). Generally speaking, M describes output configurations and F prohibits operations on the input; if M is ranked above F (M >> F), there is some kind of change involved unless the input conforms to the requirements of M. In cases where F is ranked higher than M (F >> M), the input form must be retained under all circumstances (no change). If functional considerations, like the ones discussed in the preceding sections, are admitted into the OT model, M can now be described as a predominantly speaker-oriented family of constraints with respect to speech production in that it strives to avoid the surfacing of marked sound structures in the language, i.e. structures that are costly in terms of articulation, while F can be viewed as a predominantly hearer-oriented family of constraints working in favor of speech perception in that it aims at the preservation of lexical contrast in input-output correspondences and thus attempts to achieve the highest possible perceptual distinctivity. In chapter 8, where different types of SAE alternations are analyzed, I will show that these two basic constraint families cannot be assigned to the two functional goals mentioned above in such a straightforward way. The interaction of the various F and M constraints results in a slightly more complex picture.

Considering an optimality-theoretic investigation of coronals in SAE, a functional perspective does appear rewarding. Coronals are well-known for their notoriously inconsistent and even contradictory properties with respect to both their distribution and their behavior in various types of alternation (e.g. Paradis & Prunet 1991). Only part of their exceptional behavior can be explained by claiming that they are underspecified for place (cf. chapter 4.3.1). Coronals are those consonants which are most frequent in every aspect of frequency measurement, they occur in environments where labials and dorsals are not allowed (cf. chapter 6); they are favored as epenthetic consonants (cf. chapter 8.3.1). On the other hand, however, there is crosslinguistic evidence as well as data from SAE which show frequent alteration of the featural specifications of coronal segments, or even their total obliteration. That is, coronals are frequently the targets of deletion (cf. chapter 8.2.2) and assimilation (Kiparsky 1985; Mohanan 1991; Paradis & Prunet 1991; cf. chapter 8.2.1) with coronals being the *assimilated* segments, and much less so noncoronals.

I wish to show that these controversial properties are best explained if coronals are viewed as being in the center of an area of conflict between the necessities of articulation and the demands of perception. This means that in the default case a more or less economical coronal gesture is preferred to a labial or dorsal gesture, but in cases when an unambiguous perception is guaranteed, a coronal input consonant may surface as a labial or dorsal output consonant, e.g. homorganic place assimilation in CC clusters. The representational framework used for this purpose is OT's machinery of inherently conflicting constraints.

CHAPTER 6

Distributional properties of coronals in SAE

6.1 INTRODUCTION

A blurring of the boundaries between phonetics and phonology automatically leads to an increase in the importance of functional considerations (Kager 1999: 421). An important issue to be addressed here is the question from which specific phonetic phenomena the constraints arise that can be shown to operate in the phonology of SAE. Unless this question can be answered satisfactorily, phonetically and functionally motivated accounts of phonological phenomena will fail to have predictive power, and thus will have the status of mere post-hoc explanations.

In this chapter, I would like to present the behavior of coronals in SAE with respect to their distributional properties. My analysis is embedded in a version of OT which pursues two main goals: first of all, it attempts to integrate the functionally grounded needs of speech production and perception (commonly regarded as gradient phenomena and external to the grammar) directly into the grammar by employing ranked constraints that refer to both articulatory and auditory requirements of communication, and secondly, being a model of grammar, OT must account for the contrastive, categorical properties of speech sounds (commonly labeled grammar-internal considerations).

My claim is that the very fact that coronals in SAE display inconsistent behavior (e.g. they may be the preferred epenthetic segments, they may also be subject to deletion under certain conditions) excludes the possibility of a one-dimensional explanation which refers *either* to the articulatory dimension (e.g. ease of articulation) *or* the perceptual dimension (e.g. enhancement of perceptual distinctiveness) alone. It is thus assumed that the forces of minimization of articulatory effort and of maximization of perceptual contrast, which are in constant conflict in a communication event, are responsible for the seemingly contradictory properties of coronals. These conflicting requirements will be described in OT terms of ranked constraints that refer to both production and categorization.

It is worth pointing out that not only coronals involved in alternations which are brought about by the dynamic interaction of sound segments in connected speech (cf. chapter 8 below) display contradictory behavior, inconsistencies are also found when looking at the more static properties, i.e. at the distribution of coronals in the consonant system of SAE as well as at phonotactic constraints, which express the syllable-based generalizations of coronals in SAE. Thus coronals appear in places where other consonant types are disallowed, e.g. as 'surplus' consonants at the margins of prosodic words. And conversely, as a whole class of sounds they are forbidden in certain combinations, e.g. in onset clusters with the palatal glide /j/, e.g. /*tj/, /*sj/, /*nj/, etc. whereas there are no restrictions for consonants belonging to the other place nodes, e.g. /pj/, /fj/, /kj/.

6.2 PARADIGMATIC CONSTRAINTS ON THE SAE CONSONANT INVENTORY

As discussed in chapter 2 above, the consonant systems of the various accents of English are not subject to significant variation from one to the next (Giegerich 1992: 33-4). A notable exception, for example, is Standard Scottish English (SSE), which has a slightly modified consonant inventory. In addition to the consonants common to all accents of English (cf. figure 6.1 below), SSE has a voiceless dorsal fricative /x/ as in <loch> and a voiceless labial fricative /ʍ/ as in <whine>, a sound also found in some Canadian and US accents. For speakers of this variety, <loch> and <lock> as well as <whine> and <wine> are minimal pairs (Giegerich 1992: 36-7).

Not only are these exceptions very few, they also do not affect coronal sounds, which are the focus of this investigation. Thus the following set of 13 coronals can be safely postulated without going into further detailed discussion of the different accents of English.

(6.1) Distribution of the English consonants over the PLACE node

LABIAL	**CORONAL**	DORSAL	GLOTTAL
p b f v m w	**θ ð t d s z ʧ ʤ ʃ ʒ n l r**	k g ŋ j	h (ʔ)

Needless to say that the situation is totally reversed with respect to the various vowel systems, which differ at times radically from one another and which form the basis according to which the numerous accents are distinguished (Giegerich 1992: 43; cf. also chapter 2.2.1 above).

6.2.1 The emergence of sound inventories under OT

From an OT point of view, particular inventories are the result of specific constraint rankings for the languages in question, i.e. inventories derive solely from constraints on outputs and not on input forms, since Freedom of Analysis grants GEN a virtually unconstrained production of a multitude of conceivable output candidates for some input selected from the LEXICON (McCarthy & Prince 1993: 88)—that is, virtually unconstrained, since the only restriction GEN must face is that it produces linguistic objects (Prince & Smolensky 1993: 4). Taking the notorious example from chapter 4 again, namely the fact that English sonorants are redundantly voiced, a high-ranking MARKEDNESS constraint SON/VOI ('a sonorant must be voiced') guarantees that voiceless sonorants will not surface in the language (Itō, Mester & Padgett 1995: 581; and also discussion in chapter 4.3.4 above).

In pre-OT analyses, cooccurrence restrictions on features, such as the automatic voicing of sonorants, were expressed in various ways: for instance, as an *SPE*-type structure-building rule [+sonorant] → [+voiced] (Chomsky & Halle 1968: 380-9), or as a context-free markedness filter prohibiting sonorants from being specified for voicing in the lexicon *[+sonorant, αvoiced] plus a default rule supplying the predictable feature value [+voiced] (Kiparsky 1985: 92), or as an implicational statement representing a phonetically grounded, very strong path condition [+sonorant] > [+voiced] (Archangeli & Pulleyblank 1994: 179; cf. also condition b. shown in figure 4.17 of chapter 4.3.3). It is self-evident that Archangeli & Pulleyblank's Grounding Hypothesis is largely a formalization of markedness considerations as well.

Looking at the medium-sized consonant inventory of SAE (24 consonants), obstruents display a voicing contrast in that all voiceless obstruents except /h/ have voiced counterparts (cf. figure 6.1 above). On the other hand, taking again Hawai'ian as an example of an extremely small inventory with 8 consonant segments /m, p, w, n, l, k, h, ʔ/, this consonant inventory contains voiceless obstruents only, i.e. voicelessness is redundant (Pukui & Elbert 1957: xvii). Just as sonorants are redundantly voiced and thus frequently grouped under the SV-node (SPONTANEOUS VOICE) in feature-geometric representations (e.g. Rice & Avery 1991: 103; cf. also figure 4.13 in chapter 4.3.2 above), the default value for obstruents is [–voiced] (e.g. Stampe 1979), which can be expressed as a phonetically grounded[85], medium strong path condition (cf. condition c. shown in figure 4.17 of chapter 4.3.3 above). This condition is labeled medium strong because a violation of it is frequent, especially in larger inventories. The distinction between the two inventories with respect to obstruent voicing can be analyzed as two different language-specific rankings of CON (the set of universal constraints). On the one hand, there is the MARKEDNESS constraint OBS/VOI ('an obstruent must be voiceless'; Hayes 1999: 14), on the other there is the FAITHFULNESS constraint FAITHLAR, requiring outputs to surface faithfully concerning the laryngeal feature [voice].

For a language like English with voiced obstruents in its inventory, the FAITHFULNESS constraint must crucially dominate the MARKEDNESS constraint that demands that obstruents be voiceless. Languages like Hawai'ian have exactly the opposite ranking; the MARKEDNESS constraint crucially outranks the FAITHFULNESS constraint thus guaranteeing that only voiceless obstruents will be part of the inventory. It is indeed rather likely that the candidate set produced by GEN for a specific Hawai'ian input form may contain voiced obstruents; for example, the Polynesian (cf. also chapter 2.1 above) word <tabu> meaning {taboo} is pronounced /kapu/[86] in Hawai'ian, since voiced obstruents are ruled out by the high-ranking MARKEDNESS constraint OBS/VOI. Whether or not an input form taken from the Hawai'ian lexicon actually does contain voiced obstruents, however, is in principle an irrelevant issue, considering that OT's approach is completely surface-based.

[85] The phonetic explanation is that articulatory effort is reduced if there is no vocal cord vibration during the constriction phase for plosives and fricatives (Westbury & Keating 1986: 146).

[86] Note that Hawai'ian is in fact the one curious language (of 317 investigated languages; cf. chapter 4.3.1 above) that does not have a coronal plosive (Maddieson 1984: 32); this is the reason why /t/ surfaces as /k/.

(6.2) Fraction of an OT grammar for English with voiced obstruents in its inventory: FAITHLAR >> OBS/VOI >> *PLCOR

/tabu/	FAITHLAR	OBS/VOI	*PLCOR
☞ tabu		*	*
tapu	*!		*
kapu	*!		

(6.3) Fraction of an OT grammar for Hawai'ian with an exclusively voiceless obstruent inventory: OBS/VOI >> *PLCOR >> FAITHLAR

/tabu/	OBS/VOI	*PLCOR	FAITHLAR
tabu	*!	*	
tapu		*!	*
☞ kapu			*

The language-specific ranking of constraints, which ensures that the optimal candidate, i.e. the voiceless obstruent will surface, is equivalent to the grammar of the respective language. Restrictions on underlying inventories, which then form phonological inputs as underlying representations, find no place in an OT grammar.

6.2.2 Segmental contrast in SAE

Lindblom states that "an optimal system might be one that meets not only perceptual but also memory-based and sensori-motor conditions of distinctiveness" (1986: 39). He draws attention to the following articulatory and perceptual evidence: (1) articulators have greater mobility at the front of the mouth (e.g. the tongue tip); (2) there appears to be a richer supply of structures for sensory control at anterior vocal tract locations; and (3) acoustic-perceptual effects are greater at the front than at the back, given geometrically comparable articulatory perturbations and conditions typical of, for example, voiceless consonants (ibid.). These observations are supported by an experiment designed to provide perceptual data concerning the relative distinctiveness of different consonant classes, both with respect to place and manner of articulation. Hura et al. investigated the error rates in the identification of nasals, plosives, and fricatives, and of three place categories (excluding GLOTTAL), which are of importance here. Of the three places of articulation, the DORSAL place showed the highest error rate in identification followed by the LABIAL place with the

CORONAL place exhibiting the lowest rate of subjects' confusions (1992: 64-7). On the other hand, the majority (about 75%) of the subjects' incorrect nasal responses were /n/, of all incorrect fricative responses were /ʃ/, and of all incorrect plosive responses were /t/. Thus, when perception is somehow not unambiguously possible or otherwise hampered, hearers tend to anticipate that the consonant in question is a coronal; this led Hura et al. to postulate that coronals are the "default" perceptual segments (1992: 66).

Additional suggestive evidence in support of Lindblom's assumptions comes from an empirical study by Stevens & Keyser (1989) regarding the acoustic manifestations of features. Among the entire set of distinctive features for consonants, they establish the three features [continuant], [sonorant], and [coronal] as "primary features" (1989: 81); these are claimed to be more salient and robust than all remaining or "secondary features" (1989: 82). Of the three robust features, only one is a place feature: [coronal]. The reason for selecting [coronal] as primary is that it has distinctive acoustic properties which set it apart from all other place features. Place cues are characterized by specific patterns in the relative distribution of energy in the high and low frequencies: [+coronal] segments are unique in that they share a large spectral amplitude at high frequencies in relation to the amplitude at low frequencies (1989: 97-8; cf. also chapter 3.3.2 above).

Another major point of Stevens & Keyser's study is that a primary feature like [coronal] can be represented in a given sound segment with varying degrees of strength and that it can be enhanced or weakened by its cooccurrence with specific secondary features. For a segment specified [–continuant], [–sonorant], and [+coronal]—a /t/-type sound—the secondary feature [–distributed] is the ideal choice for an enhancement of the three primary features. That is, in order to give rise to an abrupt onset of acoustic energy and exploit this perceptually distinct property maximally, the thus specified segment should be produced apically and not laminally, i.e. with a short constriction at the tongue tip.

The set of segment types which results from all eight possible configurations of the three primary features plus their enhancement through secondary features corresponds exactly to Maddieson's (1984) list of the ten most frequently-occurring consonants utilized to express contrasts in the 317 sound inventories surveyed, i.e. /m, p, w, n, s, t, l, j, k, h/ (Stevens & Keyser 1989: 86).

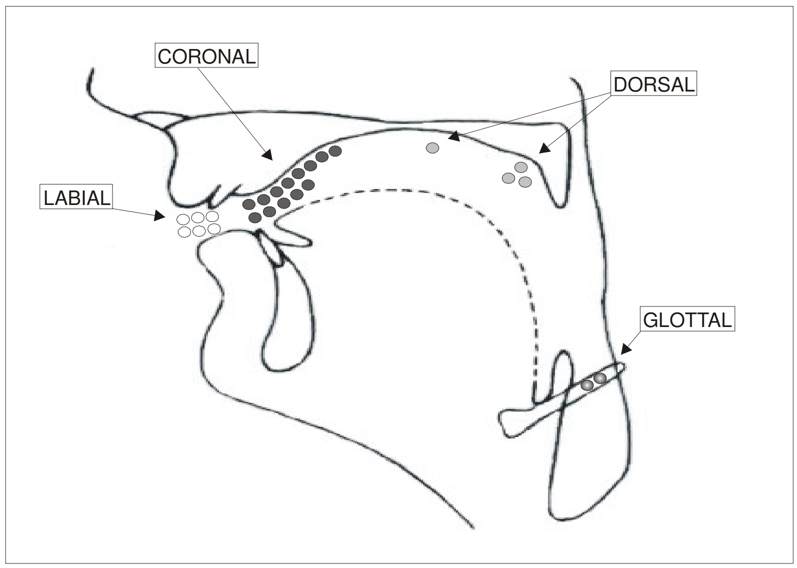
(6.4) The three primary features and their combinations (after Stevens & Keyser 1989: 86)

segment type	CORONAL	SONORANT	CONTINUANT
/t/	+	–	–
/s/	+	–	+
/n/	+	+	–
/l/	+	+	+
/p, k/	–	–	–
/h/	–	–	+
/m/	–	+	–
/w, j/	–		

Of the ten segment types, four are characterized by a coronal place feature (cf. figure 6.4 above). These segment types are labeled "optimal, since they provide the strongest representation of the contrast defined by each of the three primary features" (Stevens & Keyser 1989: 89). From the hearer's perspective they stand out due to their maximal distinctness or optimal dispersion in the available perceptual space (Lindblom 1992: 140-1). Not surprisingly, languages with a small consonant inventory select their segment types almost exclusively from this set (cf. again Hawai'ian with its eight consonant segments /m, p, w, n, l, k, h, ʔ/) while those with larger inventories additionally will have to exploit contrasts which are not maximally distinct. In this situation (e.g. SAE), clearly the coronals are the favored segments to enlarge an inventory. Taking the ten most frequently-occurring segment types, the claims of both Lindblom (1986) and Stevens & Keyser (1989) are confirmed. Anterior vocal tract locations are preferred over posterior locations: LABIAL and CORONAL segments amount to seven whereas at DORSAL and GLOTTAL places only three segments are produced. In a direct comparison between LABIAL and CORONAL, CORONAL wins with four to three segments.

Returning to the dispersion of consonants in the system of SAE, according to Lindblom, there should be a clustering of distinctive sounds at the front region of the oral cavity (e.g. around the coronal articulators) as opposed to fewer distinctive sounds at the posterior parts of the mouth (cf. figure 6.5 below). Evidently, there is an asymmetry of vocal tract sensori-motor representations manifest in the consonant system of SAE. As predicted by Lindblom, a great variety of sounds are produced at anterior vocal tract locations using the highly mobile front part of the tongue, i.e. the coronal articulators. Due to the high perceptibility of consonants produced in this region and with these articulators, the available articulatory and perceptual spaces need not be as large as at posterior locations where fewer contrasting consonantal targets are located. Phonological evidence for this claim comes from the phonotactic restrictions on SAE syllable structure (cf. section 6.3 below).

(6.5) The set of English consonants shown on a vocal tract profile illustrating a crucial asymmetry with respect to the exploitation of articulatory space



Presupposing that the system of SAE coronals can be shown to confirm the claims of Dispersion Theory, i.e. that the dispersion of coronals in the system of SAE (Lindblom 1986: 39-40; cf. also the discussion in chapter 5.5.3 above) displays an appropriate distribution over available spaces (perceptual and articulatory), an area of contention remains when Dispersion Theory is applied to OT. McCarthy points out that constraints, unlike the phonological rules of standard GP, which are designed to account for alternations as well as for the gaps in the distribution of segments in sound systems (Kenstowicz 1981: 432), evaluate candidate forms produced by GEN and not sound systems, i.e. inventories (2002: 225-7). However, as shown in section 6.2.1 above, constraint ranking may be held responsible for the type of sound that can surface in a given language.

6.3 SYNTAGMATIC CONSTRAINTS ON SAE SYLLABLE STRUCTURE

The so-called *core syllable* is associated with a single peak in sonority, which is more sonorous than its neighbors to the left and right. The syllabic element (the peak) is *the* obligatory constituent of the syllable. At maximum, two segments (X-positions) forming the onset may precede the peak, and sonority increases from left to right towards the peak. In the segments following the peak and forming the coda (maximally two X-positions), sonority decreases from left to right towards the right margin of the syllable. Peak and coda form the rhyme constituent which may contain three X-positions at the most (Giegerich 1992: 146-7; Kenstowicz 1994: 255).

There is strong evidence for assuming that peak and coda combine to form the higher-level rhyme constituent, since there are very specific collocational restrictions to be observed between peak and coda which do not hold between the onset and peak constituents. The number of consonants constituting the coda depends on the vowel type forming the peak constituent: if the vowel is lax and thereby occupies one X-position, the coda may contain two consonants. If however the vowel forming the peak is tense, thus occupying two X-positions, the number of consonants in the coda is generally restricted to just one (Kenstowicz 1994: 259).

Under certain conditions, there is a violation of either quantity constraints regarding the number of elements well-formed English syllables may contain, or of the qualitative Sonority Sequencing Principle (SSP; e.g. Selkirk 1984b: 116; Clements 1990) requiring the sonority profile of well-formed syllables to be of its specific shape: it must rise until it peaks, and then fall:

> Between any member of a syllable and the syllable peak, only sounds of higher sonority rank are permitted. (Clements 1990: 285)

In an early OT account, Hammond formalizes the SSP as the SONORITY constraint ("onsets must increase and codas must decrease in sonority"; 1997: 40), which "appears to be unviolated in English" (ibid.). The ensuing discussion of SAE onset and coda clusters will show that this rather simplistic postulate is unfortunately false.

The phonetic correlates of sonority are not altogether clear (Ohala & Kawasaki 1984: 122; Clements 1990: 290-2) because no consensus has been reached so far as to what properties should be used for its definition: perceptual, articulatory, or acoustic parameters. Nevertheless, a phonetic basis for sonority is commonly invoked by phonologists in terms of the consonants' resonance characteristics. The following scale for the relative sonority of segments seems to be widely accepted in the phonological literature (e.g. Foley 1972: 97; Hooper[87] 1972: 537; Ohala & Kawasaki 1984: 122; Kenstowicz 1994: 254):

(6.6) Sonority scale (after Ohala & Kawasaki 1984: 122)

least sonorant								most sonorant
obstruents	<	nasals	<	liquids	<	glides	<	vowels

Crosslinguistic evidence for such a sonority scale comes from Greenberg's (1978) set of 40 phonological universals, which are based on a sample of 104 languages. He formulates generalizations concerning initial and final consonant cluster types as implicational statements, e.g. the presence of initial stop + stop clusters implies initial stop + fricative combinations (universal no. 7; 1978: 254), at the same time these are expressions of relative markedness. According to Prince & Smolensky (1993: 68) implicational statements such as these represent universal harmony scales (cf. also discussion in chapter 5.6 and figure 5.11 above) which express the crosslinguistic well-formedness of units (e.g. features, segments, onsets, nuclei, codas, syllables) in specific structural contexts. The phonetically based harmony scales are then converted into constraint families governing the syllabic positions of segments with different degrees of sonority. These constraints, which account for the distribution of various segments across syllable positions in OT, are *inherently* ranked constraints, i.e. the substantive properties of the constraints themselves determine their ranking with respect to other constraints. In this way, again, it can be safely said that even standard OT attempts to ground analyses in phonetics, at least to a certain extent.

Prince & Smolensky posit the two prominence scales in figure 6.7 below as phonetic scales, which are supposedly universally aligned (1993: 137); their harmonic alignment is shown in figure 6.8 below, the constraint alignment is depicted in figure 6.9 below.

[87] Her proposal is notable also for reincorporating a definition of the syllable into GP.

(6.7) Prominence scales (after Prince & Smolensky 1993: 136)

a. Syllable Position Prominence:
P > M (i.e. a peak is more prominent than a margin)

b. Segmental Sonority Prominence (slightly modified):
Low V > High V > Glide > Liquid > Nasal > Obstruent

The original nuclear harmony scale plus its associated constraint (HNUC), which states that a higher sonority nucleus is more harmonic than a lower sonority nucleus (Prince & Smolensky 1993: 16; cf. also chapter 5.6 above), has now been rewritten as a harmony scale referring to syllable peaks which is paralleled by a harmony scale referring to syllable margins:

(6.8) Peak and margin harmony scales (after Prince & Smolensky 1993: 134-5)

a. Peak Harmony (slightly modified):
P/V ≻ P/Glide ≻ P/Liquid ≻ P/Nasal ≻ P/Obstruent

b. Margin Harmony (slightly modified):
M/Obstruent ≻ M/Nasal ≻ M/Liquid ≻ M/Glide ≻ M/V

The harmony scales above correspond to the following two dominance hierarchies:

(6.9) Peak and margin constraint hierarchies (after Prince & Smolensky 1993: 134-5)

a. Peak Hierarchy (slightly modified):
*P/OBSTRUENT >> *P/NASAL >> *P/LIQUID >> *P/GLIDE >> *P/V

b. Margin Hierarchy (slightly modified):
*M/V >> *M/GLIDE >> *M/LIQUID >> *M/NASAL >> *M/OBSTRUENT

Two types of prosodic prominence are aligned, sonority of segments and syllabicity. The basic idea encoded in these constraint families is that syllable position prominence and segmental sonority prominence are correlated in that prominence of type a. should cooccur with prominence of type b., and conversely, the lack of prominence of type a. should cooccur with a lack of prominence type b (cf. figure 6.7 above). That *P/OBSTRUENT is ranked at the top of the constraint hierarchy expresses the fact that obstruents are the least desirable and also the least likely segments in peak position whereas vowels are the ideal segments to fill a syllable peak. On the other hand, *M/V says that vowels make the worst possible syllable margins whereas obstruents typically appear at this syllable position.

In the following two sections I will show that SAE syllables sometimes violate both principles, the quantity restrictions *and* the SSP. The existence of clusters which do not conform to the patterns prescribed by the principles undermines their universality within theories in which constraints are inviolable.

6.3.1 Syllable onsets

In an attempt to devise a universal syllable theory, standard OT postulates a licensing asymmetry between onsets and codas, i.e. "codas can contain only a subset […] of the segments appearing in onsets" (Prince & Smolensky 1993: 128). This supposedly universal tendency is derived from the crosslinguistically well-supported facts that onsets are generally required in syllables while codas are frequently absent. These facts are formalized as two Basic Syllable Structure Constraints: ONSET and *CODA (ibid.). It is obvious then that with respect to syllable margins (cf. figure 6.7 above), a division into onsets and codas is indeed sensible. Unfortunately however, Prince & Smolensky's generalization does not hold for SAE syllable onsets and codas. It seems that the asymmetry works precisely vice versa: there are much more varied consonant clusters possible in codas, e.g. /pt, mp, nd, ŋk, lt, lm, rn/, which are impossible as onsets, e.g. /*tp, *pm, *dn, *kŋ, *tl, *ml, *nr/ when their order is reversed (due to the SSP; for syllable codas cf. section 6.3.2 below).

While almost any consonant segment can function as a single-consonant onset in SAE (the sole exception is the velar nasal /ŋ/), severe restrictions are placed on two-consonant onsets.

(6.10) Ranking of the complex *ONSET constraints for SAE (after Hammond 1999: 88; modified)

*O/NASOBS, *O/GLIDEOBS >> *O/NASLIQUID, *O/NASGLIDE >> *O/OBSNAS >> FAITH >> *O/OBSLIQUID, *O/OBSGLIDE

SAE onsets are generally of two distinct types: either they are composed of a plosive or a voiceless fricative plus a sonorant (except for the class of nasals), i.e. of an obstruent plus liquid or glide (cf. figure 6.10 above), or of the voiceless coronal fricative /s/ followed by an oral or nasal stop, i.e. a voiceless plosive or the nasals /m/ and /n/ excluding /ŋ/ (cf. figures 6.14 and 6.15 below). The constraints refer to the classes of sounds which are used to analyze the different degrees of sonority of the SSP (cf. figure 6.6 above). The two constraints ranked highest in the hierarchy are not ranked with respect to each other because they express a sonority reversal which is prohibited for SAE onset clusters[88] in all circumstances. Nasal plus liquid and nasal plus glide onsets are the next to be ruled out, just as obstruent plus nasal clusters[89] are generally disallowed. These constraints all outrank FAITHFULNESS, i.e. a possible input /pnɪk/ is not allowed to surface faithfully.

(6.11) Tableau for the impossible SAE lexeme *<pnick> containing the illicit onset */pn/

/pnɪk/	*O/NASOBS	*O/OBSNAS	FAITH	*O/OBSLIQUID
pnɪk		*!		
npɪk	*!		*	
pmɪk		*!	*	
☞ prɪk			*	*

FAITHFULNESS must be ranked above *O/OBSLIQUID because if the input were indeed /prɪk/, FAITHFULNESS would guarantee its surfacing.

The only segment that can be part of an onset exceeding two X-positions *and* violate the SSP is coronal /s/. In two-consonant onsets as for example /spaj/, /stej/, /skaj/ only the SSP is violated, whereas cases like /splijn/, /strowk/, /skrijm/, in addition to the SSP violation, contain otherwise illicit three-consonant onsets.

Theoretical proposals range from an analysis of /s/ as an instance of contingent extrasyllabicity[90] (Goldsmith 1990: 107-8) to one where the cluster /s/ plus voiceless plosive qualifies as a single obstruent segment (Selkirk 1982: 336), to an analysis of /s/ as an appendix to the syllable node (Giegerich 1992: 147-50) or as an appendix to the prosodic word (PW)[91] node (Golston & van der Hulst[92] 1999: 165-6), thereby

[88] For the special case /sC/ cf figure 6.14 below.

[89] /sm/ and /sn/ are dealt with in connection with the special case /sC/, cf. figure 6.15 below.

[90] Extrasyllabicity is a device that allows segments which cannot be parsed according to the existing syllables templates of a language to be invisible to the rules of syllabification. This device is constrained insofar as only segments at the peripheries of a given domain (syllable, prosodic word) may qualify as extrasyllabic segments (Kenstowicz 1994: 274).

[91] The term "phonological word" also frequently appears in the literature (e.g. in Golston & van der Hulst 1999: 157), which is then used as a synonym for the term "prosodic word".

[92] Actually, Golston & van der Hulst use the more accurate term "prependix" instead of "appendix" for the word-initial extrasyllabic position that /s/ occupies (1999: 165).

skipping the syllable and foot nodes and violating the Strict Layer Hypothesis (SLH)[93]. The latter is an account of licensed extrasyllabicity referring to the general notion of *prosodic licensing* and requiring each segment to be licensed by being part of some higher-level domain (Goldsmith 1990: 107-8). The edges of a PW are endowed with specific licensing powers regardless of the inevitable SLH violation (Kenstowicz 1994: 260). No matter how these phenomena are accounted for, the coronal /s/ is special in that it can appear in places where other sounds cannot.

Conversely however, coronals do not form onsets with the coronal liquid /l/ whereas the other plosives and the voiceless labial fricative are not restricted in this position /*tl, *dl, *θl, *ðl but pl, bl, kl, gl, fl). These facts are not captured by the very general constraint hierarchy shown in figure 6.10 above; the ranking of *O/OBSLIQUID below FAITHFULNESS would in fact allow all of the above clusters at the surface. Superficially, a solution might be the division of *O/OBSLIQUID into a specific constraint *O/COROBSLIQUID representing the special case and a more universal one applying to the general cases only—roughly analogous to Kiparsky's Elsewhere Condition (1973: 94)—with the following ranking:

*O/CorObsLiquid >> Faithfulness >> *O/ObsLiquid

This ranking, however, does not achieve the desired result either, since licit onset clusters like /tr, θr, sl/, for example, would be disallowed. It is quite unfortunate that the only solution available seems to be the postulation of constraints that refer to the segment clusters in question directly, e.g. *ONSET/[tl], *ONSET/[θl]. Constraints like these are of course no more general or theoretically desirable than the marking of certain onset clusters with an asterisk.

The sole exception to the ban on coronal plus /l/ onsets is again /s/ as in <sleep> which—in analogy to the /s/ plus plosive onsets discussed above—once more can be considered as a candidate for extrasyllabicity, this time the reason being the exceptional behavior of coronal /s/ within its own class of sounds. The prohibition on coronals plus coronal /l/ in onsets may be attributed to the effects of the Obligatory Contour Principle (OCP)[94] which disallows adjacent identical places

[93] The SLH holds that elements of a given prosodic domain (such as the prosodic word) must be exhaustively analyzed into a sequence of elements of the next lower domain (the foot) without any leftovers (Selkirk 1984a: 26; Nespor & Vogel 1986: 7).

[94] The first expression of the OCP is commonly attributed to Leben (1973), one of the earliest autosegmental analyses of tone, and is later explicitly formulated in Goldsmith (1976). It holds that adjacent identical elements are banned from lexical representations. Extended to segmental

of articulation. Further support for this comes from the class of labials /p, b, f, v, m/, which display behavior of a similar kind in that they fail to cluster with the labial glide /w/ in onsets. Since the sequence obstruent plus nasal is generally disallowed in SAE onsets[95], no further evidence can be adduced from the observation that homorganic onsets like /*tn, *dn/ are illicit. However, the third coronal sonorant /r/ *does* cluster freely with coronal obstruents except for /*ðr, *ʒr/, and /*sr, *zr/, thereby ruining the overall generality of the OCP argument. Again, /s/ has a special status in that it can precede all licit consonants (i.e. voiceless plosives or sonorants; cf. just above) *except* for /r/. The only consonant that coronal /ʃ/ can combine with to form a two-consonant coda is /r/ as in <shrink>. Thus, /s/ and /ʃ/ are in complementary distribution. One proposal for treating this complementarity within an OT framework is the following:

(6.12) Tableau for <shrink> (after Hammond 1999: 103; modified)

/srɪŋk/	FAITHCOR	FAITHMOA	*O/[sr]	FAITHPOA
srɪŋk			*!	
frɪŋk	*!			*
trɪŋk		*!		
☞ ʃrɪŋk				*

A high-ranking, specific FAITHFULNESS constraint ensures that the coronal place of articulation is preserved by all means. Next is a general FAITHFULNESS constraint demanding that segments surface faithfully with respect to their manner of articulation, i.e. only fricatives are allowed. Ranked below FAITHMOA is then a very specific MARKEDNESS constraint militating against the onset cluster /sr/. An input form /srɪŋk/ in analogy to /tr, dr/ is indeed conceivable, but it must not surface. This is achieved by the proposed ranking of the relevant constraints. However, since constraints—just like the former rules of GP—are designed to capture generalizations about language, it appears as if a constraint banning the onset cluster /sr/ is no more desirable than a rule deriving only /ʃr/ onsets. The reason seems to be

phonology, this means that two identical features or place nodes (e.g. LABIAL) may not be adjacent on a given tier (cf. also chapter 4.3.2 above).

[95] The sole exception to this generalization is again the coronal fricative /s/ as in <smile>, <snake>, which for this reason, could again be regarded as an extrasyllabic element.

that more or less idiosyncratic phonotactic restrictions concerning syllable structure cannot be stated easily in any grammar model, and therefore not in OT either.

A totally different theoretical proposal worth mentioning is the one made by Yip (1991). The so-called Cluster Condition introduced in her paper is designed to account for English onsets without the need for postulating extrasyllabicity for any of the onset members (i.e. /s/). In addition, her analysis is interesting because she comes to the conclusion that within the class of coronals only /t, d, s, z, n/, which are specified [+anterior] and [–distributed], are truly placeless, or unmarked (cf. chapter 7.2 below). The other coronal consonants are claimed to be more restricted in their freedom of occurrence (1991: 74). Yip's data are monomorphemic English words containing initial, medial, and final consonant clusters which never include more than one noncoronal, such as <chapter, factor, abdomen, clasp, prism, lift, anvil, tense>.

(6.13) The Cluster Condition for English (after Yip 1991: 63)

Adjacent consonants are limited to at most one place specification.

Yip assumes a model of underspecification in which coronals lack a CORONAL articulator node in URs (cf. discussion of coronal underspecification in chapter 4.3.4 above) and are thus underspecified for place. On their way to the surface, the missing values are inserted by default rules. Clusters which do not contain coronals must then be homorganic. The Cluster Condition works very well for the three-consonant onsets /sCl/ and /sCr/, where only the single consonant C can be a noncoronal and consequently has a place specification. It also works well for most two-consonant onsets.[96] Apart from the cases that observe the Cluster Condition, Yip's proposal does not explain why there is also a tendency to avoid homorganic coronal two-consonant onsets, e.g. /*tn, *dn, *sr, *ðr, *tl, *θl/, as expressed by the OCP. Yip's claim that the coronals which carry the specification [+distributed], i.e. /θ, ð, ʧ, ʤ, ʃ, ʒ/, behave like the noncoronals (1991: 74) does not contribute to the clarification of this issue. It might be hypothesized that English consonant clusters need *one and only one* place

[96] However, Yip does not say anything about onsets like /pj/ in <pure> or /kw/ in <quick>. Since she does not explicitly classify /j/ as coronal or /w/ as dorsal, one must assume two place specifications for these onsets, i.e. LABIAL plus DORSAL and vice versa.

specification in URs. Since the illicit clusters above consist of coronals only, these clusters would have no place specification in URs. Unfortunately, this condition on clusters is not in accordance with the facts either, as all-coronal clusters such as /sn/ in <snake, tense> or /nt/ as in <hint> are perfectly grammatical.

It can be concluded that /s/ plus plosive onsets are not governed by the SSP in the same way that other onset clusters are. In OT terms, this means that SONORITY must be outranked by various other constraints, among these a constraint stating that if an obstruent and a stop (nasal or oral) are aligned next to each other in an SAE onset, the first segment must be /s/. Since it should become apparent that obstruent + stop onsets are generally disallowed, e.g. /*pt, *pn, *tm, *kp, *fp, *θd, *zg/, a high-ranking general constraint *O/OBSSTOP must guarantee that these clusters do not surface. This constraint is so defined as to refer to all obstruent + stop clusters except for /s/ + stop because this special case is formulated as the very specific constraint *ONSET/[s]STOP which is ranked below FAITHFULNESS.

(6.14) Tableau for <stake> ruling out all obstruent + stop onsets except for /s/ + stop

/stejk/	*O/OBSSTOP	FAITH	*O/[s]STOP	SONORITY
☞ stejk			*	*
sneik		*!	*	
tnejk	*!	*		
pteik	*!	*		*

Similarly, if the input is /snejk/, it is allowed to surface due the high-ranking FAITHFULNESS constraint.

(6.15) Tableau for <snake> ruling out all obstruent + stop onsets except for /s/ + stop

/snejk/	*O/OBSSTOP	FAITH	*O/[s]STOP	SONORITY
☞ snejk			*	
steik		*!	*	*
tnejk	*!	*		
pteik	*!	*		*

Although the postulation of a general and a specific constraint referring to similar onset types seems to be a clear disadvantage, this hierarchy would ensure an input onset /sl/ in the output owing to high-ranking FAITHFULNESS.

To add further material to the controversy, however, all 13 coronals are prohibited from occurring in a cluster with the dorsal glide /j/ in SAE onsets /*tj, *dj, *sj, *nj, etc./ whereas all other consonant types may precede the glide /pj, bj, kj, gj, etc./.[97] In chapter 4.3.1 above, I showed that these observations function as evidence for the claim that coronals form a natural class of sounds. In an OT framework, this fact can be captured in terms of a high-ranking constraint *ONSET/COR[j].

Moreover, the (coronal) affricates fail to cluster with any consonant in onsets. This exceptional behavior is best accounted for by analyzing affricates as single consonants on the CV-tier but as two consonants on the segment tier, thereby satisfying the requirements on the core syllable that onsets may contain no more than two consonants, with the unique exception of coronal /s/.

Considering the amount of contradictory evidence in the distribution of coronals, it is not entirely clear whether or not the tendency to avoid homorganic two-consonant onsets should be attributed to the effects of a constraint maximizing auditory distinctness, thus banning adjacent identical places of articulation from the surface in onsets. If so, then /θ, t, d, ʃ/ plus /r/ onsets must be attained from a high-ranking constraint minimizing articulatory effort. If not, all onsets containing a coronal consonant plus the lateral (except for the sequence /sl/) must be excluded from the surface by the constraint *ONSET/COR[l].

6.3.2 Syllable codas

A similar situation to that with onsets occurs with codas, which also exceed the upper bound of elements under certain conditions. No matter whether the violating elements are simply analyzed as extrasyllabic or as appendices to a specific domain, such as the prosodic word, crucial for my investigation is that it is *exclusively* coronals that appear in such exposed positions. The situation for codas is somewhat more complicated than for onsets because the quantity restriction with respect to the maximum number of segments a coda may have (usually no more than two X-positions) is dependent on the number of segments in the peak constituent. If the peak has two X-positions, the coda generally may consist of only a single consonant.

[97] If the glide /j/ were indeed placed under CORONAL and not under DORSAL, as Hall (1997: 21) and Hammond (1999: 5) assume (also cf. discussion in chapter 3.2.2 above), the prohibition against onset clusters composed of a coronal consonant plus /j/ could undoubtedly be attributed to the effects of the OCP.

The coronal obstruents /θ, t, d, s, z, ʤ/ may violate both the SSP and the quantity constraints defining the well-formedness of syllables in SAE. In the examples <adze, fox, width>, the coronal fricatives combine with a preceding plosive and thus violate the SSP (cf. figure 6.6 above), or they violate the quantity requirements as in <paint, bounce, strange>, or they violate both as in <text, hoax>.

(6.16) SAE mono-syllables with rhymes violating either quantity constraints, or the Sonority Sequencing Generalization, or both

XXX exceeded	SSP violated	both	violating segments
mind			/d/
paint		text	/s/, /t/
beans	adze	globes	/z/
bounce	fox	hoax	/s/
warmth	width		/θ/
strange			/ʤ/

Cases like <globes, warmth, sixths> have a special status, because they have either a derivational suffix <-th>, or an inflectional suffix <-s>, or even both suffixes added. The regular inflection processes which do not change the number of syllables[98] add the unmarked coronals /s, z, t, d/ to verb and noun stems: 3rd person singular present tense <she /klajmz/, he /frijks/>, past tense and past participle <she /klajmd/, he /frijkt/>, plural formation </bijdz/, /bijts/>, and genitive inflection <a /frɛndz/ place, the /lɑfts/ location>. It is quite striking that the suffixes (inflectional *and* derivational) which violate the SSP and the quantity constraint all belong to the class of coronals, that is, as coronal obstruents the violating segments are characterized by the features [+coronal], [–sonorant]. This can possibly be explained by postulating that if the core syllable must be expanded, e.g. through the processes of derivation or inflection, then the least marked segments, i.e. coronals, will be first to emerge and fulfill the task.

This hypothesis is supported by a closer look at the collocational restrictions between peak and coda. If the peak is filled by one of the three true diphthongs[99],

[98] Cf. e.g. adjective inflection <big> vs. <bigger, biggest> or <love> vs. the progressive <loving>, where the inflected forms contain two syllables.

[99] True diphthongs are characterized by a perceptually distinct change in vowel quality and are set off from vowel sounds, such as /e/ and /o/, which may have the diphthongal realizations [ej] and [ow] as in <fate> and <tow> but may also be realized as plain monophthongs as in <hair> and <toes>

there is room left for just one consonant in the coda, otherwise the syllable cannot qualify as SAE core syllable. With the exception of the diphthong /aj/, which may be followed by any consonant type, the single consonant type that may appear after /oj/ and /aw/ in the coda is restricted to being a coronal consonant (cf. figure 6.17 a. below). In cases where the number of X-positions in the rhyme exceeds three, it comes as no surprise that only coronals may appear in this position (cf. figure 6.17 b. below).

(6.17) SAE mono-syllables with rhymes containing one of the three true diphthongs

rhyme	/aj/	/oj/	/aw/
a.	/pajl/	/vojs/	/dawn/
XXX	/tajp/	*/fojm/	*/tawb/
	/lajk/	*/mojŋ/	*/prawg/
b.	/majnd/	/pojnt/	/sawnd/
XXX	*/najmf/	*/sojmt/	*/rawmp/
exceeded	*/pajŋk/	*/bojlg/	*/shrawŋk/

Parallel to the constraints on obstruent plus nasal sequences in onsets, the sequence nasal plus obstruent is subject to restrictions in syllable codas as well. However, these are not as severe as to amount to a total prohibition of such consonant combinations. Instead, a tautosyllabic sequence of nasal (stop) plus noncoronal consonant must have the same place of articulation. This means that homorganic clusters of nasal plus obstruent (e.g. /mp/, /mf/, /nt/, /ns/, /nʧ/, /ŋk/) are licit codas[100], that the heterorganic combinations are allowed if C_2 is coronal (e.g. /mθ/, /mt/, /ŋθ/), but that sequences such as */mk/, */np/, */nf/, */nk/ are disallowed on the grounds that C_2 is not coronal (regardless of the composition of the peak). It is significant that precisely the unmarked coronal obstruents /d, z/, which are allowed to occur unsuffixed in homorganic NC clusters <sand>, <lens>, are the ones that are suffixed to a stem ending in a voiced consonant, e.g. <canned> as well as <trimmed> and <banged>, <fins> as well as <rims> and <rings> in regular verb and noun inflection. The restrictions on NC sequences in codas can be summarized: in a coda consisting

in SAE. Moreover, the three sounds /aj, oj, aw/ function as diphthongs in all accents of English (Giegerich 1992: 50-1).

[100] With respect to noncoronal homorganic nasal-obstruent clusters, the sequence nasal plus voiced obstruent is prohibited (*/mb/, */mv/, */ŋg/; cf. also chapter 8.2.1.2 below).

of two adjacent consonants, if C_1 is a stop and C_2 [–coronal], the two consonants must share the same place of articulation. Note that this generalization not only includes nasal stops but also accounts for the principal grammaticality of the codas /pt/, /kt/ where C_1 is an oral stop and C_2 is a coronal and /sp/, /sk/, /lp/, /lk/, /rp/, /rk/ where C_1 is not a stop and C_2 is not a coronal consonant. The codas */pk/, */kp/, */tp/, */tk/, however, where C_1 is an oral stop and C_2 is not a coronal, are disallowed.

In OT terms, this restrictions can be formalized as the partial constraint ranking shown in figure 6.18 below. Pulleyblank (1997: 64) introduces a family of syntagmatic constraints which are designed to impose restrictions on possible sequences of sounds. IDENTICAL CLUSTER CONSTRAINTS (ICC) refer to the four dimensions PLACE, VOICING, CONTINUANCY, and NASALITY, e.g. ICPL demanding that a cluster of consonants share the same place of articulation. By analogy to the Obligatory Contour Principle (cf. section 6.3.1 above), disallowing adjacent identical elements in URs, one could propose a family of syntagmatic constraints which require sequences of sounds to differ with respect to the four dimensions mentioned above. DIFFERENT CLUSTER CONSTRAINTS (DCC) may well turn out to be articulatorily grounded. The constraint DCPL thus prohibits heterorganic consonant clusters and thereby minimizes articulatory effort. With respect to SAE codas, its effects need to be restricted to sequences consisting of a stop plus a noncoronal.

(6.18) Tableau for <sink>

/sɪnk/	*C/STOP–CORDC	FAITHC_2	FAITHC_1	*C/STOPCOR
☞ sɪŋk			*	
sɪmt		*!	*	
sɪkt		*!	*	*
sɪnk	*!			

It is mandatory to split up FAITHFULNESS into FAITHC_1 and FAITHC_2 in order to show that /sɪŋk/ is a better candidate than /sɪmt/, although /mt/ in principle is a possible SAE coda (e.g. <undreamt>). This can be justified for two two reasons: first, a double violation of FAITHFULNESS should have more severe consequences than a single violation; and second, on the basis of crosslinguistic evidence Mohanan posits dominance relations with respect to place specifications. One of his generalizations states that the place of articulation of the second segment in a sequence is dominant with respect to that of the first (1993: 90; cf. detailed account in chapter 7.2 below). Ranked below FAITHFULNESS is a constraint militating against stop plus coronal codas, but due to its low rank, it cannot prevent an input coda /kt/ as in <act>, for example, from surfacing.

While the implicational statement above expresses a constraint on consonant combinations, Selkirk analyzes the data by postulating the following, even more general requirement, which is formulated as a positive statement: "The second consonant of the coda must be a coronal" (1982: 337). She is led to this strong claim by the observation that a coda consisting of two or more noncoronals is disallowed, and the maximum number of consonants permitted in English codas is three, with /st/ or /sθ/ in second or third place (ibid.: 338). Selkirk sees apparent counterexamples in cases like <wasp> and <risk> (cf. immediately above) and accounts for these by claiming that sequences of /s/ plus voiceless plosive count as single obstruent segments, just as they do in onsets (cf. section 6.3.1 above). Selkirk goes on to argue that cases like <elf>, <help>, <triumph>, <damp> and <rank> (cf. immediately above) likewise do not call into question the validity of the collocational restriction on bi-consonantal codas: if there is a simple vowel (one X-position) in the peak, the following sonorant also belongs to the peak, which results in codas composed of only a single consonant. Inflectional endings, which may extend the syllable template to four-consonant rhymes, are discounted by virtue of their being suffixes. Inflection attaches to a word category (stem) and not to the syllable and thus falls outside the domain of syllabification (1982: 338-40). Tempting as this claim may be, the rescue operations she has to stipulate to uphold it, i.e. to count sonorants as peak constituents if the peak has only a single X-position vowel, appear somewhat questionable. The preceding analysis thus seems preferable to Selkirk's.

The core syllable is governed by the SSP. If however the upper limit of permitted elements is exceeded, the violation does not involve any random set of consonants, but a well-defined natural class: only coronals may violate the core syllable pattern. It is this class of sounds that may appear as 'extra' elements of the syllable. A possible explanation is that if something 'extra' comes into play, languages will choose unmarked elements. In addition to being unmarked, coronals are characterized by a high degree of perceptibility, which appears to be particularly rewarding in clustering with other consonants (cf. section 6.2.2 above).

CHAPTER 7

An OT model of correspondences between phonetics *and* phonology

7.1 INTRODUCTION

In the preceding chapter, I argued that coronal unmarkedness is responsible for rich coronal inventories. On the other hand, it is the reason why coronals disappear in alternations. Using the example of place assimilation, I will briefly illustrate how alternational phenomena have been dealt with in previous models of generative phonology. After presenting data from the *SBCSAE*, I will point to some of the difficulties and apparent contradictions in connection with the behavior of coronals in place assimilation alternations. I will argue that these contradictions can be resolved if a functional approach is pursued and if the asymmetries involving coronals are analyzed as conflicts between articulatory and perceptual demands on the sides of the speaker and the hearer. I will present a model of grammar that contains the relevant grammatical correlates of the speech production and perception processes in the form of articulatory and perceptual representations and the formalizations of articulatory and perceptual demands as OT constraints. In chapter 8, this model is then used to account for weakening or lenition phenomena (assimilation and deletion) and strengthening or fortition phenomena (dissimilation and epenthesis) involving coronals in SAE.

Constraints referring to the facts of articulation had been invoked before the advent of OT. Westbury & Keating (1986) devised a model to account for the notion of naturalness with respect to stop consonant voicing. In simulating the effects of various articulatory conditions on voicing, they claim that their model provides the possibilities for defining and testing the notion ease of articulation. The conclusion

Westbury & Keating draw is that ease of articulation is not the primary determinant of phonetic form. Specifying the limiting properties of the articulation and perception mechanisms, they suggest that there must be other, equally powerful principles at work that are needed to account for phonetic form besides ease of articulation; these are "communicative efficiency, acoustic invariability, and perceptual requirements" (1986: 164). Cole & Kissebirth argue in a very similar fashion suggesting that for an adequate description of assimilation *two* forces must be recognized: perceptibility, i.e. features should be perceptible, and articulator stability, i.e. the changes from the neutral, steady state of articulators should be minimal, the two things, however, are inherently in conflict with one another (1994: 4).

Constraints referring to the facts of perception had been proposed before the advent of OT as well. Ohala (1990) presents results focusing on assimilation that add to the body of evidence, pointing to the crucial role of the hearer in initiating certain sound changes. Sound change appears to be a nonteleological phenomenon (McMahon 2000: 84), neither speaker nor hearer chooses, consciously or not, to make a change. Ohala compares sound change to medieval scribal errors in the copying of manuscripts (1990: 266). Bybee argues that a change may be made by hearers in *analogy* to some other form in the language (1994: 292; cf. also chapter 5.4 above). The fact that assimilation has perceptual roots has been discussed at length by many researchers, e.g. Ohala (1990), Byrd (1992; 1996), Hura et al. (1992).

7.2 PREVIOUS GENERATIVE ANALYSES OF PHONOLOGICAL ALTERNATIONS

Assimilation of place of articulation can serve as a good illustration of the changes in phonological theory from rules to representations to constraints (cf. also chapter 4). In early generative phonology, rules were used to depict the changes in phonological representations, to relate one level of representation to another (cf. chapter 4.1.1):

> The two representations [phonological and phonetic] are systematically related by phonological rules that delete, insert, or change sounds in precise contexts.
> (Kenstowicz 1994: 7)

In *SPE* terms, nasal place assimilation, for instance, is expressed by a rule stating that a nasal will have the same feature values (+ or –) with respect to coronality, anteriority, and backness as the following consonant:

(7.1) *SPE*-type nasal place assimilation rule (after McCarthy 1988: 86)

$$[+\text{nasal}] \rightarrow \begin{bmatrix} \alpha \text{ coronal} \\ \beta \text{ anterior} \\ \gamma \text{ back} \end{bmatrix} \quad / \;\_\_\_ \quad \begin{bmatrix} \alpha \text{ coronal} \\ \beta \text{ anterior} \\ \gamma \text{ back} \end{bmatrix}$$

(where α, β, γ are variables over + and –)

The deficiencies of this rule-based formalism soon became apparent and among them was the fact that a rule which is equally 'simple' (in terms of feature specifications) and should thus be equally 'good' may in fact be an implausible or even impossible rule (McCarthy 1988: 86):

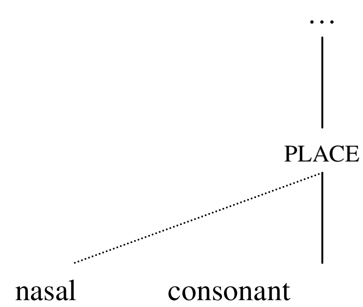
(7.2) Impossible nasal place assimilation rule (after McCarthy 1988: 86)[101]

$$[+\text{nasal}] \rightarrow \begin{bmatrix} \alpha \text{ back} \\ \beta \text{ coronal} \\ \gamma \text{ anterior} \end{bmatrix} \quad / \;\_\_\_ \quad \begin{bmatrix} \alpha \text{ coronal} \\ \beta \text{ anterior} \\ \gamma \text{ back} \end{bmatrix}$$

For obvious reasons, rules like these do not show *descriptive adequacy* let alone *explanatory adequacy* (cf. the discussion of Chomsky's three criteria for the evaluation of grammars in chapter 1.1.1 above). Subsequently, theories of representation gained priority (e.g. Anderson 1985; cf. also the discussion in chapter 4). McCarthy expresses this notion in the following way: "if the representations are right, then the rules will follow" (1988: 84). The type of representation McCarthy deems to be right is feature geometry showing nasal place assimilation as the insertion of an association line whereby the influence of a specific feature or node (in this case the PLACE node) is extended over a wider domain than just a single segment:

[101] McCarthy's criticism is of a very similar kind to that formulated by Schachter (1969) almost 20 years previously; cf. chapter 4.2.1 above.

(7.3) Nasal place assimilation in feature geometry (after McCarthy 1988: 87)



Applying Chomsky's three criteria to this representation, the level of *descriptive adequacy* is certainly achieved, since within the frameworks of feature geometry the coronal nasal (in fact all coronal consonants) is frequently represented as placeless, i.e. it either lacks the CORONAL articulator node (Avery & Rice 1989: 179-200) or even the superordinate PLACE node (Paradis & Prunet 1989: 319; Paradis & Prunet 1991: 6; also cf. discussion of coronal underspecification in chapter 4.3.1 and 4.3.4 above). The lack of place features then predisposes it to assimilate the place features of the following consonant.

(7.4) Examples of coronal stops assimilating to the place of articulation of a following noncoronal consonant in SAE (from *SBCSAE*)

a.	that boy	/tb/ → [pb]	(sbc0002.wav: 00'27")
b.	that means	/tm/ → [pm]	(sbc0002.wav: 01'01")
c.	don't you	/tj/ → [ʧ]	(sbc0004.wav: 05'36")
d.	that guy	/tg/ → [kg]	(sbc0002.wav: 08'49")
e.	did mom	/dm/ → [bm]	(sbc0001.wav: 21'25")
f.	did you	/dj/ → [ʤ]	(sbc0002.wav: 14'19")
g.	did go	/dg/ → [gg]	(sbc0013.wav: 08'39")
h.	ten percent	/np/ → [mp]	(sbc0002.wav: 11'19")
i.	ten minutes	/nm/ → [mm]	(sbc0002.wav: 20'18")
j.	in fact	/nf/ → [ɱf]	(sbc0002.wav: 04'04")
k.	ten grand	/ng/ → [ŋg]	(sbc0010.wav: 13'33")

When looking at the behavior of coronals in nasal place assimilation more closely, parallels can be observed between constraints on the distribution of consonants in

SAE syllable structure (if C_1 is a stop (nasal *or* oral) and C_2 is [–coronal], the two consonants must share the same place of articulation, cf. chapter 6.3.2 above) and constraints on the output of SAE assimilation processes. The generalization about place assimilation across words, i.e. within the domain of the phonological phrase (PhP), which can be derived from these data and which also governs SAE syllable structure (cf. chapter 6.3 above) is the following: if C_1 is a coronal stop (nasal *or* oral) and C_2, which follows C_1, is [–coronal], the place node of C_2 spreads to C_1.[102] Clearly, there is an asymmetry with respect to coronals: C_1 assimilates to C_2. The segments which undergo place assimilation (C_1) are coronal stops (nasal *and* oral), the segments which trigger place assimilation (C_2) are noncoronals.

As Mohanan (1991; 1993) points out, this type of asymmetry is not restricted to English but is extremely frequent crosslinguistically:

> It is quite common to find place assimilations in which coronals assimilate to non-coronals without non-coronals assimilating to coronals. (1991: 311)
> [C]oronals are more likely to undergo and less likely to trigger place assimilation than labials, and labials are more likely to undergo and less likely to trigger place assimilation than [dorsals]. (1993: 71)

Mohanan concludes from the facts of place assimilation that noncoronal place features are "dominant" or marked features since segments with dominant features will trigger spreading, while [coronal] is "non-dominant" or unmarked since segments with nondominant feature specifications will undergo spreading (1991: 311).

(7.5) Dominance scale with respect to PLACE (after Mohanan 1993: 91)

most dominant				**least dominant**
DORSAL	>	LABIAL	>	CORONAL

Not only are coronals least dominant with respect to place specification, followed by the more dominant labials, and finally the dorsals, which are most dominant on the scale, there are further dominance relations relevant for place assimilation. As can be

[102] Zwicky also comments on the parallels between place assimilations that apply in these two different domains—"within morphemes and across morpheme boundaries [sic]" (1972b: 281). Invoking a notion of Natural Phonology (cf. chapter 4.2.2 above), Zwicky assumes that the syllable-internal constraints discussed in chapter 6.3 above are restricted versions of the more general processes that are active in the larger domain of the PhP.

seen from the data in figure 7.4 above, C_1 assimilates to C_2 and not vice versa, so C_2 is dominant with respect to C_1. For example, an assimilation /mt/ → *[mp] as in <ham toast> that would give *[hæmpowst] is not only unattested in English but unlikely if not impossible crosslinguistically (Mohanan 1993: 81).

A possible explanation for this is offered by Ohala (1990). Notably, he sees the cause for this type of assimilation not in the domain of articulation (on the side of the speaker) but in perception (on the side of the hearer). Ohala reports that the place cues for plosives in prevocalic position are generally very strong whereas the place cues for word-final unreleased plosives are rather weak and thus frequently misperceived (1990: 261). This can be taken as an explanation for the assimilations a.-g. in figure 7.4 above since word-final plosives are typically produced without releasing the built-up air pressure, especially in cases when another stop follows. On the other hand, articulatory requirements can equally be held responsible—at least in part—for these assimilations. It is not implausible to presume that in order to save one articulatory gesture, the originally coronal plosives have labial and dorsal release gestures in anticipation of the following labial and dorsal stops (principle of minimization of articulatory effort; cf. chapter 5 above). For the nasals, which are quite distinct as a class but tend to be confused among themselves, Ohala assumes even less salient place cues than for unreleased plosives:

> [W]hen joined to a following stop it is not surprising that the listener has relatively less trouble hearing the nasal consonant as such but takes the place cue from the more salient stop release. (1990: 261)

This describes exactly what one finds in the data of figure 7.4 h.-k. above: hearers may perceive the nasal consonant (C_1) simply as a [+nasal] segment without a specification for PLACE, as it were, and simply rely on the following consonant (C_2) for the place cue. This also explains why /mt/ → *[mp] is an impossible assimilation. An experiment conducted by Ohala himself showed that when speakers of American English were asked to identify a sequence of two consonants, where C_1 and C_2 have different places of articulation, in 93% of the cases the place of articulation of C_2 dominated the percept over the place of articulation of C_1. That is, for /anpa/, for example, [ampa] was heard and not [anta] (1990: 263). As a basis for discussion, Mohanan offers the following statements regarding dominance:

(7.6) Proposed dominance relations in place specifications (after Mohanan 1993: 90)

a. an onset is dominant with respect to a coda
b. a less sonorous segment is dominant with respect to a more sonorous segment
c. the second segment in a sequence is dominant with respect to the first

Clearly, by referring to sonority and the so-called strength hierarchy[103] in statement b., Mohanan makes an effort to ground place assimilation in phonetics, that is, the stronger the segment, the better its chances of survival since it will more easily resist the forces that alter its properties than the weaker, nondominant segment. The weaker segments are both the more sonorous segments, i.e. they show a relatively high degree of periodic acoustic energy, and the more open segments, i.e. they offer relatively little resistance to airflow. On closer inspection, however, a. and c. are phonetically motivated statements as well: they are congruous with Ohala's (1990) observation presented above that the place cues for plosives in prevocalic, i.e. in onset position are stronger than the ones in postvocalic, i.e. in coda position.

The phonetically grounded dominance relations formulated by Mohanan (cf. figure 7.6 above) serve as explanations for alternations which actually make coronals disappear. A dominance scale like the one presented in figure 7.5 above is an expression of relative markedness or gradual markedness[104] and corresponds directly to Prince & Smolensky's notion of coronal unmarkedness (1993: 180) and subsequent constraint formulation (1993: 181): *PL/DOR , *PL/LAB >> *PL/COR (cf. chapter 4.3.4 above).

Inevitably, the asymmetrical behavior of coronals must be analyzed as a conflict between articulatory and perceptual demands and, as will be shown below, even as a conflict within the domains of articulation and perception respectively. A further investigation of assimilation clarifies how the sequence of the two adjacent

[103] It is generally claimed that plosives are stronger than fricatives, fricatives stronger than nasal stops, and nasal stops stronger than continuant sonorants (plosives > fricatives > nasal stops > continuant sonorants > glides > vowels), and that voiced plosives are weaker on the hierarchy than their voiceless counterparts (e.g. Zwicky 1972a: 275-82; Lass 1984: 177-83; Mohanan 1993: 101). This classification is important for resolving the issue of whether the voicing of voiceless plosives (e.g. /t/-voicing in SAE) is a weakening or strengthening alternation (cf. chapter 8.2.1.2 below).

[104] For convenience, I repeat Greenberg's (1966: 58-9) essential criteria for markedness presented in figure 4.6 of chapter 4.2 above. Marked elements are:

a. not as frequent both language-internally and crosslinguistically
b. subject to more restrictions in distribution
c. more complex in form
d. lost in neutralization alternations

consonants C_1 (coronal, target of place assimilation) and C_2 (noncoronal, trigger of place assimilation) is characterized by the fact that the two share the same place of articulation. From a physiological point of view articulatory effort is minimized since the energy for one extra gesture is saved. However, the product of the assimilation process (a labial or dorsal stop) is itself more complex from an articulatory perspective than the original coronal stop because the deviation in distance from the tongue position at rest to a coronal gesture is less than to a labial or dorsal gesture.

Therefore, following Mohanan (1993: 99) and Cole & Kisseberth's suggestion (1994: 4), I adopt a decomposition of Passy's (1890) general principle of *minimization of articulatory effort* (cf. chapter 5.3 above) into two parts, analogous to Flemming's (1997) binary division of the principle of *maximization of perceptual contrast* (cf. chapter 5.5.3 above). *Maximization of perceptual contrast* and *minimization of articulatory effort* are each divided into two subprinciples (cf. figure 7.7 below), which express the (potentially conflicting) quantitative and qualitative requirements of the two main principles respectively. Just as principles 1.a. and 1.b. are in conflict with each other (contrasting sounds share both an articulatory and a perceptual space and fitting more segments into the same space automatically reduces the perceptual distinctness of each single segment in that space; cf. figure 5.8 in chapter 5.5.3 above), 2.a. and 2.b. are also potentially conflicting principles.

(7.7) The two conflicting principles of *maximization of perceptual contrast* and *minimization of articulatory effort*, each decomposed into two conflicting sub-principles or goals

1. Maximize perceptual contrast:
    a. Maximize the number of contrasts.
    b. Maximize the distinctiveness of contrasts.
2. Minimize articulatory effort:
    a. Minimize the number of articulatory gestures.
    b. Minimize the deviation of articulatory gestures from configurations of least effort.[105]

[105] With this statement, Mohanan transforms Greenberg's sixth criterion for phonological markedness, articulatory (formal) complexity, into a goal which sound systems strive for (cf. discussion in chapter 4.2 above): "A particular articulation is to be considered more complex [and more marked] than some other if it includes an additional articulation defined in terms of departure of an organ from the position it normally has in the absence of speech." (Greenberg 1966: 70-1)

Given a choice between the sounds /m/ and /n/, a sound system might favor /n/ instead of /m/ on the grounds of 2.b. If, however, both /m/ and /n/ occur in the sound system and the sequences /mp/ vs. /np/ are to be evaluated (cf. Ohala's example /ampa/ vs. /anpa/ (1990: 263) discussed just above), 2.a. will favor /mp/. It is this sequence that saves one articulatory gesture and is not inferior perceptually since the manner feature [+nasal] without a PLACE specification[106] seems to be sufficient for the perception process, and the following consonant (C_2) is the one that hearers rely on for the place cue.

As an interim summary, I claim that place assimilation produces perceptually viable articulatory simplifications. Its targets thus far discussed are the coronal stops /t/, /d/, and /n/. It is commonly assumed that within the class of coronals the segments with the feature specification [+anterior] and [–distributed] are the truly unmarked or underspecified coronals (Paradis & Prunet 1991: 6; Yip 1991: 74; cf. also chapter 4.3 above), as opposed to /θ, ð, ʧ, ʤ, ʃ, ʒ/, which carry the feature specification [+distributed]. Employing this classification, the following coronals are unmarked coronals: /t, d, s, z, n/. The coronal /r/ which is commonly classified [–anterior] due to its [+retroflex] specification is excluded from the class of unmarked coronals. Likewise /l/ does not qualify on the grounds of its [+lateral] specification: laterals are neither [–distributed] nor [+distributed] because the feature is defined as appling to the midsaggital region of the oral cavity (Halle & Clements 1983: 6). If this definition were rejected, /l/ would be classified [+distributed], and would thereby be excluded in any case. Since place assimilation generally affects unmarked segments, it is not surprising that not all 13 SAE coronals are targets of simplification processes to exactly the same extent, but that there is a hierarchy even within the class of coronals with the most unmarked segments being the most likely ones to undergo an alternation.[107]

Turning to the fricatives now, these segments are reported to be perceptually more distinct as a class than the plosives or even the nasals, and are thus less susceptible to place assimilation (Hura et al. 1992: 66; Kohler 1992: 209; Byrd 1996:

[106] Cf. discussion of coronal underspecification in chapter 4.3.4 above; nasals are assumed to lack the articulator node (e.g. CORONAL) or even the superordinate PLACE node (ØPLACE → CORONAL). Coronals are prone to assimilation because they are underspecified at the point where spreading of the PLACE node takes place.

[107] There is also evidence from typology corroborating the assumption that the sound segments /t, d, s, z, n/ are indeed truly unmarked: these five are the segments with the highest crosslinguistic inventory frequency (Maddieson 1984; cf. also discussion in chapter 4.3.1 above).

217-31), especially the ones that are specified [+strident], i.e. /s, z, ʧ, ʤ, ʃ, ʒ/ (Hura et al. 1992: 70). Nevertheless, within the class of coronal fricatives, the unmarked alveolar fricatives /s/ and /z/, which carry the feature specifications [+anterior] and [–distributed], do participate in alternations.

(7.8) Examples of coronal fricatives assimilating to the place of articulation of a following consonant in SAE (from *SBCSAE*)

a.	this year	/sj/ → [ʃj]	(sbc0003.wav: 07'57")
b.	as your	/zj/ → [ʒj]	(sbc0008.wav: 03'25")
c.	guess she	/sʃ/ → [ʃʃ]	(sbc0002.wav: 05'03")
d.	does she	/zʃ/ → [ʒʃ]	(sbc0002.wav: 05'01")

Place is in fact assimilated, but unlike in the examples above involving stops, the class membership of the fricatives does not change; the assimilated segments are still coronals. Considering the fact that there are no dorsal fricatives in SAE, this comes as no great surprise. Looking at the labial consonants, there are two labial fricatives /f/ and /v/, but an assimilation like for example /sf/ → *[ff] as in <this friend> or /zf/ → *[vf] as in <as fast> is improbable if not impossible in SAE.

The data on place assimilation indeed support Yip's claim that five of the 13 SAE coronals should be regarded as unmarked coronals, i.e. /t, d, s, z, n/. The fact that the coronal stops /t, d, n/ assimilate to the place of articulation of a following coronal consonant simply adds further evidence: /t, d, n/ are dental before /θ/ as in <eighth>, <hundredth>, <tenth> and retroflex before /r/ as in <tree>, <dream>, <enroll> (Clements 1985: 236). Comparable to the cases involving /s, z/, class membership is retained, and the stronger, marked or dominant coronals act as trigger of the alternation while the weaker, unmarked or nondominant coronals are the targets of place assimilation. Considering again the conflicting perceptual and articulatory principles shown in figure 7.7 above, if speakers have a choice between a dental stop and an alveolar stop, they will probably choose the alveolar gesture on the grounds of principle 2.b., since this gesture deviates least from the position the articulator usually has in the absence of speech. If, however, both the stop and /θ/ occur in a cluster, principle 2.a. will possibly favor an overarching dental articulation for the sequence since one articulatory gesture is saved and perception is seemingly unaffected.

While assimilations like the ones above are commonly classified as weakening alternations, i.e. segmental integrity and unequivocality may be lost, there is a well-known strengthening process in English which involves the class of [+strident] fricatives and which serves exactly the opposite purpose. In the formation of SAE genitives and plurals, the original morphemes with the shape /əz/ are retained as allomorphs in nouns ending in a stem-final [+strident] fricative. Here perceptual disambiguation appears to be the prime issue; potential ambiguities arising from sequences such as /ss, ʧs, ʃs/ and /zz, ʤz, ʒz/ seem intolerable and are thus avoided.

I wish to show that assimilations originate as articulatory simplifications. For their phonologization, however, specific perceptual conditions have to prevail; the perceptual effects which these alternations produce must not be so severe as to result in stable phonemic contrasts.

> [I]n the production of connected speech only those articulatory modifications are possible that produce acoustic results of a low degree of perceptual salience, i.e. productive processes depend on auditory control. […]. What is not very distinctive for a listener anyway may be reduced by a speaker more easily to yield to the principle of economy of effort. (Kohler 1992: 209)

In certain contexts, e.g. in the case of the coronal fricatives assimilating to the place of articulation of a following consonant shown in figure 7.8 above, the change in place of articulation (or even in the laryngeal features, cf. chapter 8.2.1.2 below) of the segments does not produce perceptually salient contrasts that are significant enough for context-free phonemic evaluation. The manner features, which do not change (the assimilated segments in figure 7.8 above are still fricatives), seem to be sufficient for the identification of the relevant lexical item.[108] I will argue that there is a tendency for goal 2 (*minimize articulatory effort*) to focus on place or laryngeal features instead of on manner features for articulatory simplification, and a tendency

[108] Consider such an extreme case as the neutralization of the voiceless stops /p, t, k/ to a glottal stop characteristic of Cockney in word-final position; in this variety <whip, wit, wick> are all pronounced homophonously (Wells 1982: 323). Although the place contrast between the three stops is lost, the output of the alternation [ʔ] is still a stop. According to Lombardi (2001b) [glottal] is in fact the truly unmarked place of articulation ranked even below *PL/COR. Lombardi (2001b: 29) modifies Prince & Smolensky's (1993: 181) place hierarchy to include a lowest ranked constraint against the appearance of glottal stops: *PL/DOR , *PL/LAB >> *PL/COR >> * PL/PHA (cf. chapter 4.3.5 above).

for goal 1 (*maximize perceptual contrast*) to focus on manner features instead of on place or laryngeal features for perceptual distinction (cf. figure 7.7 above).

Phonetically oriented research has underlined the importance of perceptual information for an accurate account of place assimilation. Among the influential proposals highlighting the contribution of perceptual factors are Ohala (1990), whose approach is discussed above, Kohler (1991, 1992), Byrd (1992; 1996), Hura et al. (1992), Nolan (1992), Jun (1996), and Boersma (1998); Boersma (1998), for example, states that place assimilation is mostly restricted to nasals and explains this by saying that /n/ and /m/ are perceptually more alike than /t/ and /p/, and that hearers will rely less on place information for nasals than for plosives (1998: 144-5). However, it is not true that place assimilation is restricted to nasals crosslinguistically (cf. the SAE data in the tables of figures 7.4 and 7.8 above; these data show that coronal stops in general—and not only the coronal nasal—*and* coronal fricatives assimilate to the place of articulation of a following consonant, i.e. there is <te[m p]ercent> but there are also <tha[p b]oy, tha[k g]uy, thi[ʃ j]ear> in the data. Thus, a broader generalization seems preferable, namely that hearers of casual or fast speech generally rely less on place information than on manner features. Manner features are perceptually more salient than place features, i.e. /m/ is still a nasal, /p/ and /k/ are still plosives, and /ʃ/ is still a fricative.

Ohala (1990) also indicates that the environment in which assimilation takes place is of crucial relevance for whether an alternation may occur or not. A prevocalic position for stops, i.e. an onset position, is generally regarded as a position of prominence—or dominance, using Mohanan's (1993) term—whereas a stop in postvocalic or word-final position, mostly a coda position, is rather weak and frequently subject to misperception (1990: 261). Ohala's findings are in line with Kohler's (1992) account of place assimilation in German who presents a range of examples showing that the place of nasals, for example, is assimilated in <geben> [ge:bm] or <sagen> [za:gŋ] whereas the same process is unthinkable in cases like <gib nicht> *[pm] or <sag nicht> *[kŋ] although the respective segments are also adjacent. Kohler attributes the nonoccurrence of assimilation in the latter examples to the strong word-initial position of the nasal in <gib nicht>; only in a weak word-final position as in <geben> is assimilation possible (1992: 209). Further phonetic evidence for onsets as positions of prominence is supplied by an experimental study conducted within the model of Articulatory Phonology. Byrd (1996) measured the articulatory timing in English consonant sequences using EPG and found that onset clusters are less overlapped and less variable in their timing than coda clusters (1996: 217-31). That is, onset consonants are more stable and less susceptible for alternations like assimilation to take place than coda constituents.

In terms of an optimality-theoretic investigation there is additional support from Beckman's analysis of *positional faithfulness* (1999). She assumes that FAITHFULNESS constraints (F) are sensitive to positions of prominence. That is, depending on the position of a specific segment, the output string will be faithful to the input or else unfaithful.

(7.9) Privileged vs. nonprivileged positions in phonological systems (after Beckman 1999: 3)

privileged positions	**nonprivileged positions**
root	affix, clitic, function word
root-initial syllable	noninitial syllable
stressed syllable	unstressed syllable
syllable onset	syllable coda
tense/long vowel	lax/short vowel

Beckman establishes a list of privileged phonological positions which "enjoy some perceptual advantage in the processing system, via either psycholinguistic or phonetic prominence, over the complement of non-privileged positions" (1999: 3). Both types of position are included in the table above. The first two positions in the table are attributed psycholinguistic prominence, given that these bear the heaviest burden of lexical storage, lexical access and retrieval, and processing (Beckman 1999: 3), while the other three are assumed to be positions of phonetic prominence. In such privileged positions, featural and segmental contrasts frequently remain faithful to the input while contrasts may be suspended in nonprivileged positions. Segments found in dominant positions often act as triggers for alternations, for example place assimilation, while those in nondominant positions may assume the role of target of place assimilation. Thus, positional privilege is not only determined by perceptual salience, that is by phonetic factors, but also by an array of phonological phenomena:

(7.10) Phonological asymmetries diagnostic of positional privilege (after Beckman 1999: 4)

- positional maintenance of contrasts which are neutralized elsewhere
- positional triggering of phonological alternations
- positional resistance to alternations which apply elsewhere

Phonological asymmetries with respect to coronals, e.g. in place assimilation, are analyzed as dependent on at least two parameters: the types of segment involved (their perceptual and articulatory qualities) *and* the positions in which these segments occur. Coronal asymmetries on this view arise from the fact that positional FAITHFULNESS constraints dominate context-free FAITHFULNESS constraints and MARKEDNESS constraints, e.g. Prince & Smolensky's (1993: 181) context-free markedness hierarchy: *PL/DOR , *PL/LAB >> *PL/COR (Beckman 1999: 4).

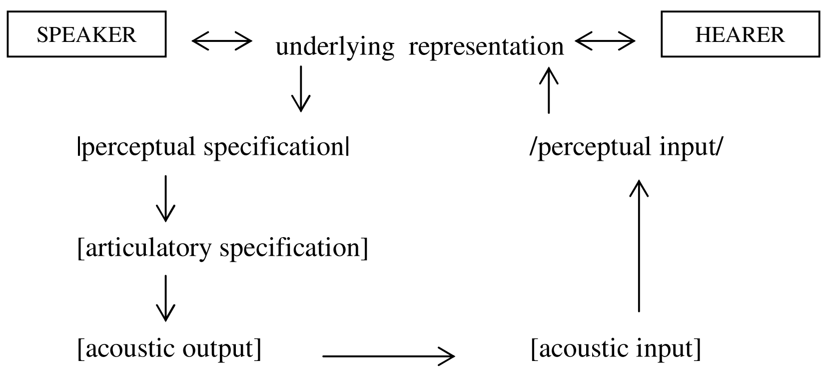
7.3 PERCEPTUAL AND ARTICULATORY CONSTRAINTS IN AN OT GRAMMAR

What is needed for an adequate phonological description within an OT model are thus constraints referring to the contextual positions of the segments involved and constraints referring to both the perceptual *and* the articulatory requirements of speech, which are in constant conflict. Speech perception and production must also involve a process of translation between underlying representations (URs) of utterances and the peripheral output and input representations. Figure 7.11 below shows a model of communication containing the essential components which are subsequently formalized as OT constraints.

7.3.1 Communication model

An ideally identical *underlying representation* of an utterance is stored in the lexicons of language users (SPEAKER and HEARER). When a speaker wants to produce a linguistic token, a *perceptual specification* in terms of perceptual features, which corresponds to the UR, will be converted into an *articulatory specification* in terms of articulatory gestures (a gestural score in Browman & Goldstein's (1989) model; cf. figure 7.13 below). The product of the articulation process is an *acoustic output*. The acoustic output then constitutes the *acoustic input* to the hearer who in turn constructs a *perceptual input* which (ideally) should match the UR of the speaker, provided that the information flow works successfully. The step from the perceptual input to the UR marks the recognition and categorization of a given linguistic item. A further component, the perceptual result or output, could be added on the side of the speaker who is not only perceived by others but who also perceives herself speak (Boersma 1998: 143-4).

(7.11) Communication circuit from underlying representation (SPEAKER) to underlying representation (HEARER)



This functions as the speaker's—who is now the hearer—own perceptual input which can be checked against the UR of the intended utterance for the purpose of self-monitoring, e.g. to judge whether the principle of m*aximization of perceptual contrast* has been obeyed. In case the two do not match, a change might be necessary with respect to the perceptual and articulatory specifications of the utterance.

External evidence for a communication model like the one given in figure 7.11 above comes from research in aphasic speech. On the basis of numerous individual disorders, a model is assumed that distinguishes between a "phonological output lexicon" and a "phonological input lexicon" (Lesser & Milroy 1993: 56-8). The phonological input lexicon is accessed after an auditory analysis of an item has been accomplished and is then matched with the "semantic system" (ibid.) in order to get at the specific meaning of the item. This input lexicon seems akin to the *perceptual input* of figure 7.11 above, which is constructed by hearers from an acoustic input. In turn, if a given item is to be produced, the phonological output lexicon is then accessed, which provides—one might assume—an *articulatory specification* of how the item should be assembled from articulatory gestures, eventually leading to an acoustic output (cf. figure 7.11 above).

Since the needs of perception generally aim at preserving contrast and militate against the loss of distinctivity (cf. figure 7.7 above), I will argue that these goals are best expressed as FAITHFULNESS constraints, which try to ensure that a phonological contrast as specified in UR will indeed turn up at the surface. The needs of articulation, on the other hand, which generally aim at reducing the expenditure of articulatory energy and thereby strive to weaken phonological contrast, will analogously be formalized as MARKEDNESS constraints, trying to ensure that only the most unmarked structures will surface.

7.3.2 Perceptual and articulatory representations

Proceeding from the UR of an utterance, I will now describe the nature of the three major aspects of the communication model. First, a *perceptual specification* is made in terms of perceptual features, i.e. of how the utterance should sound. As discussed in chapter 3.1.1 above, Jakobson, Fant & Halle are convinced of the primacy of the perceptual dimension for the determination of the set of distinctive features (1951: 12). Perceptual features refer to properties such as continuancy, sonorancy, stridency, nasality, periodicity of sound, friction, silence, and release bursts.

(7.12) Perceptual specification for |tɛns| (after Boersma 1998: 28)

<table>
<tr><td></td><td>|t|</td><td>|ɛ|</td><td>|n|</td><td>|s|</td></tr>
<tr><td>timing</td><td>C or X</td><td>V, X, or μ</td><td>C, X, or μ</td><td>C, X, or μ</td></tr>
<tr><td>CORONAL</td><td>burst</td><td></td><td colspan="2">+</td></tr>
<tr><td>aspiration</td><td>+</td><td colspan="3"></td></tr>
<tr><td>friction</td><td colspan="3"></td><td>sibilant</td></tr>
<tr><td>sonorancy</td><td></td><td colspan="2">+</td><td></td></tr>
<tr><td>nasality</td><td colspan="2"></td><td>+</td><td></td></tr>
<tr><td>F1</td><td></td><td>open mid</td><td></td><td></td></tr>
<tr><td>F2</td><td></td><td>max</td><td></td><td></td></tr>
</table>

As Chomsky & Halle pointed out, these features may come in numerous possible gradings but specific languages choose a limited number of +/– categories from the scales (1968: 164). A |t|[109], for example, is specified for one thing by the presence of aspiration and a release burst—both the formation of the constriction and the release burst are reported to be perceptually salient (Stevens 1998: 243-55; cf. also chapter 3.3.2 above)—for another by the absence of friction, sonorancy, nasality, etc. A perceptual specification not only includes a specification for each individual segment category[110], but also a specification of the timing of each segment, i.e. X stands for an X-position on the timing tier (cf. chapter 6.3 above), and μ for the mora as a weight unit (Bird & Klein 1990: 48). A perceptual specification for <tense> might thus be similar to the one shown schematically in figure 7.12 above.

Second, suitable representations of *articulatory specifications* are provided by the model of Articulatory Phonology (Browman & Goldstein 1986; 1989; 1990; 1992; cf. also chapter 3.3.3 above) in the form of gestural scores. Browman & Goldstein (1986) are very explicit in their analysis of the appearance of 'epenthetic stops' like the coronal [t] shown in figure 7.13 below:

> [T]he occurrence of so-called epenthetic stops in English words like *tense* is dialect-dependent. […] such 'stops' are to be analysed in terms of variation in the relative timing of oral and velic gestures (rather than actual segment insertion; […]). In general, then, languages can differ from one another in the timing of (roughly the same) articulatory gestures. (1986: 221).

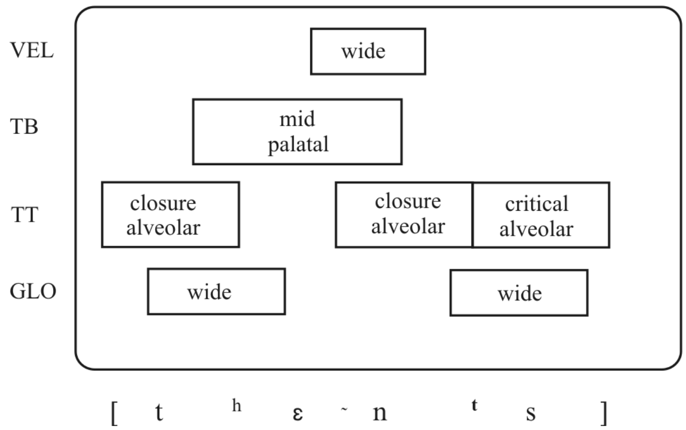
The representation shown in figure 7.13 below has gestural overlaps between each of the four segments involved. For example, the delay in the onset of voice (wide glottis) after the release of the alveolar closure gesture is the reason for the aspiration of [t^h]. The widening of the velum before the next alveolar closure gesture gives rise to a partial anticipatory nasalization of the vowel [ɛ], indicated by the symbol ~. And finally, the lowering of the velum prior to the release of the alveolar closure gesture causes the 'epenthetic' plosive [^{t}s] to appear after the nasal (Bird & Klein 1990: 49).

---

109 The vertical lines are used to refer to perceptual specifications.

110 There seems to be quite unequivocal evidence that a UR to which the perceptual specification must correspond involves discrete and abstract segment-sized units (represented by the vertical lines to the left and right of each segment in figure 7.12 above). The necessity for discrete categories is derived from the extreme variability to which the speech signal is subjected, especially on the side of the perceiver, and from the finite storage capacity of the human memory system (e.g. Cutler 1992: 290).

(7.13) Gestural score for <tense>[111] (after Bird & Klein 1990: 49)



Third, the representation in figure 7.14 below shows the *perceptual input*, which is perceived by the hearer in terms of perceptual features, as an uncategorized event, represented between slashes. Anticipatory nasalization of the vowel is again indicated by the symbol ~, 'cont' stands for a [+continuant] coronal /s/.

(7.14) Perceptual input for /t^hɛ~n^ts/ (after Boersma 1998: 29)

<table>
<tr><td></td><td colspan="2">/t^h</td><td colspan="2">ɛ~</td><td colspan="2">~n</td><td>^{t}s/</td></tr>
<tr><td>silence</td><td colspan="4"></td><td>+</td><td colspan="2"></td></tr>
<tr><td>CORONAL</td><td>burst</td><td colspan="4"></td><td colspan="2">burst cont</td></tr>
<tr><td>aspiration</td><td>aspiration</td><td colspan="6"></td></tr>
<tr><td>friction</td><td colspan="6"></td><td>sibilant</td></tr>
<tr><td>sonorancy</td><td></td><td colspan="4">+</td><td colspan="2"></td></tr>
<tr><td>nasality</td><td colspan="2"></td><td colspan="3">+</td><td colspan="2"></td></tr>
<tr><td>F1</td><td></td><td colspan="2">open mid</td><td colspan="4"></td></tr>
<tr><td>F2</td><td></td><td colspan="2">max</td><td colspan="4"></td></tr>
</table>

A schematic constraint tableau containing the three major aspects of communication—specification, articulation, and perception—thus takes on the following form:

[111] The articulator L (lips) is not included in the score because it remains more or less inactive during the articulation of <tense>; VEL, TB, TT, and GLO correspond to the articulators velum, tongue body, tongue tip, and glottis respectively.

(7.15) Schematic constraint tableau: Constraint$_1$ dominates Constraint$_2$ (C_1 >> C_2) (after Boersma 1998: 146)

|spec|	C_1	C_2
[art$_1$] /perc$_1$/	*!	
☞ [art$_2$] /perc$_2$/		*

Categorization in terms of discrete, segment-sized units takes place only afterwards in the hearer's recognition system which converts the perceptual input into a matching UR. I will comment on the categorization process in chapter 8.2.1.1 below.

7.3.3 Formalization of perceptual and articulatory representations as OT constraints

Based on the discussion in section 7.2 above, the following implicational statements can be set up. They serve as a basis for the formulation of constraints which mirror the phonetic motivations of phonological patterns and their direct integration into phonological analysis:

- Targets of place assimilation (PLACE): If dorsals are targets so are labials, and if labials are targets so are coronals (cf. Mohanan's (1993: 91) dominance scale).

  DORSAL[112] > LABIAL > CORONAL
- Targets of place assimilation (MANNER): If fricatives are targets so are plosives, and if plosives are targets so are nasals (cf. the high perceptual distinctivity of fricatives; Hura et al. 1992: 66; Kohler 1992: 209; Byrd 1996: 217-31).

  FRICATIVE > PLOSIVE > NASAL
- Targets of laryngeal assimilation (VOICING): If dorsals are targets so are labials, and if labials are targets so are coronals. I argued in chapter 6.3.2 above that only coronals are allowed to violate the SSP and the quantity constraints of the SAE syllable. This means that if the core syllable is expanded, e.g. by the processes of derivation or inflection, then the relatively unmarked coronals are used. These coronals then are potential targets of laryngeal assimilation.

  DORSAL > LABIAL > CORONAL

[112] Dorsals are at the top of the hierarchy because of crosslinguistic evidence presented by Mohanan (1993), because of the facts of place assimilation in Korean (Jun 1996), and because they have the additional acoustic feature [compact] (Jakobson 1949: 284) which makes them perceptually more salient than both labials and coronals (cf. chapter 5.5.3 above).

- Targets of assimilation or weakening (POSITION): If onsets are targets so are codas (Ohala 1990: 261; Kohler 1992: 209; Mohanan 1993: 90; Byrd 1996: 217-31; Beckman 1999: 3).

  SYLLABLE ONSET > SYLLABLE CODA

- Targets[113] of assimilation or weakening (POSITION): If stressed syllables are targets so are unstressed syllables (Kahn 1976: 56-61; Beckman 1999: 3).

  STRESSED SYLLABLE > UNSTRESSED SYLLABLE

These crosslinguistic implications are formalized as OT constraints in accordance with Prince & Smolensky's postulate that OT is subdivided into two subtheories (1993: 67; cf. discussion in chapter 5 above): the theory of substantive universals of phonological well-formedness (gradient phonetic phenomena) which are converted into rankable constraints in the theory of formal universals (categorical phonological phenomena). In Prince & Smolensky's terms (1993: 68), these implicational statements represent universal harmony scales (cf. also the table shown in figure 5.11 above) which express the crosslinguistic well-formedness of units (e.g. features, segments, onsets, nuclei, codas) in specific structural contexts.

Harmonic rankings of FAITHFULNESS constraints, which assess identity of correspondent elements by means of a family of input-output (IO) correspondence constraints, e.g. IDENTIO/PLACE (cf. presentation of Correspondence Theory in chapter 5.5.2 above; McCarthy & Prince 1995: 16-7), are generally motivated by the principle of *maximization of perceptual contrast*:

- The FAITHFULNESS constraint enforcing identity for place features in dorsals is universally ranked above the corresponding FAITHFULNESS constraints for labials and coronals (cf. the crosslinguistic tendency for coronals to lose their place of articulation in favor of a labial or dorsal place feature; Mohanan 1991: 311; Mohanan 1993: 71).

  IDENTIO/PLDOR >> IDENTIO/PLLAB >> IDENTIO/PLCOR[114]

- The FAITHFULNESS constraint enforcing identity for place features in fricatives is universally ranked above the corresponding FAITHFULNESS constraints for

[113] I will refrain from formulating implications for triggers since it is mostly the marked, noncoronal segments that act as triggers (Mohanan 1993: 71). If, however, this was done for PLACE in place assimilation, the result would be the exact inverse of the target implication: CORONAL > LABIAL > DORSAL, i.e. if coronals are triggers so are labials, and if labials are triggers so are dorsals.

[114] In cases where faithfulness does not refer to specific place features in the hierarchy, a general FAITHFULNESS constraint IDENTIO/POA (place of articulation) can be postulated.

plosives and nasals (cf. the crosslinguistic unattestedness of cases like *[anta] instead of [ampa] for /anpa/; Ohala 1990: 263; Mohanan 1993: 81).

IDENTIO/PLFRI >> IDENTIO/PLPLO >> IDENTIO/PLNAS[115]

- The FAITHFULNESS constraint enforcing identity for laryngeal specifications in dorsals is universally ranked above the corresponding FAITHFULNESS constraints for labials and coronals.

  IDENTIO/LARDOR >> IDENTIO/LARLAB >> IDENTIO/LARCOR[116]

- The FAITHFULNESS constraint enforcing identity for features (F) of segments located in syllable onsets is universally ranked above the corresponding FAITHFULNESS constraint for syllable codas (Beckman 1999: 20; Lombardi 1999: 270).

  IDENTIO/FONSET >> IDENTIO/FCODA

- The FAITHFULNESS constraint enforcing identity for features (F) of segments located in stressed syllables is universally ranked above the corresponding FAITHFULNESS constraint for unstressed syllables (Beckman 1999: 127).

  IDENTIO/FSÝL >> IDENTIO/FSỲL

Harmonic rankings of MARKEDNESS constraints, e.g., Prince & Smolensky's (1993: 181) constraint formulation *PL/DOR, *PL/LAB >> *PL/COR, are generally motivated by the principle of *minimization of articulatory effort*.

- Diverging slightly from Prince & Smolensky's (1993) proposal of constraints militating against a specific place of articulation, the MARKEDNESS constraints enforcing coronal unmarkedness that I will use are formulated as bans on specific articulatory gestures. A set of constraints referring to the active articulators (cf. chapter 3.3.3 above) is more in line with the framework of Articulatory Phonology (Browman & Goldstein 1986; 1989; 1990; 1992), which supplies the articulatory representations for this investigation.

  *GESTURE/DOR >> *GESTURE/LAB >> *GESTURE/COR

- Pulleyblank introduces a class of MARKEDNESS constraints, the so-called IDENTICAL CLUSTER constraints[117] (ICCs), whose purpose is to represent the

[115] In cases where faithfulness does not refer to specific manner features in the hierarchy, a general FAITHFULNESS constraint IDENTIO/MOA (manner of articulation) can be postulated.

[116] In cases where faithfulness does not refer to specific place or manner features in the segments involved, a general FAITHFULNESS constraint with respect to voicing IDENTIO/LAR (laryngeal) can be postulated.

forces that "impose articulatory inertia" with respect to the four featural dimensions PLACE, VOICING, CONTINUANCY, and NASALITY (1997: 64). ICCs inhibit changes in the positions of the articulators during the production of adjacent sound segments. Being in constant conflict with IDENTITY constraints, one of their prime functions is to enforce assimilatory alternations. However, the definitional prerequisite that clusters be identical with respect to some feature makes ICCs unsuitable for accounting for partial assimilations where two segments are made to be more similar but not identical concerning the property in question, e.g. in <thi[ʃ j]ear>. Instead of ICCs, I thus propose a similar constraint family with a slightly wider scope, namely ASSIMILATE/PLACE, ASSIMILATE/VOICE, etc. The advantage of an ASSIMILATE constraint family is obvious in that all types of assimilatory alternations can be represented including what has formerly been called allophonic, phonemic, and coalescent assimilation (e.g. Jones 1960: 219). On the other hand, a conceivable disadvantage of postulating ASSIMILATE constraints is a loss of generality with respect to the more static properties of sound segments, e.g. phonotactic constraints (cf. chapter 6.3 above). Looking at the distribution of segments in syllable onsets and codas, identical clusters are a frequent structural requirement in SAE.

- Based on the so-called Production Hypothesis[118], Jun (1996) alternatively proposes a whole family of MARKEDNESS constraints which are joined together under the even more general term WEAKENING constraints. They have the effect of conserving articulatory effort on the side of the speaker by reducing or eliminating consonantal gestures, which then may, for example, lead to place assimilation (1996: 223). WEAKENING constraints are in constant conflict with PRESERVATION constraints evaluating perceptual properties (1996: 223). A drawback of Jun's proposal is that WEAKENING constraints are not further subcategorized with respect to their specific functions, e.g. assimilation *or* deletion. It seems safe to claim that all constraints having as their aim the reduction of articulatory effort, such as IDENTICAL CLUSTER constraints and ASSIMILATE constraints, can be subsumed under the large family of WEAKENING

[117] There is a synonymous term for this family of constraints in the OT literature, which is AGREE; Lombardi (1999), for instance, has the constraints AGREE/PLACE and AGREE/VOICE.

[118] The Production Hypothesis states that "[s]peakers make more effort to preserve the articulation of speech sounds with powerful acoustic cues, whereas they relax in the articulation of sounds with weak cues." (Jun 1996: 224)

constraints. By analogy, the constraints that work in favor of maximizing segmental contrast can be said to belong to the family of STRENGTHENING constraints (cf. chapter 8.3 below). Typical language-specific rankings of WEAKENING constraints according to Jun (1996: 224) are:

a. WEAK >> PRES/PLPLO >> PRES/PLNAS (all stops can potentially be targeted by place assimilation; this is the case for SAE)
b. PRES/PLPLO >> WEAK >> PRES/PLNAS (only nasals targeted)
c. PRES/PLPLO >> PRES/PLNAS >> WEAK (neither is targeted)

Again employing place assimilation as a suitable example for illustrating the workings of constraints, a complete communication event for <input> from underlying representation (SPEAKER) to underlying representation (HEARER), as shown in figure 7.11 above, could be illustrated as follows:

(7.16) Schematic communication event for <input>

specify:	|ɪnpʊt|
articulate:	[ɪmpʊt]
ranked CONSTRAINTS (=OT grammar):	FAITH IO/MOA FAITH IDENTIO/PLONSET >> IDENTIO/PLCODA FAITH IO/PLDOR FAITH IO/PLLAB MARK ASSIM/PL FAITH IO/PLCOR MARK *GEST/DOR >> *GEST/LAB >> *GEST/COR
perceive:	/ɪmpʊt/
categorize:	|ɪnpʊt|

The main stages of the communication event captured in figure 7.16 above are specification, articulation, perception, and categorization. The lexical representation for <input> is assumed to be |ɪnpʊt|. What is articulated, however, and what is perceived is [ɪmpʊt] and /ɪmpʊt/ respectively, that is, the processes of articulation and perception are governed by the OT constraints; this is exactly where their sphere of influence is located. I also assume that the average SAE hearer upon perceiving

/ɪmpʊt/ will categorize it as |ɪnpʊt|, in this way recovering the intended meaning of the utterance.

CHAPTER 8

Alternational properties of coronals in SAE

8.1 INTRODUCTION

In this chapter, I will attempt to analyze the two main, opposing types of alternation that are found in coronals in SAE, using the grammar model presented in the previous chapter. First, weakening or lenition phenomena, i.e. assimilation and deletion, are discussed in section 8.2 below. These alternations generally function to reduce the contrasting properties of a sound with respect to (a) neighboring sound(s). Browman & Goldstein (1989: 216-7) claim that the timing of articulatory movements is adjusted in such a way as to have the perceptual effect of place assimilation on coronals in coda position or even the deletion of a sound, e.g. <tents> [tɛnts] → [tɛns]. Second, strengthening or fortition phenomena, i.e. epenthesis and dissimilation, are discussed in section 8.3 below. These generally function to intensify the contrasting properties of a sound with respect to (a) neighboring sound(s), possibly by adjusting the timing of articulatory movements so as to have the perceptual effect of inserting an additional sound, e.g. <tense> [tɛns] → [tɛnts] (Bird & Klein 1990: 49), or rendering the neighboring sounds less similar.

In his paper titled "What Is Lenition?", Bauer (1988) shows that a unified characterization of what constitutes *weakening* and *strengthening* is one of the controversial issues of phonological theory. He stresses the necessity of distinguishing between a phonetic and a phonological definition of lenition and fortition. I think Bauer is right when he argues that a phonological definition is not unambiguously possible since 'strength' or 'weakness' then depends on the function of a segment in a given sound system, which may be subject to a large degree of variation crosslinguistically, and not on the segment's phonetic structure. Depending on the perspective of the linguist, /t/-voicing in SAE may be seen as a lenition or a fortition. On the one hand, voiceless obstruents usually count as the stronger segments (cf. discussion of sonority in chapter 6.3 above), on the other hand voiced obstruents require one extra (laryngeal) gesture. Within the conceptual frame of markedness theory, the very term *weakening* suggests that this alternation should give rise to less complex, and thus unmarked segments. A voiced obstruent, however, is marked with respect to its voiceless counterpart (cf. Greenberg's (1966: 58) criteria for markedness presented in chapter 7.2 above). According to Bauer, if a circular definition is to be avoided, the only halfway satisfactory definition of lenition is a *phonetic* specification in terms of some sort of reduction in constriction degree or constriction duration, which then necessarily *differs* for consonants and vowels (1988: 389-90). This type of definition would not arrive at any consistent pattern for vowel lenition phenomena, since a weak vowel is *not* the result of a smaller degree of constriction in the oral cavity. It seems to be valid for consonants, however, and will be adhered to implicitly for the purpose of this investigation, at least with respect to the demands of articulation.

Fortitions are aimed at maximizing perceptual contrast for the hearer, and they may have the concomitant effect of minimizing articulatory effort for the speaker. Although lenitions, on the other hand, are said to have this latter effect as their primary goal, there may in fact be cases of lenition where perceptual distinctivity is enhanced as well. It is generally held that the effects of fortitions become salient in slower, more formal speech styles, while lenitions are more likely to operate in faster, more colloquial styles (Donegan & Stampe 1979; Browman & Goldstein 1990; Kohler 1991; Kirchner 1998). This may be part of the reason why examples of assimilation and deletion in the *SBCSAE* are much more frequent than examples of epenthesis and dissimilation.

8.2 MINIMIZATION OF SEGMENTAL CONTRAST

One of the main assumptions is that the weakening alternations which I will discuss in this section—assimilation and deletion—are both characterized by a loss of contrast with respect to some feature(s) or other(s). A further assumption is that these alternations are mainly connected speech phenomena, that is, if speech was articulated very carefully, word by word as it were, and according to the lexical specifications of underlying representations, the number of lenition phenomena would be greatly reduced. Taking assimilation as an example, only lexically determined forms like <impossible> would be left for investigation. However, when naturally occurring speech is analyzed, such as the conversations recorded for the *SBCSAE*, assimilation is one of the main means of guaranteeing fluidity of speech.

Generally, it can be claimed that weakening alternations work in favor of the demands of articulation and—at least to a certain extent—to the disadvantage of the needs of perception: by minimizing articulatory effort, perceptual contrast is reduced or even abolished. I will show under which precise circumstances it is possible to dispense with perceptual distinction without at the same time putting successful communication in jeopardy.

8.2.1 Assimilation

OT analyses embedded in the general framework of Correspondence Theory view assimilation as a correspondence relation $\mathscr{R}$ between input and output strings S_1 and S_2 (McCarthy & Prince 1995: 16; cf. also chapter 5.5.2 above). The identity of correspondent elements is assessed by a family of FAITHFULNESS constraints; IDENTITY–INPUTOUTPUT(FEATURE) constraints demand identical specifications between input and output strings with respect to some feature F and militate against assimilation. Only if IDENT–IO(F) is outranked by some MARKEDNESS constraint can assimilation take place. Rankings, it must be assumed, will vary according to speech tempo and style, i.e. fast, slow, casual, formal, etc. Thus, the ranking of specific FAITHFULNESS constraints and specific MARKEDNESS constraints may indeed be reversed depending on the register involved. It is hypothesized that for a casual speech utterance of <input>, for example (cf. also figure 7.16 in chapter 7 above), the MARKEDNESS constraint ASSIM/PLACE must crucially dominate the FAITHFULNESS constraint demanding identity for place features in coronals. The SAE-specific ranking of constraints for this speech style determines the optimal candidate. As shown in figures 7.16 above and 8.1 below, the highest ranked constraint is a FAITHFULNESS constraint pertaining to the demands of perception, a general constraint requiring input-output identity for manner of articulation features (MOA).

(8.1) Place assimilation in <input> (a selection of candidates and constraints)

|ɪnpʊt|	IO/MOA	IO/PLLAB	ASSIM/PL	IO/PLCOR
[ɪnsʊt] /ɪnsʊt/	*!	*		
[ɪntʊt] /ɪntʊt/		*!		
[ɪnpʊt] /ɪnpʊt/			*!	
☞ [ɪmpʊt]/ɪmpʊt/				*

If, for example, <income> were considered by the constraint hierarchy, the dorsal plosive would surface faithfully just as the labial does in <input> since IO/PLDOR is ranked even higher than IO/PLLAB in the hierarchy (cf. chapter 7.3.3 and figure 7.16 above). The MARKEDNESS constraint ASSIM/PL, which pertains to the demands of articulation, intervenes between IO/PLLAB and IO/PLCOR to make sure that the coronal nasal assimilates the place feature of the dorsal to give [ɪŋkʌm].

8.2.1.1 Place assimilation

Place assimilation is more varied and more frequent than other types of assimilation in English, such as manner or voice assimilation (Roach 1983: 106; Gimson[119] [1962] 2001: 285-6) and one of the most common phonological alternations in the world's languages (Spencer 1996: 149).

Many analyses of place assimilation limit themselves to investigations of *nasal* place assimilation (e.g. Padgett 1995a; Pater 1995; Boersma 1998: 144-6; Kager 1999: 59-64). I assume this to be the case because nasals are the class of consonants which are most susceptible to assimilation (Lombardi 2001b: 19; cf also the constraint hierarchy presented in chapter 7.3.3 above: IDENTIO/PLFRI >> IDENTIO/PLPLO >> IDENTIO/PLNAS). Since nasal place assimilation is extremely common in the world's languages (Lass 1984: 199; Kager 1999: 60), it is frequently treated as the default type of place assimilation (Sagey 1986; Padgett 1995a). The SAE data in the tables of figures 7.4 and 7.8 in chapter 7 above show, however, that coronal stops in general (coronal nasals *and* plosives) and *even* coronal fricatives assimilate to the place of articulation of a following consonant, i.e. in the data, there is <te[m p]ercent> but there are also <tha[p b]oy, tha[k g]uy, and thi[ʃ j]ear>. The generalization that holds for place assimilation involving nasals can be extended to

[119] The sixth edition (2001) of Gimson's book *Pronunciation of English* is edited by Alan Cruttenden.

place assimilation in general in SAE. Hearers of casual or fast speech generally rely less on place information than on manner features. Manner features are perceptually more salient than place features, i.e. /m/ is still a nasal, /p/ and /k/ are still plosives, and /ʃ/ is still a fricative, which is expressed in the ranking of the relevant constraints.

(8.2) Place assimilation in <ten percent> using the place hierarchy (sbc0002.wav: 11'19")

|tɛn pɜrsɛnt|	IO/MOA	IO/PLLAB	ASSIM/PL	IO/PLCOR
[tɛm fɜrsɛnt] /tɛm fɜrsɛnt/	*!	*		*
[tɛn tɜrsɛnt] /tɛn tɜrsɛnt/		*!		
[tɛn pɜrsɛnt] /tɛn pɜrsɛnt/			*!	
☞[tɛm pɜrsɛnt] /tɛm pɜrsɛnt/				*

Place assimilation as in <ten percent> is attributed by Browman & Goldstein (1989: 215-7) to the hiding of an articulatory gesture due to increased gestural overlap in connected speech. According to the notional framework of Articulatory Phonology, the coronal closure gesture is produced (their example is <seven plus>) but is hidden by the following labial plosive. The hidden gesture is then "*perceived* as assimilation" (1989: 216; emphasis added), as the labial nasal /m/.

These assumptions are confirmed by an examination of the perceptual effects of gestural overlap in consonant clusters undertaken by Byrd (1992). In order to determine what is causing the percept of place assimilation, the only variable in the experiment was the *degree* of overlap within a consonant cluster, not the gestural magnitude. All gestures were actually produced with the same magnitude, which means that gestures were not changed or even deleted (cf. also discussion of deletion in section 8.2.2 below). In the typical assimilation environment VC_1C_2, Byrd found that C_1 may perceptually be lost in favor of C_2 because the C_2 closure overlaps the C_1 closure gesture to such an extent that only the place of articulation of C_2 is perceived, although both gestures of C_1 and C_2 are fully intact (1992: 4). The cause for this asymmetrical behavior is seen in the perceptual weight accorded the C_2V onset formants as a position of dominance (1992: 21). She also found that place assimilation is more readily perceived when C_1 is a coronal than when C_1 is a labial or dorsal consonant (cf. also Byrd 1996: 237). This asymmetry is attributed to the perceptual dominance of dorsal and labial closures over a coronal closure, which is due to the greater velocity of the TT closure gesture (1992: 21). In concluding her paper, she claims that the gestural model of AP not only accounts for gradient assimilatory phenomena from the point of view of articulation but also from the perspective of perception.

In chapter 7.2 above, I presented experimental evidence from Ohala, who claims that the release burst of C_2 in nasal plus consonant (NC) sequences is the most relevant feature with respect to the perception of place cues (1990: 261). Together with Byrd's results, his argument can be extended to cases like <tha[k g]uy> and <thi[ʃ j]ear>: C_1 is perceptually not as salient as C_2. The place feature of C_1 may be hidden by the place feature of C_2. The manner feature of C_1, which is *not* hidden, seems sufficient for the categorization process that leads to the encoded meaning. When compared to [s], the word-final [ʃ] does approach a dorsal place, but the segment does not acquire a dorsal specification; [ʃ] is still a coronal.[120]

(8.3) Place assimilation in <this year> using the place hierarchy (sbc0003.wav: 07'57")

|ðɪs jijr|	IO/MOA	IO/PLDOR	ASSIM/PL	IO/PLCOR
[ðɪs dijr] /ðɪs dijr/	*!	*		
[ðɪk jijr] /ðɪk jijr/	*!			*
[ðɪs jijr] /ðɪs jijr/			*!	
☞ [ðɪʃ jijr] /ðɪʃ jijr/				*

This situation is shown in figure 8.3 above; the MARKEDNESS constraint ASSIM/PL enforces an assimilation of place features while IO/PLCOR is violated by the winning candidate as the place of articulation of the coronal is retracted. ASSIM/PL thus causes a change of the articulatory place in the oral cavity, in much the same way as in the example of figure 8.2 above.

Browman & Goldstein point out that when two gestures are on the same articulatory tier, and specifically the TT (tongue tip) articulator, which is responsible for the production of coronals, these gestures cannot overlap "without perturbing each other's tract variable motions" (1989: 215). They analyze the example <ten themes>, whereby the target of the alveolar gesture is slightly advanced to produce a dental nasal (C_1) in anticipation of the dental gesture required for the fricative (C_2),

[120] It is always a somewhat arbitrary matter where to draw the border line between two places. In fact, data like these might count as evidence for palato-alveolars to be regarded as dorsals.

as an instance of "blending between the dynamical parameters of the two gestures rather than hiding" (1989: 217-8).

Since no dorsals and labials are involved in the alternation, the ranking used in the tableau of figure 8.3 above, with the MARKEDNESS constraint ASSIM/PL outranking the FAITHFULNESS constraint IO/PLCOR would select [tɛn̪ θijmz] as the optimal candidate. However, since it is again crucial that C_1 (coda) and not C_2 (onset) is targeted by the alternation, the use of different constraints seems preferable to account for the phenomena in question. Instead of the FAITHFULNESS constraints enforcing identity for place features in the segments involved, positional FAITHFULNESS constraints (cf. figure 7.9 in chapter 7 above) governing syllable onsets and codas can be shown to select the same candidate as optimal.

(8.4) Place assimilation in <ten themes> using positional FAITHFULNESS constraints

|tɛn θijmz|	IO/MOA	IO/PLONSET	ASSIM/PL	IO/POA
[tɛθ θijmz] /tɛθ θijmz/	*!			*
[tɛn sijmz] /tɛn sijmz/		*!		*
[tɛn θijmz] /tɛn θijmz/			*!	
☞ [tɛn̪ θijmz] /tɛn̪ θijmz/				*

The proposed ranking works equally well for the example <this year> presented in figure 8.3 above and for <ten themes> (cf. figure 8.4 above). Both tableaux (figure 8.3 above and figure 8.5 below), which show place assimilation targeting a fricative, are identical with respect to the candidates involved and the violations incurred.

(8.5) Place assimilation in <this year> using positional FAITHFULNESS constraints (sbc0003.wav: 07'57")

|ðɪs jijr|	IO/MOA	IO/PLONSET	ASSIM/PL	IO/POA
[ðɪs dijr] /ðɪs dijr/	*!	*		*
[ðɪk jijr] /ðɪk jijr/	*!			*
[ðɪs jijr] /ðɪs jijr/			*!	
☞ [ðɪʃ jijr] /ðɪʃ jijr/				*

In the tableau of figure 8.5 above, I substitute the the IDENTIO/PLACE constraint hierarchy IDENTIO/PLDOR >> … >> IDENTIO/PLCOR (cf. figure 8.3 above) for Beckman's (1999: 20) generalization referring to the structural position of a specific segment. A positional FAITHFULNESS constraint requiring identity of place features in onsets strictly dominates a FAITHFULNESS constraint demanding identity of place features in general: IDENTIO/PLONSET >> IDENTIO/POA (cf. Lombardi 2001b: 26 for a similar ranking proposal). If this hierarchy is substituted for the IDENTIO/PLACE hierarchy, i.e. IO/PLDOR is to replace IO/PLONSET and IO/PLCOR is to replace IO/POA, the optimal candidate selected is still the same, <thi[ʃ j]ear>.

With respect to place assimilation, it can be said that these two constraint subhierarchies somehow work hand in hand. They do seem to have a common goal, but they cannot be substituted for each other. This means that constraints referring solely to position and neglecting the proneness of coronals to become the targets of assimilations and, more generally even, of weakening alternations, do not achieve optimal results in accounting for the data from SAE. The two subhierarchies are possibly candidates for the controversial proposal of local constraint conjunction by Smolensky (1995), who claims that many languages rule out forms that violate two constraints simultaneously while forms that violate only one of the two may survive.

(8.6) The local conjunction of C_1 and C_2 in domain D (after Smolensky 1995: 4)

a. C_1&C_2 is violated when there is some domain of type D in which both C_1 and C_2 are violated.
b. Universally C_1&C_2 >> C_1,C_2

The proposal is controversial because it qualifies OT's basic assumption of strict domination. Padgett (2002), for example, shows that constraint conjunction is not required of the theory. He assumes—as I do—that universal constraint subhierarchies like *PL/DOR, *PL/LAB >> *PL/COR (Prince & Smolensky 1993: 181) are grounded in scales having some phonetic or psycholinguistic relevance (2002: 5). From this, he infers that restrictions on triggers of assimilation (cf. this section) and dissimilation, i.e. OCP effects (cf. section 8.3.2 below), can well be explained without relinquishing the assumption of strict domination (2002: 10-18).

Based on Itō (1989), a different proposal is made by McCarthy (1993). He formulates the constraint CODA-CONDITION to account for the fact that many languages restrict the types of consonants that may appear in syllable codas. This constraint simply states that a coda cannot license place features. For SAE, however, this formulation is obviously too strong, since it would produce *[hæntowst] for <ham toast> or *[hæŋkænərɪ] for <ham cannery>, which are clearly impossible assimilations (cf. chapter 7.2 above).

In the preceding chapters, I have presented abundant evidence for the claim that coronals have quite asymmetrical properties, both in SAE and crosslinguistically. On the one hand, CORONAL is the most frequent place feature crosslinguistically (cf. chapter 4.3.2 above); coronals are 'plain' and 'basic' segments for both perceptual processes (cf. chapter 6.2 above) *and* articulatory processes (cf. chapter 3.2 above); and coronal segments are frequently found in positions where labials and dorsals cannot appear (cf. chapter 6.3 above). On the other hand, they are frequently the target of weakening alternations, such as assimilation, which is discussed in this section.

In OT terms, this asymmetry finds its clearest expression in the diametrically opposed rankings of one set of MARKEDNESS constraints and one set of FAITHFULNESS constraints as shown in figure 8.7 below.

(8.7) Diametrically opposed rankings of two sets of constraints

IDENTIO/PLDOR >> IDENTIO/PLLAB >> IDENTIO/PLCOR (FAITHFULNESS)
*GESTURE/DOR >> *GESTURE/LAB >> *GESTURE/COR (MARKEDNESS)

Boersma offers an interesting explanation for the asymmetries displayed by coronals (1998: 180-4). He estimates an average language to have three times as many coronal segments as labial or dorsal segments. In SAE, this estimated figure is roughly accurate for both the inventory frequency and the occurrence frequency of coronals (Wang & Crawford 1960: 136; Denes 1963: 894), that is, a coronal gesture is used about three times as often as a labial or dorsal gesture in this variety. Elaborating on the probabilistic recognition strategies of the hearer, Boersma further assumes that in each 100 articulations of segments that differ with respect to the three articulatory places LABIAL, CORONAL, and DORSAL, about the same number of misperceptions will occur. However, since coronals are much more frequent than labials and dorsals, the probability that a labial or dorsal gesture is perceived as CORONAL is approximately three times as high as the probability that a coronal gesture is perceived as LABIAL or DORSAL. Boersma demonstrates his probabilistic calculation with a direct comparison between CORONAL and LABIAL places.

(8.8) Numbers of occurrences of produced and perceived place values calculated from 100 place specifications (after Boersma 1998: 182)

prod↓ perc→	LAB	COR	total produced
LAB	21.8	3.2	25
COR	2.3	72.7	75
total perceived	24.1	75.9	100

With respect to the assimilation of place shown in figure 8.2 above, this means that for a specification |tɛn pɜrsɛnt| an assimilated form [tɛm pɜrsɛnt] is three times as likely to be perceived and categorized as |tɛn pɜrsɛnt| as a form [tɛn tɜrsɛnt] because the hearer will readily reconstruct a coronal when a labial is pronounced but will not quite as readily reconstruct a labial when a coronal is pronounced. Thus, due to their exceptionally high frequency, coronals can be reconstructed even if labials or dorsals are produced. A further important argument for [tɛm pɜrsɛnt] being favored over [tɛn tɜrsɛnt] is of course the positions of the relevant segments in coda and onset. Positional privilege and perceptual advantage are awarded to the dominant onset position which commonly acts as trigger and not as target for assimilation while the segment(s) in the nondominant coda position are commonly the target(s) of assimilation (Beckman 1999: 3; cf. also Ohala's (1990: 263) example discussed in chapter 7.2 above, where [ampa] was heard for /anpa/ and not [anta]).

8.2.1.2 Laryngeal assimilation

It is frequently pointed out in the literature that the default laryngeal feature value for obstruents is [–voiced] if a binary feature is assumed (e.g. Stampe 1979), or alternatively that obstruents do not have a laryngeal specification by default if a privative feature VOICE is assumed (Lombardi 1999: 270). Crosslinguistic evidence shows that in cases where one finds consonant clusters that differ in laryngeal specification, i.e. one consonant without laryngeal features and one marked for [voice], usually the result of laryngeal assimilation is a cluster unmarked for [voice] and not the other way around. Likewise, it has been observed that in languages with a voicing contrast in obstruents, neutralization may affect voiced obstruents (e.g. in German; Trubetzkoy 1939) in nonprivileged positions turning them into voiceless obstruents (Beckman 1999: 3).

Following Prince & Smolensky (1993: 178-85), Lombardi (1999: 271) postulates a MARKEDNESS constraint *LAR for obstruents: "do not have laryngeal features", equivalent to Hayes' MARKEDNESS constraint OBS/VOI: "an obstruent must be voiceless" (1999: 14), which I introduced in chapter 6.2.1 above. MARKEDNESS constraints like these are phonetically grounded in that they pertain to the needs of articulation. In the inventory of SAE, obstruents come in pairs, all voiceless ones have voiced counterparts. However, for an obstruent to be voiced, one extra laryngeal gesture is needed, that is, voiceless obstruents are simpler in articulatory terms than voiced ones.

Consequently, the devoicing of a voiced obstruent in connected speech, as for instance in <five third-graders> [faif θɜrd], is both a common and expected phenomenon in English generally (Spencer 1996: 105) and more specifically in SAE as well (cf. figure 8.9 below for examples taken from the *SBCSAE*).

(8.9) Examples of voiced obstruents assimilating to the voicelessness of a following obstruent in SAE (from *SBCSAE*)

a.	dog keeps	/gk/ → [kk]	(sbc0007.wav: 08'58")
b.	five car	/vk/ → [fk]	(sbc0007.wav: 06'54")
c.	five third-graders	/vθ/ → [fθ]	(sbc0004.wav: 15'55")
d.	five hundred	/vh/ → [fh]	(sbc0014.wav: 19'41")
e.	these people	/zp/ → [sp]	(sbc0014.wav: 13'32")
f.	these tendons	/zt/ → [st]	(sbc0001.wav: 02'12")
g.	these classes	/zk/ → [sk]	(sbc0002.wav: 00'26")
h.	these figures	/zf/ → [sf]	(sbc0002.wav: 10'04")
i.	these shoes	/zʃ/ → [sʃ]	(sbc0001.wav: 04'09")

Looking at the table in figure 8.9 above, the situation is very similar to the one of place assimilation depicted in figures 8.2 and 8.3 above. Describing the sequence of target and trigger segments, all these instances can be called cases of anticipatory assimilation across either word or morpheme boundaries. Although it is indeed a crosslinguistic fact that anticipatory assimilation is much more frequent than preservative assimilation (e.g. Cho 1990; Kiparsky 1995; Lombardi 1999), the mere statement of a directional asymmetry is in no way an explanatory notion. In search of a deeper level of analysis, however, I pointed out in the preceding discussion that faithfulness or nonfaithfulness to segments in positions of privilege or nonprivilege plays one of *the* crucial roles for alternations in general and assimilation in particular.

(8.10) Anticipatory laryngeal assimilation in <five third> (sbc0004.wav: 15'55")

|faiv θɜrd|	IO/LARONSET	OBS/VOI	ASSIM/LAR	IO/LAR
[faif ðɜrd] /faif ðɜrd/	*!	*	*	*
[faiv ðɜrd] /faiv ðɜrd/	*!	*		*
[faiv θɜrd] /faiv θɜrd/		*!	*	
☞ [faif θɜrd] /faif θɜrd/				*

In the tableau shown in figure 8.10 above, a specific FAITHFULNESS constraint, demanding input *onsets* to surface faithfully with respect to the feature [voice], must accordingly be ranked high above a general FAITHFULNESS constraint requiring identity between input and output laryngeal specifications. Intervening between the two FAITHFULNESS constraints in the hierarchy are two MARKEDNESS constraints motivated by the needs of articulation, one pressuring obstruents to be voiceless, the other demanding obstruent clusters to agree in voicing.

The same constraint ranking that effects these assimilations may in fact also be responsible for alternations in derivatives like <five> vs. <fifth>, <wide> vs. <width> and irregularly inflected forms like <leave> vs. <left>, <lose> vs. <lost>. All of these base forms originally contain a voiced obstruent and the complex forms surface in an all-voiceless word-final obstruent cluster, i.e. <five> + <-th> yields [fɪfθ]. In these cases, where a single syllable is affected, the IO/LARONSET constraint is assumed to be present in the hierarchy without having any effects.

So far, the generalization that I proposed for place assimilation can well be extended to laryngeal assimilation, namely that C_1, which is a coda constituent, is perceptually not as relevant as C_2, which is located in a position of prominence, i.e. the onset. The manner feature of C_1 (/f/ is still a fricative) seems sufficient for the categorization process leading to the encoded meaning. Manner features thus appear to be perceptually more salient than laryngeal *and* place features. This implies that hearers of casual or fast speech seem to rely less on laryngeal and place information than on manner features.

The regular English plural and past tense formations are further instances of laryngeal assimilation. Here, onset and coda position cannot be contrasted either because the formation of the plural and the past tense takes place entirely in syllable codas. It is commonly assumed that the plural suffix is either unspecified for voicing (Kenstowicz 1994: 150) or that its underlying form is /z/ (Lombardi 1999: 288). Both proposals account for the occurrence of the three different plural allomorphs. If the root ends in a voiced sound segment, i.e. a voiced consonant or a vowel, the plural suffix retains the feature [voice], e.g. [dɑgz], [kawz]. If the stem ends in a voiceless consonant, the plural suffix assimilates to voicelessness, e.g. [pʌps], and finally, if the root ends in a sibilant, the suffix surfaces as [ɪz], thereby changing the syllable structure, e.g. [hɔr.sɪz][121].

In an OT analysis of the English plural formation, Lombardi invokes a constraint similar to the SSG formalized as SONORITY (cf. chapter 6.3.1 above), which she calls HARMS' GENERALIZATION[122], requiring voiced obstruents to be closer to the syllable peak than voiceless obstruents (1999: 288), thereby ruling out the incorrect form *[pʌpz].

(8.11) Plural formation in <pups> (after Lombardi 1999: 289)

|pʌp + z|	HARMS'GENL.	IO/LAR	*LAR
[pʌpz] /pʌpz/	*!		*
[pʌbz] /pʌbz/		*	*!*[123]
☞ [pʌps] /pʌps/		*	

This analysis, however, comes to the undesirable result that [pʌps] and *[pʌbz] are almost equally good candidates. Only with respect to the lowest ranked constraint *LAR does *[pʌbz] get two—according to Lombardi *fatal*—violation marks, whereas [pʌps] does not violate *LAR at all. However, the results presented by Lombardi are

[121] Focusing on syllable structure, the voicing of the plural allomorph in the second syllable in <horses> [.sɪz] is caused by the preceding vowel just as in [kawz]; the vowel [ɪ] is epenthesized for perceptual reasons in order to separate the two sibilants.

[122] The generalization stems from a paper by Robert T. Harms (1973) titled "Some Nonrules of English" and distributed by the Indiana University Linguistics Club at Bloomington.

[123] Lombardi (1999) has a double violation mark for cases like these where the constraint *LAR is violated by two obstruents.

not as straightforward as they should be. Since the higher ranked IO/LAR is violated by both candidates in equal fashion, the violation of the lower ranked constraint *LAR should be of no consequence as constraints in OT are ranked with respect to one another by way of strict domination (cf. chapter 4.3.4 above). An analysis referring to the exceptional properties of coronals, using the constraint hierarchy IDENTIO/LARDOR >> IDENTIO/LARLAB >> IDENTIO/LARCOR (introduced in chapter 7.3.3 above) instead of the general FAITHFULNESS constraint IO/LAR, achieves better results.

(8.12) Plural formation in <pups> using a constraint hierarchy focusing on PLACE

|pʌp + z|	IO/LARDOR	IO/LARLAB	ASSIM/LAR	IO/LARCOR
[pʌbz] /pʌbz/		*!		
[pʌpz] /pʌpz/			*!	*
☞ [pʌps] /pʌps/				*

The winning candidate only violates the lowest ranked FAITHFULNESS constraint IDENTIO/LARCOR which requires laryngeal input and output specifications of coronals to be identical. Considering the general instability of coronals for which I have presented evidence in the preceding chapters, it comes as no surprise that coronal segments are subjected to laryngeal assimilation and labials or dorsals are not. The other two candidates clearly violate higher ranked constraints. One violates the higher ranked FAITHFULNESS constraint IDENTIO/LARLAB, the other violates the intervening MARKEDNESS constraint ASSIM/LAR demanding obstruent clusters to assimilate in voicing.

(8.13) Plural formation in <dogs>

|dɑg + z|	IO/LARDOR	IO/LARLAB	ASSIM/LAR	IO/LARCOR
☞ [dɑgz] /dɑgz/				
[dɑgs] /dɑgs/			*!	*
[dɑks] /dɑks/	*!			*

The generalization that obstruents are voiceless by default and show a tendency to become voiceless when specified for VOICE does hold in certain SAE contexts (e.g. in the obstruent clusters presented in figure 8.9 above, which might also be analyzed as instances of neutralization of contrast). Considering the circumstances of the plural formation, which seems to be a genuine example of laryngeal assimilation, it becomes apparent, however, that this generalization can only be upheld for cases where there is a voiceless trigger, e.g. <pups>; in an all-voiced environment voicedness can dominate over voicelessness.

The distribution of the past tense allomorphs (as well as the past participle allomorphs) can be stated in parallel terms as the generalizations concerning the plural formation in English. Three different allomorphs are distinguished, namely /d/, /t/, and /ɪd/, the first of which is assumed to be the underlying form. If the stem ends in a voiced sound segment, i.e. a voiced consonant or a vowel, the past tense suffix retains the feature [voice], e.g. <starred> [stɑrd], <stayed> [stejd]. If the stem ends in a voiceless consonant, the past tense suffix assimilates to voicelessness, e.g. <stopped> [stɑpt], and finally, if the stem ends in a coronal plosive, e.g. <started>, the suffix surfaces as [ɪd], thereby changing the syllable structure, i.e. [stɑr.dɪd]. The tableau in figure 8.14 below uses the same constraint hierarchy as the one for plural formation shown in figure 8.13 above.

(8.14) Past tense formation in <stopped>

|stɑp + d|	IO/LARDOR	IO/LARLAB	ASSIM/LAR	IO/LARCOR
[stɑbd] /stɑbd/		*!		
[stɑpd] /stɑpd/			*!	*
☞ [stɑpt] /stɑpt/				*

In SAE, a well-known and well-investigated case of laryngeal assimilation is /t/-voicing: the coronal plosive /t/ becomes voiced if surrounded by voiced segments and if it is part of an unstressed syllable (Trager 1942: 147; Lehmann 1953: 271-5; Sharf 1960: 105-9; Kenyon (1994: 315-20), includes spectrographic data comparing pairs of lexemes like <writer> and <rider>). Another term for this alternation is flapping of /t/ *and* /d/ in the environment mentioned above (Kahn 1976: 56-61). Authors who use this term commonly point to the fact that in this context, contrast between the two coronal plosives is neutralized to [ɾ] (cf. figures 8.15 and 8.16 below). I mentioned in chapter 7.2 above that articulatory complexity as a means of distinguishing marked and unmarked categories is important in connection with one of Greenberg's (1966: 58) criteria for markedness, i.e. the unmarked element is the element that contrasts *neutralize* to. Neutralization and assimilation both are characterized by a certain degree of loss of contrast. However, while neutralization by definition has as its result a less complex and thus unmarked segment, assimilation frequently gives rise to more complex and thus marked segments (e.g. place assimilation discussed in section 8.2.1.1 above). A voiced plosive is more complex than its voiceless counterpart with respect to an extra laryngeal gesture, so it remains questionable whether the terms neutralization or even lenition as a description for this type of alternation are really suitable. It does not matter whether this sound is called a voiced /t/ or a flap, it is a *voiced* segment (Oswald 1943: 24-5). Thus, the laryngeal feature VOICE is assimilated in cases like <kettle> (cf. figure 8.15 below), which are specified for the voiceless coronal plosive /t/ in UR.

In a study of 54.000 American English stop consonants taken from the TIMIT corpus[124], Byrd (1993) investigated 1557 sentences containing two potential flap sites, the word <water> and the phrase <suit in>. She found a significant effect with respect to the position of the flap; 99% of the speakers had a flap word-internally in <water> while only 19% had one at the word-final site <suit in> (Byrd 1993: 109). The reason for these very distinct results may be the fact that the syllable structure in the second example must be altered in order for the alternation to take place, i.e. [suw.ɾɪn].

(8.15) Examples of /t/-flapping in SAE (from *SBCSAE*)[125]

a.	international	SON[D]V	(sbc0012.wav: 12'30")
b.	kettle	V[D]SON	(sbc0012.wav: 12'15")
c.	dirty	SON[D]V	(sbc0006.wav: 14'33")
d.	political	V[D]V	(sbc0012.wav: 11'58")
e.	categorizing	V[D]V	(sbc0012.wav: 12'53")

[124] The TIMIT corpus of read speech has been designed to provide data for the acquisition of acoustic-phonetic knowledge and for the development and evaluation of automatic speech recognition systems. TIMIT contains a total of 6300 sentences read by 630 speakers, i.e. 10 sentences per speaker. The speech was recorded at Texas Instruments (TI) and transcribed at the Massachusetts Institute of Technology (MIT) = TIMIT.

[125] Apart from the common IPA symbol [ɾ], the flap is also represented as [D] following Trubetzkoy's (1939) archiphoneme notation.

(8.16) Homophonous examples of /d/-flapping in SAE (from *SBCSAE*)

a.	tendency	SON[D]V	(sbc0012.wav: 13'05")
b.	cradle	V[D]SON	(sbc0012.wav: 12'41")

The flap becomes the onset of the second (unstressed) syllable (cf. tableaux in figures 8.21 and 8.22 below). For the SAE data shown in figure 8.15 above, this means that one of the notorious markedness statements, namely the fact that sonorants are redundantly voiced (cf. discussion in chapter 4.3.4 above), turns up as a high-ranking MARKEDNESS constraint in an OT analysis of the data. SON/VOI ('a sonorant must be voiced') ranked above OBS/VOI guarantees that voiceless sonorants will not surface in English (Itō, Mester & Padgett 1995: 581) even if the MARKEDNESS constraint ASSIM/LAR is present in the hierarchy to require an assimilation of laryngeal features. The target of the voicing alternation is the voiceless coronal plosive /t/ if located between two vowels or a vowel and a sonorant. It is in this environment that /t/ assimilates the laryngeal feature and surfaces as a voiced coronal obstruent—the so-called flap, which constitutes a single rapid contact between the tongue tip and the alveolar ridge.

(8.17) SAE laryngeal assimilation in <dirty> (sbc0006.wav: 14'33")

|dɜrtɪ|	SON/VOI	IO/LARDOR	IO/LARLAB	ASSIM/LAR	IO/LARCOR
[dɜr̥tɪ] /dɜr̥tɪ/	*!				
[dɜrtɪ] /dɜrtɪ/				*!	
☞ [dɜrɾɪ] /dɜrɾɪ/					*

A high-ranking SON/VOI thus prevents an assimilated *voiceless* cluster as in [dɜr̥tɪ], which satisfies ASSIM/LAR just as well as the winning candidate [dɜrɾɪ] does (cf. tableau in figure 8.17 above).

Specifically with respect to nasals—which are also sonorants—when followed by an obstruent, there is crosslinguistic evidence for a MARKEDNESS constraint of the form *NC$_{\text{VOICELESS}}$ banning nasal plus voiceless obstruent sequences, a constraint which is said to be phonetically grounded. During the production of the sequence nasal plus obstruent, there is air still flowing out of the nose which in turn facilitates the voicing of the following obstruent (Kager 1999: 61). At first sight, the data taken from the *SBCSAE* are in accordance with this claim; <in[ɾ]ernational> (8.15 a.) is an example of obstruent voicing following a nasal. However, when considering the other examples of figure 8.15 above, it is immediately evident that not only nasals, but all sonorants (including vowels) trigger voicing and that not all obstruents are affected, but only the coronal obstruent /t/. Again, the ranking with the MARKEDNESS constraint ASSIM/LAR intervening between the two FAITHFULNESS constraints accounts for laryngeal assimilation in <dirty> *and* for the nonoccurrence of laryngeal assimilation in cases like <murky> *[mɜrgɪ] and <chirpy> *[ʧɜrbɪ].

However, another problem arises when comparing the pair <atom> with a flapped /t/ and <atomic>, where no flapping occurs; the constraint ranking proposed in figure 8.17 above does not account for the unflapped /t/ in <atomic>. The conditions for flapping—as set up by Kahn (1976: 56-61)—are the following: the sequence vowel or sonorant + /t/ + vowel or sonorant is prohibited, if and only if the coronal plosive belongs to an unstressed syllable, i.e. /t/ may not be aspirated. Consequently, the constraint hierarchy must be modified to include a positional FAITHFULNESS constraint enforcing identity for features (F) of segments located in stressed syllables over features of segments located in unstressed syllables (Beckman 1999: 7-8), that is IDENTIO/FSÝL >> IDENTIO/FSỲL (cf. chapter 7.3.3 above).

(8.18) SAE laryngeal assimilation in <atom>

<table>
<tr><th>|æ.təm|</th><th>IO/LARSÝL</th><th>SON/VOI</th><th>IO/LARDOR[126]</th><th>ASSIM/LAR</th><th>IO/LARCOR</th></tr>
<tr><td>[æ.təm] /æ.təm/</td><td></td><td></td><td></td><td>*!</td><td></td></tr>
<tr><td>☞ [æ.ɾəm] /æ.ɾəm/</td><td></td><td></td><td></td><td></td><td>*</td></tr>
</table>

In the tableau of figure 8.18 above the high-ranking constraint IDENTIO/FSÝL remains inactive since the target segment is the onset of the unstressed syllable. In figure 8.19 below, however, this constraint prevents flapping from taking place in a stressed syllable.

[126] In the tableaux shown in figures 8.18 through 8.20 above, the constraint IO/LARLAB in the hierarchy IDENTIO/LARDOR >> IDENTIO/LARLAB >> IDENTIO/LARCOR is dropped merely for reasons of space since it plays no role in the alternations.

(8.19) High-ranking positional FAITHFULNESS constraint IO/LARSÝL preventing assimilation in an SAE pronunciation of <atomic>

|ə.tɑ.mɪk|	IO/LARSÝL	SON/VOI	IO/LARDOR	ASSIM/LAR	IO/LARCOR
☞ [ə.tɑ.mɪk] /ə.tɑ.mɪk/				*	
[ə.ɾɑ.mɪk] /ə.ɾɑ.mɪk/	*!				*

Alternatively, an OT analysis is conceivable which uses another high-ranking FAITHFULNESS constraint IDENTIO/FASP. This constraint would enforce identity for laryngeal features in an aspirated segment, as is the case in <atomic>.

(8.20) High-ranking IO/LARASP preventing assimilation in an SAE pronunciation of <atomic>

|ət^{h}ɑmɪk|	IO/LARASP	SON/VOI	IO/LARDOR	ASSIM/LAR	IO/LARCOR
☞ [ət^{h}ɑmɪk] /ət^{h}ɑmɪk/				*	
[əɾɑmɪk] /əɾɑmɪk/	*!				*

Both constraints—IO/LARASP and IO/LARSÝL—yield the same result. In the light of the preceding and ensuing discussion, however, I suggest that positional FAITHFULNESS constraints, which in a very general way refer to positions of privilege, are much better suited for capturing generalizations within phonological systems than one specific constraint referring to the aspiration of segments.

For an analysis of Byrd's example <suit in> (1993: 109), IO/LARSÝL is clearly superior to IO/LARASP. Assimilation is not possible in cases where the coronal belongs to the stressed syllable, while assimilation is allowed in cases where the coronal is part of the following unstressed syllable.

(8.21) High-ranking positional FAITHFULNESS constraint IO/LARSÝL preventing assimilation in an SAE pronunciation of <suit in> (example from Byrd 1993: 109)

|suwt.ɪn|	IO/LARSÝL	SON/VOI	IO/LARDOR	ASSIM/LAR	IO/LARCOR
☞[suwt.ɪn] /suwt.ɪn/				*	
[suwɾ.ɪn] /suwɾ.ɪn/	*!				*

Within OT, resyllabification as shown in figure 8.22 below is actually the domain of ALIGNMENT constraints that evaluate the coincidence of edges of morphological and phonological constituents (McCarthy & Prince 1993). Since my investigation is mainly concerned with the segmental dimension, in particular with coronals, I will not elaborate on the theoretical implications of alignment.

(8.22) High-ranking positional Faithfulness constraint IO/LarSýl permitting assimilation in an SAE pronunciation of <suit in> (example from Byrd 1993: 109)

|suw.tɪn|	IO/LarSýl	Son/Voi	IO/LarDor	Assim/Lar	IO/LarCor
☞[suw.ɾɪn] /suw.ɾɪn/					*
[suw.tɪn] /suw.tɪn/				*!	

In both cases (8.17 and 8.21), if the coronal plosive is part of an unstressed syllable, which is a nondominant position (cf. figure 7.9 in chapter 7 above), the voicing of the coronal stop is possible in SAE. The manner features [–continuant, +obstruent], which, it is assumed, guarantee perceptibility, are retained in the assimilated segment. Laryngeal assimilation can be represented as a Markedness constraint Assim/Lar dominating a Faithfulness constraint IO/Lar (figure 8.10 above), which in most cases is decomposed into the hierarchy IdentIO/LarDor >> IdentIO/LarLab >> IdentIO/LarCor (figures 8.12-14 and 8.17-22 above). The exceptional behavior of the coronal segments becomes evident in both laryngeal assimilation and place assimilation. In these cases, the Markedness constraint Assimilate(F) is ranked exactly between the Faithfulness constraints IdentIO/LarLab or IdentIO/PlLab and IdentIO/LarCor or IdentIO/PlCor.

Laryngeal assimilation is subject to a similar conflict between perceptual and articulatory goals as the coronal stops assimilating to the place of articulation of a following noncoronal consonant in SAE, which was discussed in the previous section. Evaluating a voiceless and a voiced plosive in isolation, goal 2.b (cf. figure 7.7 in chapter 7 above) will favor /t/ because it is simpler in articulatory terms. For the production of /t/ the vocal cords are inactive while they must vibrate for the production of /d/. However, if the coronal is located in an all-voiced environment *and* is part of an unstressed syllable as in <dirty>, goal 2.a outranks goal 2.b. A continuous vibration of the vocal cords across three segments reduces articulatory effort in comparison to a sequence of sounds where voicing is present in the first segment, after which there is an offset of voice for the production of /t/, and finally a voice onset again for the third segment. Thus when there is a choice between /t/ and /d/ in isolation, /t/ amounts to less effort for the speaker than /d/, but intervocalically

/d/ minimizes articulatory effort (Mohanan 1993: 99) *without* having the effect of perceptual confusion for the hearer. Articulatory simplification, i.e. the prolonged voicing throughout a cluster consisting of [+son] [–cont, +obs, +cor] [+son], can take place without running the risk of perceptual ambiguity.

On the basis of the SAE assimilation data and my analysis of these data, I propose an OT grammar fragment, which is of the following shape:

(8.23) OT grammar fragment for SAE assimilation alternations consisting of ranked constraints

IO/MOA	FAITHFULNESS
IO/LARSÝL	FAITHFULNESS
SON/VOI	MARKEDNESS
IO/PLONSET or IO/LARONSET	FAITHFULNESS
IO/PLDOR or IO/LARDOR	FAITHFULNESS
IO/PLLAB or IO/LARLAB	FAITHFULNESS
OBS/VOI	MARKEDNESS
ASSIM/PL or ASSIM/LAR	MARKEDNESS
IO/PLCOR or IO/LARCOR or IO/PLACE	FAITHFULNESS
*GESTURE/DOR	MARKEDNESS
*GESTURE/LAB	MARKEDNESS
*GESTURE/COR	MARKEDNESS

In summarizing the preceding two sections, I would like to stress the fact that assimilatory phenomena in SAE are characterized by goal 2 (minimize articulatory effort) generally affecting place or laryngeal features in coronals while goal 1 (maximize perceptual contrast) generally is concerned with manner features which therefore remain unchanged, i.e. manner features are commonly not subject to assimilation. I have shown that the two conflicting goals can be represented as OT constraints making contradictory requirements. In SAE assimilation alternations, both goals are generally satisfied in that articulatory effort is reduced without the loss of perceptual contrast.

8.2.2 Deletion

Within Correspondence Theory, deletion is analyzed as a correspondence relation $\mathscr{R}$ between input and output strings S_1 and S_2 (McCarthy & Prince 1995: 16; cf. also chapter 5.5.2 above). The identity of correspondent elements is assessed by a set of Faithfulness constraints, which belong to the Maximality–InputOutput family, and demand the underlying phonological string S_1 to be *maximally* present in the surface string S_2. In their function of stating that each segment of some specific input form must have a correspondent in its output form, Max–IO constraints militate against deletion. If, however, Max–IO is outranked by the Markedness constraint Delete(S), deletion of (a) segment(s) may occur. I mentioned in section 8.2.1 above that rankings are dependent on speech tempo and style, i.e. fast, slow, casual, formal, etc. and that rankings of specific Faithfulness constraints and specific Markedness constraints may in fact be reversed depending on the register involved. The SAE-specific ranking of constraints for each speech style then determines the optimal candidate.

The example shown in figure 8.24 below is a typical and frequently occurring case of consonant cluster simplification. In particular, the final coronal plosives /t,d/ are deleted when preceded and followed by consonant segments. Spencer claims that the deletion of the coronal plosives is not only a well-known feature of connected speech in English—in fact, in all of his examples of consonant deletion, only coronal plosives are dropped (cf. Spencer's tables; 1996: 224)—but is also extremely common crosslinguistically (1996: 64-6).

(8.24) Deletion of final coronal plosive /t/ in <perfect memory> in casual speech (example from Browman & Goldstein 1989: 215)

|pɜrfɛkt mɛmərɪ|	Max/Dor[127]	Max/Lab	Del/Cor	Max/Cor
[pɜrfɛt mɛmərɪ] /pɜrfɛt mɛmərɪ/	*!			
[pɜrfɛkt ɛmərɪ] /pɜrfɛkt ɛmərɪ/		*!		
[pɜrfɛkt mɛmərɪ] /pɜrfɛkt mɛmərɪ/			*!	
☞ [pɜrfɛk mɛmərɪ] /pɜrfɛk mɛmərɪ/				*

[127] The constraints Max-IO/Dor, Max–IO/Lab, and Max–IO/Cor are abbreviated for reasons of space.

This type of alternation is said to affect virtually all speakers of American English in various styles (Guy 1980: 2; Neu 1980: 53; Browman & Goldstein 1989: 215), i.e. speakers of SAE are included. The highest ranked constraints in a casual speech utterance of <perfect memory> are FAITHFULNESS constraints pertaining to the demands of perception, requiring input-output correspondents for dorsals and labials. The MARKEDNESS constraint DELETE/COR[128], which works in favor of the demands of articulation by eliminating coronal plosives wherever possible, intervenes between MAX–IO/LAB and MAX–IO/COR to make sure that only coronal stops are deleted and no other consonants.

Browman & Goldstein (1989: 215-7) have also addressed the phenomenon of deletion. Their analysis supplies phonetic evidence for the assumption that assimilation and deletion are both weakening alternations of a very similar kind. From an articulatory perspective, gestural scores vary with repect to an increase or a decrease in overlap between two articulatory gestures, the results of which are hidings and blendings of the relevant gestures (cf. section 8.2.1.1 above).

> [W]ith sufficient overlap, one gesture may completely obscure the other *acoustically*, rendering it *inaudible*.
>
> (Browman & Goldstein 1989: 215; emphasis added)

Clearly, Browman & Goldstein elaborate on the notion that a gesture may in fact be produced by the speaker but not perceived by the hearer. This is a circumstance that can well be accounted for by an OT grammar model, realizing the potential conflicts between the demands of speech perception and production (cf. figure 8.25 below). Browman & Goldstein analyze the example <perfect memory>, where the medial coronal, located between the dorsal plosive and the labial nasal, is deleted. They report that upon examination of the articulatory data of the speaker, they found that "nothing was actually deleted from a gestural viewpoint" (1989: 215). The coronal closure gesture was produced with much the same magnitude as when the two lexical items were produced in isolation. However, the coronal gesture was completely overlapped by the constrictions of the preceding dorsal gesture and the following labial gesture, all of which are on separate tiers, i.e. TB, TT, and L. In the midst of

[128] On the basis of the data from the *SBCSAE*, the MARKEDNESS constraint DELETE/COR is defined to apply only to coronal *plosives* in SAE. It is assumed to be part of the large WEAKENING constraint family, just like the various ASSIMILATE constraints (cf. chapter 7.3.3 above).

substantial overlap among these three gestures, the coronal "closure gesture was acoustically hidden" (ibid.).

(8.25) Modified analysis of Browman & Goldstein's (1989: 215) example <perfect memory> employing the constraint HIDE instead of DELETE

<table>
<tr><th>|pɜrfɛkt mɛmərɪ|</th><th>MAX/DOR</th><th>MAX/LAB</th><th>HIDE/COR</th><th>MAX/COR</th></tr>
<tr><td>[pɜrfɛkt mɛmərɪ] /pɜrfɛt mɛmərɪ/</td><td>*!</td><td></td><td></td><td></td></tr>
<tr><td>[pɜrfɛkt mɛmərɪ] /pɜrfɛkt ɛmərɪ/</td><td></td><td>*!</td><td></td><td></td></tr>
<tr><td>[pɜrfɛkt mɛmərɪ] /pɜrfɛkt mɛmərɪ/</td><td></td><td></td><td>*!</td><td></td></tr>
<tr><td>☞ [pɜrfɛkt mɛmərɪ] /pɜrfɛk mɛmərɪ/</td><td></td><td></td><td></td><td>*</td></tr>
</table>

A tableau adapting Browman & Goldstein's findings (cf. figure 8.25 above) presents an articulatory specification with all segments that are part of the perceptual specification (or UR) actually produced. With respect to the hearer's percept, however, the coronal gesture is hidden in the winning candidate and thus not perceived.

(8.26) Examples of deletions/hidings of medial coronal plosives located in CCC and CCCC clusters in SAE (from *SBCSAE*)

a.	hands	/ndz/ → [nz]	(sbc0001.wav: 11'42")
b.	hand me	/ndm/ → [nm]	(sbc0007.wav: 15'57")
c.	told me	/ldm/ → [lm]	(sbc0006.wav: 10'09")
d.	scared to	/rdt/ → [rɾ]	(sbc0001.wav: 02'51")
e.	handbook	/ndb/ → [nb]	(sbc0005.wav: 02'55")
f.	handwash	/ndw/ → [mw]	(sbc0013.wav: 15'51")
g.	learned that	/rndð/ → [rnð]	(sbc0011.wav: 03'16")
h.	expects	/kts/ → [ks]	(sbc0004.wav: 09'03")
i.	perfect weather	/ktw/ → [kw]	(sbc0003.wav: 08'18")
j.	most people	/stp/ → [sp]	(sbc0001.wav: 00'33")
k.	best time	/stt/ → [st]	(sbc0013.wav: 09'47")
l.	best friend	/stfr/ → [sfr]	(sbc0006.wav: 09'29")
m.	first couple	/rstk/ → [rsk]	(sbc0001.wav: 03'03")
n.	first moved	/rstm/ → [rsm]	(sbc0002.wav: 05'51")

The data in figure 8.26 above present an array of SAE examples collected from the *SBCSAE*. In clusters of three or more consonants, a medial coronal plosive (/t/ or /d/) is deleted (even if only perceptually) irrespective of the nature of the surrounding consonants, i.e. whether these are sonorants or obstruents. Thus if the cluster exclusively consists of coronals as in 8.26 a. and g., not the liquid, the nasal, or the fricative but the plosive is deleted. A phonetically grounded explanation is offered by Hume & Johnson. They claim that a plosive sandwiched between two or more other consonants is a context of very poor perceptibility (2001: 9). Consequently, deletion of the plosive not only serves to reduce articulatory effort (in cases where the closure gesture is indeed *not* produced), it also serves to enhance syntagmatic perceptual contrast by removing an element which is poorly audible and functionally not mandatory.

Deletion happens in cases where the cluster belongs to a single syllable (8.26 a. and h.) and where the cluster extends over syllable and root boundaries. Since mainly /t/ and /d/ are affected, it comes as no surprise that in the case of the past tense and past participle markers, even inflectional morphemes are deleted (cf. 8.26 d. and g.). In the given contexts (<I am scared to do …> and <I learned that …>) the inflectional elements are obviously not crucial for the categorization of these verb forms as being inflected for past tense.

Of further significance are the examples containing a cluster of four consonants in their input specification. Here, a medial consonant could either be position two or three in the cluster. At first sight, it seems odd why in 8.26 l. the second consonant is deleted and in 8.26 g., m., and n. it is the third consonants which are deleted. The reason for deletion cannot be position in a sequence alone since in each case the coronal is affected. On closer inspection, however, it is obvious that position must be part of the cause. The appropriate generalization appears to be that in a cluster of more three or more consonants *coronal* plosives in *coda position* tend to be deleted or, in Browman & Goldstein's terms, the TT gestures tend to be hidden.

Referring back to the discussion of positional faithfulness in section 8.2.1.1 above, an OT analysis making allowances for the nonprivileged position in which the alternation takes place is more desirable than the ones shown in figures 8.24 and 8.25 above.

(8.27) Deletion of coronal plosive /t/ in coda position in the casual speech example <perfect weather> (sbc0003.wav: 08'18")

|pɜrfɛkt wɛðər|	MAX-IO /ONSET	MAX-IO /DOR,LAB	DEL /COR[129]	MAX-IO /COR
[pɜrfɛkt ɛðər] /pɜrfɛkt ɛðər/	*!	*	*	
[pɜrfɛt wɛðər] /pɜrfɛt wɛðər/		*!	*	
[pɜrfɛkt wɛðər] /pɜrfɛkt wɛðər/			*!	
☞ [pɜrfɛk wɛðər] /pɜrfɛk wɛðər/				*

Example 8.26 d. <scared to> is interesting for a further reason as is 8.26 f. <handwash>. In these two examples, deletion seems to pave the way for a further weakening alternation, namely assimilation. If the medial plosives are deleted in the sequences <scared to> and <handwash>, the context is identical to the contexts discussed in sections 8.2.1.2 and 8.2.1.1 above on laryngeal and place assimilation. After the deletion of the coronal plosive in the coda of [skɛrd], an ideal environment for the flapping of /t/ arises. The coronal is part of an unstressed syllable and, after cluster simplification, is now surrounded by sonorants. It assimilates to the voiced environment and becomes a flap [skɛrɾə] in much the same way as the coronal stop does in the example [dɜrɾɪ] depicted in figure 8.17 above. Likewise <handwash> becomes [hæmwɔʃ]. After the medial stop is removed, /n/ is now the coda-final stop, and being a coronal stop, it is prone to assimilate the place feature (LABIAL) of the following consonant in onset position. I gave a list of examples of coronal stops assimilating to the place of articulation of a following noncoronal consonant in SAE in figure 7.4 in chapter 7.2 above.

In older, rule-based generative models such phenomena were dealt with under the headings of rule ordering and serial derivation (cf. discussion in chapter 4.1.1 above). If a rule A (deletion) produces as its output a form that serves as an input form or structural description for rule B (assimilation), then rule B will apply; the two rules are said to be organized in a feeding order (Kiparsky 1968). OT's grammar model is characterized by a total rejection of serial derivations by means of rules and rule orderings. The ranking of the relevant constraints alone is responsible for the output form. The computation of the optimal output candidate takes place across the

[129] For reasons of consistency and transparency I want to adhere to the MARKEDNESS constraint DELETE instead of MARKEDNESS constraint HIDE, since the alternation is commonly called deletion (cf. also assimilation and ASSIMILATE and epenthesis and EPENTHESIZE).

whole set of ranked constraints, i.e. DELETE and ASSIMILATE constraints assess the candidates all in one parallel operation, as proposed in figure 8.28 below.

(8.28) Deletion and place assimilation in <handwash> (sbc0013.wav: 15'51") as the interaction of FAITHFULNESS and MARKEDNESS constraints

<table>
<tr><th>|hændwɔʃ|</th><th>MAX-IO/ONS</th><th>IO/PLLAB</th><th>DEL/COR</th><th>ASSIM/PL</th><th>IO/PLCOR</th></tr>
<tr><td>[hændɔʃ] /hændɔʃ/</td><td>*!</td><td></td><td>*</td><td></td><td></td></tr>
<tr><td>[hædwɔʃ] /hædwɔʃ/</td><td></td><td></td><td>*!</td><td>*</td><td></td></tr>
<tr><td>[hænwɔʃ] /hænwɔʃ/</td><td></td><td></td><td></td><td>*!</td><td></td></tr>
<tr><td>☞[hæmwɔʃ] /hæmwɔʃ/</td><td></td><td></td><td></td><td></td><td>*</td></tr>
</table>

As mentioned above, the coronal plosive is preferred as a target for deletion over the nasal for reasons of perceptibility, while the coronal nasals are the class of consonants which are most susceptible to place assimilation for the same reasons. There are rare cases where other coronals (excepting the two plosives) are the targets of a deletion alternation.

(8.29) Examples of deletions of coronals other than /t/ and /d/ in SAE (from *SBCSAE*)

a.	sixth	/ksθ/ → [ks]	(sbc0004.wav: 09'44")
b.	months	/nθs/ → [ns]	(sbc0002.wav: 14'00")

These examples are different in two ways. First, it is not necessarily a medial consonant which is deleted and, second, the target is not a coronal plosive. In <sixth>, the target is the last of three consonants, the dental coronal fricative /θ/, and in <months>, the same segment is deleted in medial position. What these examples have in common is that the deleted segments are part of the nondominant coda position and that the segments violate the quantity constraints and the Sonority Sequencing Generalization for SAE codas (cf. discussion in chapter 6.3.2 above). Since the violating segments are exclusively coronals, the targets of deletion again are coronals.

The discussion shows that assimilation and deletion are similar with respect to yet another feature. Although the positional prerequisites are there, i.e. different consonants are located in the nondominant coda position, it is the coronals which are selected for deletion and not labial or dorsal segments. Whatever makes coronals undergo assimilation, makes them special with respect to deletion as well. Segments with nondominant, unmarked features are more prone to assimilation and deletion than those with marked features (Mohanan 1991: 319). To quote Browman & Goldstein again:

> [T]hree (superficially) different fluent speech processes—deletion of final alveolar stops in clusters, assimilation of final alveolar stops and nasals to following labials and velars, and assimilation of final alveolar stops to other tongue tip consonants—can all be accounted for as the consequence of increasing overlap between gestures in fluent speech. (1989: 219)

These phonetically grounded assumptions, which characterize the basic tenets of Articulatory Phonology, are reinforced by a recent investigation of coarticulation in connected speech by Farnetani (1997). Using a large number of experimental results, she also argues that there are no clear-cut distinctions between various connected speech processes. Rather, she favors the notion of a continuum, of which deletion would be the endpoint on a scale of weakening phenomena (1997: 399-403).

The physiological similarity of these weakening alternations is mirrored in the formal similarity of the constraint rankings required for the description of both types of alternation. The subhierarchy which accounts for deletion phenomena in SAE, i.e. MAX-IO/ONSET >> MAX-IO/DOR, LAB >> DEL/COR >> MAX-IO/COR, is highly reminiscent of the ranking proposed for assimilatory alternations shown in figure 8.23 above. A MARKEDNESS constraint which selects coronals intervenes between two FAITHFULNESS constraints, one higher-ranked constraint referring to dorsals and labials and one lower-ranked constraint referring to coronals.

In concluding this section on weakening alternations, I wish to stress that connected speech phenomena like assimilation and deletion involving coronals can be viewed as continuous and not categorical phenomena (Farnetani 1997). Continuous alternations are well-suited for a description under OT. I have shown, however, that an alternation like assimilation is not necessarily categorical from the side of the speaker but it must be categorical from the perspective of the hearer in order for communication to succeed.

8.3 MAXIMIZATION OF SEGMENTAL CONTRAST

Clearly, fortitions like epenthesis and dissimilation work in favor of the demands of speech perception by optimizing syntagmatic contrast. However, it is not quite clear whether they are aimed at reducing articulatory effort in their function of eliminating weak contrasts, or whether they in fact do just the opposite, i.e. increase articulatory effort. Either way, they are represented as MARKEDNESS constraints.

8.3.1 Epenthesis

Seen from a crosslinguistic perspective, one function of epenthesis is to enhance contrast by inserting (a) vowel(s) into clusters of consonants or to break up vowel clusters by inserting (a) consonant(s) into the cluster. Frequently, these operations are referred to as repair strategies (Paradis 1988; Archangeli & Pulleyblank 1994: 261-2) since they may change the syllable structure of the affected input forms in order to restore an unmarked CV structure. Examples from English are the plural formation, where the lax vowel /ɪ/ is epenthesized between two *coronal* sibilants, the past tense (and past participle) formation, where /ɪ/ is epenthesized between two *coronal* plosives (cf. section 8.2.1.2 above), or cases like [fɪ.ləm] for <film> in varieties of British English. The epenthesis of /ɪ/ between two coronals in the English plural and past tense formations enhances the perceptibility of the two surrounding sibilant and plosive consonants. Optimizing contrast in this environment is claimed to be all the more important because one of the coronals has the function of an inflectional morpheme (Hume & Johnson 2001: 8).

In Japanese, for example, the default epenthetic vowel is /ɯ/, a lax unrounded high back vowel. Since Japanese with one exception[130] only tolerates CV(V), i.e. open syllables, an English word <bus>, for example, is incorporated into Japanese as [bɑ.sɯ]. The vowels inserted tend to be the unmarked lax vowels [ɪ,ə,ɯ,ʊ], which is the same class of vowels that are targets of deletion. This generalization is worth mentioning because it parallels my findings with respect to coronals in that the universally unmarked segments tend to take part in both weakening *and* strengthening alternations.

[130] The cxception is the coronal nasal /n/, which is the only consonant allowed in syllable codas, e.g. in forms like <Fuji-*san*>, which is the Japanese name for Mt. Fuji.

Epenthesis also works the other way around, by breaking up vowel clusters and supplying an onset for the second syllable, e.g. [sɔː.rən] in <I saw an ad>. Vowel epenthesis mentioned above is not part of this investigation for obvious reasons. Epenthesis of [r] does effect a coronal but is mainly a feature of varieties of British English[131]. Supposedly, the [r] sound as a voiced liquid placed between two vowels reduces articulatory effort compared to a glottal stop, which is usually found in its place.

This section is concerned with cases of epenthesis in the context of consonant clusters. Applying Browman & Goldstein's gestural model of Articulatory Phonology, epenthesis of /t/ in <dense> or <tense>, for example, is viewed as an adjustment of the timing of articulatory movements.

> [T]he occurrence of so-called epenthetic stops in English words like *tense* is dialect-dependent. […] such 'stops' are to be analysed in terms of variation in the relative timing of oral and velic gestures (rather than actual segment insertion; […]). In general, then, languages can differ from one another in the timing of (roughly the same) articulatory gestures.
> (Browman & Goldstein 1986: 221).

More specifically, epenthesis is seen as a premature lowering of the velum during the articulation of the nasal and before the release of the tongue tip closure gesture, which then has the perceptual effect of an intrusive stop (Bird & Klein 1990: 49).

Correspondence Theory, as a subtheory of OT, regards epenthesis as a specific correspondence relation $\mathscr{R}$ between input strings S_1 and output strings S_2 (McCarthy & Prince 1995: 16; cf. also chapter 5.5.2 above). Correspondent segments are assessed by the DEPENDENCY–INPUTOUTPUT family of FAITHFULNESS constraints which require that every segment of S_2 has a correspondent segment in S_1, in other words, S_2 is *dependent* on S_1. Contrary to the correspondence relation expressed by MAX–IO constraints, which prohibit deletion and demand every input segment to be present in the output, DEP–IO constraints militate against epenthesis by stating that all output segments must be present in the input. Only if DEP–IO is outranked by some MARKEDNESS constraint can epenthesis take place. These MARKEDNESS constraints either belong to the family of WEAKENING constraints (assimilation, deletion) or to the family of STRENGTHENING constraints (epenthesis, dissimilation).

[131] However, I did find *one* such example in the *SBCSAE*: <That was the ugliest set of shoes I ever saw in [sɔ.rɪn] my life> (sbc0006.wav: 07'58").

(8.30) Examples of epenthesis in coda consonant clusters in SAE (from *SBCSAE*)

a.	dense	/ns/ → [nts]	(sbc0001.wav: 19'53")
			(sbc0006.wav: 04'53")
b.	once	/ns/ → [nts]	(sbc0001.wav: 09'43")
			(sbc0011.wav: 00'23")
c.	dance	/ns/ → [nts]	(sbc0001.wav: 17'35")
			(sbc0002.wav: 17'46")
d.	ounce	/ns/ → [nts]	(sbc0011.wav: 06'19")
e.	chance	/ns/ → [nts]	(sbc0006.wav: 00'44")
			(sbc0008.wav: 12'04")
f.	month	/nθ/ → [ntθ]	(sbc0004.wav: 07'58")
			(sbc0006.wav: 25'12")
g.	seventh	/nθ/ → [ntθ]	(sbc0005.wav: 14'02")
h.	strength	/ŋθ/ → [ŋkθ]	(sbc0013.wav: 12'06")

It is commonly claimed that the effects of fortition phenomena, such as epenthesis, become salient in slower, more formal speech styles, while lenitions are more likely to operate in faster, more colloquial styles (cf. section 8.1 above). However, the selection of numerous examples shown in figure 8.30 above stems from the same corpus of casual, informal speech (cf. chapter 2.1 above for a description of the *SBCSAE*) that I have previously consulted for examples of deletion and assimilation. The SAE-specific ranking of constraints for this speech style determines the optimal candidate.

(8.31) Epenthesis of the coronal stop /t/ in <dense> (sbc0001.wav: 19'53")

|dɛns|	DEP/DOR	DEP/LAB	EPEN	DEP/COR
[dɛnks] /dɛnks/	*!			
[dɛnps] /dɛnps/		*!		
[dɛns] /dɛns/			*!	
☞ [dɛnts] /dɛnts/				*

In the example <dense>, the FAITHFULNESS constraints DEP/DOR and DEP/LAB, which prohibit epenthesis of dorsals and labials, dominate the MARKEDNESS constraint EPENTHESIZE, which works in favor of minimizing articulatory effort and enhancing perceptibility. EPENTHESIZE intervenes between DEP–IO/LAB and DEP–

IO/COR to make sure that only coronal stops are epenthesized and no others. Mohanan argues for coronals as default epenthetic elements on the basis of evidence from deletion and assimilation; the segments that undergo weakening alternations asymmetrically are also the segments that are inserted in epenthesis (1991: 318-9).

Considering the tableau of figure 8.31 above, a similar ranking of constraints can be held responsible for the insertion of the coronal /t/ as the one used for assimilation and deletion. The argument is based on the relative unmarkedness of coronals when compared to other segments, i.e. on Prince & Smolensky's subhierarchy *PL/DOR, *PL/LAB >> *PL/COR (1993: 181). If anything is epenthesized, it should be the least marked or default segment, that is, a *coronal* segment. Following Prince & Smolensky (1993), Beckman (1999) has a similar proposal for epenthesis (cf. figure 8.32 below). The high-ranking MARKEDNESS constraint EPENTHESIZE demands that a segment be inserted between the nasal and the fricative. The MARKEDNESS constraints representing the place hierarchy express the fact that *DORSAL, *LABIAL is ranked above *CORONAL, i.e. the most likely segment to be inserted is a coronal. The presence the low-ranking FAITHFULNESS constraint DEP-IO has in the hierarchy is of no consequence for the alternation.

(8.32) Epenthesis of /t/ in SAE <dense> (after Beckman 1999: 28)

|dɛns|	EPEN	*DORSAL[132]	*LABIAL[133]	*CORONAL	DEP-IO
☞ [dɛnts] /dɛnts/				*	*
[dɛnps] /dɛnps/			*!		*
[dɛnks] /dɛnks/		*!			*
[dɛns] /dɛns/	*!				

In her paper titled "Coronal Epenthesis and Markedness", Lombardi (2003) also claims that coronals are *relatively*[134] unmarked segments and that they should be favored crosslinguistically as epenthetic segments. She does concede, however, that

[132] Beckman's constraints *DORSAL, etc. and Smolensky's *PL/DOR, etc. have exactly the same function as *GESTURE/DOR, etc., which I proposed in figure 8.7 in section 8.2.1.1 above.

[133] A broken line in a tableau means that the two constraints separated by this line are not ranked with respect to each other, and their order might just as easily be interchanged.

[134] She says *relatively* unmarked since the truly unmarked place is PHARYNGEAL, making glottal stops the default epenthetic segments. I presented Lombardi's modified place hierarchy and constraint subhierarchy *PL/DOR, *PL/LAB >> *PL/COR >> * PL/PHA (2001b: 29) in chapter 4.3.5 above.

> there is one small empirical problem, which is that the existence of /t/ as a purely phonologically-driven epenthetic consonant is not particularly well supported by the evidence. (Lombardi 2003: 4)

The glottal stop, which surfaces much more frequently as epenthetic segment than any coronal, does not pose a real problem because its appearance is predicted by her modified constraint hierarchy *PL/DOR, *PL/LAB >> *PL/COR >> * PL/PHA (cf. note 16 at the bottom of the previous page). When a segment is epenthesized, it appears in the output for phonotactic reasons alone. That is, it lacks a correspondent segment in the input and thus cannot be subject to the evaluation of any FAITHFULNESS constraint. Since DEP–IO constraints have no influence on the nature of epenthetic segments whatsoever, their featural make-up is fully dependent on MARKEDNESS constraints. As a consequence, the least marked segment will surface as epenthetic segment as is indeed the case in <tense> where a coronal plosive /t/ is inserted by default.

A problem, however, becomes apparent in cases where marked segments, e.g. labials and dorsals, are inserted and this is a situation encountered in SAE, e.g. in <warm[p]th> and <streng[k]th> (sbc0013.wav: 12'06"). This means that the constraint ranking proposed in figure 8.31 above which decomposes DEP–IO/PLACE according to the place hierarchy cannot be upheld. Insertions of [p] and [k] must be explained otherwise since high-ranking DEP-IO/DOR would rule out the output form <streng[k]th>. With the exception of EPENTHESIZE and the MARKEDNESS constraints representing the place hierarchy, other MARKEDNESS constraints must be at work.

I claim that EPENTHESIZE is a STRENGTHENING constraint, i.e. a MARKEDNESS constraint that enhances perceptibility and possibly reduces articulatory effort. In the one-syllabic input form <month>, for example, the coda consists of two coronals, an alveolar nasal followed by a dental fricative. In several actual speech samples of <month> found in *SBCSAE*, a voiceless coronal stop is inserted between the nasal and the fricative (e.g. sbc0004.wav: 07'58", sbc0006.wav: 25'12"). It is hypothesized that the velic lowering during the articulation of the nasal described by Bird & Klein (1990: 49), which happens prior to the release of the (nasal) TT closure gesture and gives the impression of an epenthesized stop, in fact may reduce articulatory effort for the production of the following fricative, relative to an articulation without an 'inserted' stop. After a short blockage, the air can escape unhindered in a fricative-specific continuant fashion. If this is indeed the case, it would explain why, in cases where a labial or dorsal nasal is present in an input form, a lowering of the velum before the release of the nasal L or TB closure gestures will have the perceptual effect of a labial or dorsal intrusive plosive in cases like <warm[p]th> and <streng[k]th>. In section 8.2.2 above, I claimed that a plosive sandwiched between two other consonants is frequently deleted because this context is characterized by poor perceptibility. At first sight this is a plain contradiction. However, what I wish to point out is that in cases of epenthesis, the plosive is not inserted for its own sake (that is, for the sake of its own perceptibility), but for the sake of the perceptibility of the surrounding segments. This is why in many cases a relatively unmarked coronal segment will surface as epenthetic element while in some cases a more marked labial or dorsal segment is chosen as the optimal candidate.

(8.33) Epenthesis of [k] in <strength> (sbc0013.wav: 12'06") as interaction between Strengthening and Weakening constraints on the one hand and Faithfulness on the other hand

|strɛŋθ|	Epen	Assim/Pl	*Ges/Dor	*Ges/Lab	*Ges/Cor	Dep/IO
[strɛŋθ] /strɛŋθ/	*!					
[strɛŋps] /strɛŋps/		*!		*		*
[strɛŋtθ] /strɛŋtθ/		*!			*	*
☞ [strɛŋkθ] /strɛŋkθ/			*			*

Cases like <warm[p]th> and <streng[k]th> must be explained as interaction of Epenthesize with Assimilate constraints: given the choice between any epenthetic segment, goal 2.b in figure 7.7 in chapter 7 above will favor a coronal over any other consonant. If, however, a labial or dorsal segment is already present in the environment, i.e. /m/ or /ŋ/, 2.a will favor the sequences /mp/ or /ŋk/ respectively instead of /mt/ or /ŋt/, and /p/ or /k/ will surface as epenthetic segments.

To summarize this section, I claim that coronal unmarkedness is a valid concept, although it may not be quite as uncontroversial as frequently assumed. A further generaliztion to be deduced from the facts of epenthesis is that the target position in which the alternations take place is the nonprominent coda position, as is the case with deletion and assimilation phenomena. For this reason, an unviolated Faithfulness constraint Dep-IO/Onset, which is not included in the tableau merely for reasons of space, is assumed to dominate the highest-ranked constraint Epenthesize.

8.3.2 Dissimilation

Dissimilation has the effect of producing a difference between two or more segments with respect to some feature specification and thus serves to enhance syntagmatic perceptual contrast. Dissimilatory alternations are much less frequent than assimilatory phenomena, both crosslinguistically and in the SAE variety (Spencer 1996: 59; O'Grady et al. 2001: 48, 296). In the relatively small corpus that I consulted (*SBCSAE*), I did not find instances of dissimilation. The lack of data parallels the relative amount of research that has been undertaken; there are far fewer studies of dissimilation than studies of assimilation (Suzuki 1998: 10). Within GT, dissimilation was traditionally expressed in terms of feature-changing rules of the type $A \rightarrow [-F] / \_\_\_ [+F]$. In chapter 7.2 above, I showed that the older rule-based approaches encountered various types of difficulties with their analyses of assimilatory phenomena. The criticism mentioned above can be extended to dissimilation in much the same way. In the more recent autosegmental models of feature geometry, assimilation is characterized as the association (or 'spreading', or 'linking') of a feature or node (cf. chapter 7.2 above) while dissimilation is seen as an effect of the *delinking* of a feature or node from a given segment. A subsequent operation links the remains of the segment to the opposite feature value, which is assumed to be the default value (McCarthy 1988: 88; Yip 1988: 80-6). An explanation for the dissimilatory nature of the delinking operation is supplied by the Obligatory Contour Principle (OCP), which I discussed in chapter 6.3.1 above in connection with syllable onsets. The OCP states that adjacent identical elements on a given tier are banned from lexical representations (Goldsmith 1976; McCarthy 1988: 88; Yip 1988: 66). Possible consequences of the OCP are that URs which violate it are prohibited, that it will block alternations that create surface forms which violate the OCP, or that it will enable alternations which 'repair' violations of it (Yip 1988: 65). A principle with such a variety of effects clearly has severe drawbacks for a derivational approach to phonology. A constraint-based framework like OT, on the other hand, can handle various OCP effects provided that it is integrated into the theory as a violable constraint. Under such circumstances, problems will not arise since all possible output candidates are evaluated in parallel in search of the optimal candidate. Violability, however, must be an option because the OCP is not observed in each and every case. It has been shown that the OCP is *not* a candidate for a principle of UG (Odden 1995: 461-4). Odden summarizes his investigation of OCP effects: "Ultimately, languages retain the option of doing nothing about OCP violations." (1995: 464)

In a recent OT-based investigation of dissimilation, Suzuki (1998) lists various unresolved problems in connection with the OCP, foremost of which is the inability to account for the typological diversity of dissimilatory phenomena or for phenomena where identity avoidance is not observed (1998: 16-26). The main reason for this is that pre-OT principles in GT are thought to be inviolable. If the OCP is appropriated for use in OT, and this is what Suzuki proposes, it must function as a violable, rankable MARKEDNESS constraint that allows flexible interaction with other constraints, e.g. FAITHFULNESS constraints (1998: 47).

(8.34) Account of dissimilation using a typical ranking of OT constraints

MARKEDNESS OCP(F) >> FAITHFULNESS IDENT–IO(F)

Following Holton (1995: 170), the MARKEDNESS constraint OCP can be defined as a general constraint on identity avoidance of adjacent features on a given tier, which in principle can be violated—as *all* OT constraints can be. The OCP(F) constraint works on a language-particular basis, and on selected domains.

A typologically relatively common example of dissimilation is a cooccurrance restriction on adjacent coronals if they agree in their specification for the features [sonorant] or [continuant] (Yip 1989: 362-3). That is, the dissimilated segments are either two adjacant coronals specified [+son] or two adjacant coronals specified [+cont]. If adjacant coronal sonorants additionally agree in their specification for the feature [lateral], dissimilation is even more likely. A frequently cited example comes from Latin, in which the suffix <-alis> is subject to dissimilation when the stem contains a lateral, e.g. <sol-a*r*is> and not *<sol-alis> vs. <nav-a*l*is> (Kenstowicz 1994: 509; Spencer 1996: 71; cf. also <vela*r*> vs. <corona*l*>, which are derived from Latin).

An example of this type from English is <colonel>, pronounced as [kɜrnəl]. Wilson (1993) claims in his *Columbia Guide to Standard American English* that this lexical item is pronounced in the dissimilated fashion in all Standard American English varieties.

(8.35) Dissimilation of one of the two laterals in <colonel> (example from Wilson 1993)

|kɜlnəl|	OCP/LAT	IO/PLDOR	IO/PLLAB	IO/PLCOR	IO/LAT
[kɜlnəl] /kɜlnəl/	*!				
[kɜŋnəl] /kɜŋnəl/		*!			*
[kɜmnəl] /kɜmnəl/			*!		*
☞ [kɜrnəl] /kɜrnəl/				*	*

The highest ranked OCP constraint prohibits adjacent[135] [lateral] specifications while the IDENTITY constraints demand correspondent segments in input and output forms to have identical values for PLACE and for the feature [lateral]. The IDENTITY-IO/PLACE constraint is again decomposed into its individual places according to the place subhierarchy in order to illustrate once more the exceptional behavior of coronals. The dissimilation affects an input coronal /l/, and the optimal output form contains a dissimilated coronal [r].

The question as to whether the OCP has perceptual or articulatory correlates is discussed in Boersma (1998: 415-40). Epenthesis of /ɪ/ in the English plural and past tense formation, which constitutes a violation of FAITHFULNESS—more precisely of the DEP-IO constraint family—can be taken as an example of an OCP effect. Either two coronal sibilants or two coronal plosives would be adjacent in plural forms like *[hɔrss] or in past tense forms like *[stɑrdd] (cf. section 8.2.1.2 above). In section 8.3.1 above, I presented evidence for this type of epenthesis having the function of facilitating the perceptibility of the two sibilant or plosive consonants. This is in line with Boersma's account of the OCP. He claims that due to articulatory timing in sequences of [ss] or [dd], the *perceptual* identity of one of the identical obstruents is lost (1998: 416) thereby violating segmental FAITHFULNESS. Satisfaction of the OCP through vowel epenthesis creates a perceptually clear boundary between the otherwise adjacent identical segments. Consequently, one can argue that this type of OCP effect has perceptual causes. However, it also has an effect on articulation in that adjacent identical segments are avoided in the output since one of the two

[135] As mentioned above, the OCP is a product of nonlinear theories of underspecification and feature geometry (cf. chapters 4.3.1 and 4.3.2 above). In these models, two segments that carry identical feature specifications may syntagmatically be separated by intervening segments and still be adjacent on a tier with underspecified or privative features. The feature [lateral], for example, has its own tier in representations of feature geometry (Rice & Avery 1991: 103; cf. also figure 4.14 in chapter 4.3.2 above).

segments would have to be deleted, and thus a violation of MAX–IO constraints would be the undesired result.

(8.36) Past tense formation in <started>

|stɑrt + d|	OCP/CONT	MAX–IO	SON/VOI	ASSIM/LAR	DEP-IO
[stɑrtd] /stɑrtd/	*!				
[stɑrt] /stɑrt/		*!			
[stɑr̥tɪd] /stɑr̥tɪd/			*!		*
[stɑrtɪd] /stɑrtɪd/				*!	*
☞ [stɑrdɪd] /stɑrdɪd/					*

Employing the constraints which have been proposed in the previous sections, their interaction select [stɑrdɪd] as the perceptually optimal SAE output candidate. The highest-ranking OCP constraint prohibits adjacent coronal segments with identical specifications for the feature [continuant]. Deletion of a segment is ruled out by MAX–IO. SON/VOI guarantees that the liquid [r] is not devoiced, and ASSIM/LAR assures SAE-specific assimilation of laryngeal features in the sequence $[\text{+son}]_{\text{Seg}}$ $[\text{+cor, –son, –cont}]_{\text{Seg}}$ $[\text{+son}]_{\text{Seg}}$. If this constraint were violated, the second best candidate would surface, i.e. [stɑrtɪd]. This is the optimal candidate for varieties of British English, for example, which have a much lower-ranked constraint ASSIM/LAR in their hierarchy.

According to Boersma (1998: 437), the example shown in figure 8.35 above is a case of "dissimilation at a distance". This type of dissimilation, where the two affected segments are syntagmatically not adjacent, i.e. where other segments may intervene, is said to have articulatory causes. Boersma formalizes this facet of the OCP as the *REPEAT[136] constraint, which simply expresses the crosslinguistic tendency that the repetition of the same articulatory gesture is an unfavored occurrence (1998: 416).

A further instance of dissimilation is <fifths>, where many speakers dissimilate [fɪfθs] to [fɪfts] in the coda consonant cluster in order to break up the sequence of three [+cont] fricatives with a stop (O'Grady et al. 2001: 48). Dissimilation in <fifths> is perfectly in line with Yip's analyses mentioned above, where dissimilation

[136] I think it not to be of the utmost importance to decide whether *REPEAT, MAXCONTRAST(F), or OCP is the most suitable name for the constraint. For reasons of clarity and simplicity, I decided to adopt the OCP formulation for both subtypes of the OCP.

is relatively common when two adjacent coronals agree in their specification for the feature [continuant] (Yip 1989: 362-3).

(8.37) Dissimilation of a [+cont] obstruent cluster in <fifths> (example from O'Grady et al. 2001: 48)

|fɪfθs|	OCP/CONT	MAX–IO	ASSIM/LAR	DEP-IO	*GEST/COR
[fɪfθs] / fɪfθs/	*!				
[fɪfs] /fɪfs/		*!			
[fɪfθəs] /fɪfθəs/				*!	
☞ [fɪfts] /fɪfts/					*

In principle, the same ranking as in figure 8.36 above is responsible for the optimal candidate [fɪfts]. The constraint SON/VOI does not appear in the hierarchy because it is not relevant for the evaluation of the given candidates. Instead I included the MARKEDNESS constraint *GESTURE/COR in the tableau as the lowest-ranked constraint. Immediately ranked above *GESTURE/COR are of course *GESTURE/DOR and *GESTURE/LAB, which are not violated and which are left out merely for reasons of space. This shows again that only the very low-ranked *GESTURE/COR is violated, expressing the fact that coronals and not dorsals or labials are affected by the alternation. Moreover, what the analyzed examples of dissimilation have in common, is that the alternations are targeted at the nonprominent coda position and that mostly coronals are affected.

8.4 SUMMARY

In OT, the language-specific ranking of the universal set of constraints equals the grammar of the language in question. Looking at the tableaux presented in the sections on assimilation, deletion, epenthesis, and dissimilation, I would like to make the tentative claim that the overall ranking of the relevant constraints does not change in any important respects, that is, there are no glaring inconsistencies in the proposed subhierarchies that would yield contradictory results. Of course, when considering cases of assimilation, MAX–IO or DEP–IO constraints are not included in the tableaux, but that does not mean that they are not present in the hierarchy. The fact that they do not show up in the tableaux means that they are not active in this particular operation and therefore may be left out simply for reasons of space.

The general ranking for assimilation proposed in figure 8.23 above is repeated in figure 8.38 below. I claim that the different rankings for assimilation and deletion phenomena differ only insofar as the relevant FAITHFULNESS constraints IDENT–IO(F) are replaced by the corresponding MAX–IO constraints, and that the MARKEDNESS constraints ASSIM/PL or ASSIM/LAR are replaced by the MARKEDNESS constraint DEL/COR. If both deletion and assimilation take place at the same time, as is the case in <handwash> [hæmwɔʃ] shown in figure 8.28 above, the individual ranking proposed for this example MAX-IO/ONS >> IO/PLLAB >> DEL/COR >> ASSIM/PL >> IO/PLCOR does not deviate appreciably from the general subhierarchy displayed in figure 8.38 below.

(8.38) Comparison of subhierarchies for SAE assimilation and deletion alternations consisting of ranked constraints

assimilation		**deletion**	
IO/MOA	FAITH		
IO/LARSÝL	FAITH		
SON/VOI	MARK		
IO/PLONSET or IO/LARONSET	FAITH	MAX-IO/ONSET	FAITH
IO/PLDOR or IO/LARDOR	FAITH	MAX-IO/DOR	FAITH
IO/PLLAB or IO/LARLAB	FAITH	MAX-IO/LAB	FAITH
OBS/VOI	MARK		
ASSIM/PL or ASSIM/LAR	MARK	DELETE/COR	MARK
IO/PL or IO/LAR	FAITH	MAX-IO	FAITH
*GESTURE/DOR	MARK		
*GESTURE/LAB	MARK		
*GESTURE/COR	MARK		

Instead of one MARKEDNESS constraint intervening between the FAITHFULNESS constraints IO/PLLAB and IO/PLCOR there are two MARKEDNESS constraints, i.e. DEL/COR and ASSIM/PL. The presence of these two constraints in one hierarchy does not disturb the general order of the constraints since they are on one level and

ASSIM/PL can be incorporated without running the risk of making controversial claims.

Comparing the rankings needed for accounts of epenthesis and dissimilation phenomena in SAE, there are some minor differences between the two. The general FAITHFULNESS constraint DEP-IO is lowest-ranked in the hierarchy in order to enable epenthesis. It changes places with the *GESTURE constraint family for dissimilation. High in the dissimilation subhierarchy are the MARKEDNESS constraint OCP, which militates against adjacent identical feature specifications, and the FAITHFULNESS constraint MAX-IO prohibiting the deletion of segments. A high-ranking FAITHFULNESS constraint DEP-IO/ONSET in the epenthesis subhierarchy signals that epenthesis only occurs in syllable codas.

(8.39) Comparison of subhierarchies for SAE epenthesis and dissimilation alternations consisting of ranked constraints

epenthesis		**dissimilation**	
DEP-IO/ONSET	FAITH	OCP	MARK
OBS/VOI	MARK	MAX-IO	FAITH
EPENTHESIZE	MARK	SON/VOI	MARK
ASSIM/PLACE	MARK	ASSIM/LAR	MARK
*GESTURE/DOR	MARK	DEP-IO	FAITH
*GESTURE/LAB	MARK	*GESTURE/DOR	MARK
*GESTURE/COR	MARK	*GESTURE/LAB	MARK
DEP-IO	FAITH	*GESTURE/COR	MARK

I have shown that the common ground in all rankings is that constraints referring to *positions of privilege* and constraints referring explicitly to *coronals* (vs. labials and dorsals) play a crucial role in this investigation of the behavior of coronals in SAE.

CHAPTER 9

An OT critique—some unresolved issues

9.1 INTRODUCTION—THE NATURE OF OT

For obvious reasons, I will not present a general critique of OT in this last chapter; instead, I will comment on certain crucial points of the specific version of OT which I have presented in this dissertation. Viewed from the perspective of the preceding analysis, I think some of these issues are in need of further research.

At the beginning of the 1990s, the time was ripe for a constraint-*based* model of generative phonology (GP) to replace the older derivational frameworks. This is impressively shown in Goldsmith's (1993) book with the telling title *The Last Phonological Rule: Reflections on Constraints and Derivations* and in LaCharité & Paradis' (1993) paper on the emergence of constraints in GP, which includes the authors' presentation of three constraint-based approaches to phonology. In addition to OT, they introduce Declarative Phonology (DP) and the Theory of Constraints and Repair Strategies (TCRS). Of these three models, clearly OT is the theory of language which receives most recognition and in which most research is being done.

In chapter 4, I have shown that the notion of constraint as a restriction on the phonological shape of an output is not an entirely new concept. Previously, constraints had frequently been invoked *in addition* to rules. What is new in these recent models, however, is the constraints' crucial importance for the theory. While in all three models it is only the constraints which are responsible for the output form, there are some differences with respect to the exact nature and function of constraints in each of the frameworks. Under DP, constraints are nonuniversal, language-particular but inviolable tools that function as redundancy rules; however, the constraint *ranking* is supposed to be universal. TCRS has both language-specific and universal constraints which function as principles and parameters. The ranking of the constraints and the repair strategies in cases of violation of the constraints are assumed to be universal in this model (LaCharité & Paradis' 1993: 135-49). What I find most appealing about OT is the idea of a finite set of constraints, all of which are both universal and violable. In contrast to DP, the *ranking* of the universal constraints is language-specific, and it is this specific ranking which equals the grammar of the language in question.

(9.1) Basic OT architecture (after McCarthy 2002: 10; repeated from chapter 4.1.2 above)

$\boxed{\text{Lexicon}} \rightarrow$ input $\rightarrow \boxed{\text{Gen}} \rightarrow$ candidates $\rightarrow \boxed{\text{Eval}} \rightarrow$ output

At the very center of an OT model are three main components which are subject to critical scrutiny in the following sections: first, there is the LEXICON with the basic forms of morphemes listed from which inputs are constructed; then there is the GENERATOR, which creates a multitude of output candidates for the EVALUATOR, which in turn selects the optimal output candidate on the basis of the language-specific ranking of the universal set of constraints in CON.

9.2 THE INPUT—AN OT LEXICON

Archangeli & Langendoen are very specific about the fact that the question of the nature of the input is one of OT's most central unresolved issues pending further research (1997: 200).

A crucial property of an OT LEXICON in comparison to older generative models of the lexicon is that underlying or lexical representations are not subjected to the restrictions of any type of constraint. Lexical representations *cannot* be under the control of one or more constraints because the lexical items' evaluation of the constraints in EVAL is an operation that takes place after the generation of candidates for a given input (cf. figure 9.1 above). If this were not the case, OT would find itself in the paradoxical situation of having to recognize constraints on constraints. In chapter 6 above, I showed that even sound inventories, which in standard GP arise from severe constraints on the input in terms of feature cooccurrence restrictions, are the outcome of the language-particular constraint ranking. This problematic feature of OT, which basically states that all lexicons are identical, is usually discussed under the heading Richness of the Base. I repeat Smolensky's clarifying statement from chapter 4.1.2 above:

> The source of all systematic cross-linguistic variation is constraint reranking. In particular, the set of *inputs* to the grammars of all languages is the same. (1996a: 3)

The sole requirement on an input form is that it must be a well-formed linguistic object. Apart from OT's rather vague assumption that UG provides a set of linguistic units for language representation, i.e. features, segments, syllables, and lexical categories such as N, V, and A, from which inputs are constructed, it is not entirely clear what exact shape inputs might assume (Archangeli 1997: 13-4).

The theoretical precondition that in principle anything can be the input to a certain surface form (provided that it is a linguistic object) creates a problem for learnability issues under OT. I mentioned in chapter 4.1 above that OT, being a completely surface-based approach to language, focuses on the output form since this is *the* one candidate among many which is computed by EVAL as the optimal candidate. The question now is how children acquiring a language are able to select the optimal output form, not yet knowing the language-specific ranking of the set of universal constraints, from a very abstract, vague, or remote input form. It is indeed legitimate to ask what kind of input form children would hypothesize for an output form [kæt], for example. Tesar & Smolensky state the initial task of the language learner in quite explicit terms:

> Under the assumption of innate knowledge of the universal constraints, the primary task of the learner is the determination of the dominance ranking of these constraints which is particular to the target language. (1993: 1)

The input, however, is of crucial importance for the construction of a language-specific constraint ranking in CON, considering that e.g. FAITHFULNESS constraints, which make reference to *correspondences* between input and output forms, are involved in most evaluation operations. This precisely constitutes the theoretical impasse; inputs or underlying representations (URs) are required to establish EVAL but the input itself can be anything because EVAL determines the winner. Hale & Reiss describe the "vicious circle" in which the language learner finds herself in the following way: "the child needs a ranking to get URs and needs URs to get a ranking" (1998: 666). In response to this paradox, Prince & Smolensky (1993: 192) propose a somewhat dubious rescue mechanism, termed principle of Lexicon Optimization, which is elaborated in a paper by Itō, Mester & Padgett:

> Of several potential inputs whose outputs all converge on the same phonetic form, choose as the real input the one whose output is the most harmonic.
> (1995: 593)

This of course means, choose the one which most closely resembles the output, and one is tempted to add: choose as the real input the output itself. What Lexicon Optimization in fact demands of learners is to start language acquisition 'backwards', as it were. It is assumed that children hearing the lexical item [kæt] construct an input form |kæt|, from which they can proceed to build the dominance hierarchy of constraints in CON. This 'principle' does not appear to be very satisfactory and even to date OT researchers have not come up with a better solution yet. In their recent book on learnability in OT, Tesar & Smolensky simply state:

> For the moment, we retain the original context of this principle [Lexicon Optimization], and assume the correct grammar has already been learned; lexicon optimization is then used to acquire new lexical entries. (2000: 78)

It is not clear *when* the child has acquired the grammar, i.e. the language-specific constraint ranking. McMahon summarizes the paradoxical and hopeless situation with respect to learnability issues under OT in the following comment:

> It is impossible to evaluate acquisition under OT fully at present: issues of constraint ranking cannot be resolved in the current state of indecision about the types of constraints and interactions permitted, […] the nature of the input as one of the major unresolved issues in the theory. (2000: 56)

Apart from the fact that the assumption of an input identical to the output form reverts to theoretical circularity, it also creates an obvious problem for all types of alternation. In the previous chapter, I showed how important a certain input form, e.g. |æ.təm|, is for the computation of the optimal variety-specific candidate. If the theory requires inputs equal to outputs, i.e. [æ.ɾəm], then inputs are indeed superfluous since each and every winner falls out directly from EVAL anyway, independent of the input form. The theory then has returned to Flemming's proposal within the framework of Dispersion Theory, which I presented in chapter 5.5.3 and which, after some discussion, was rejected. Flemming (1995; 1997) claims that only surface contrasts should be evaluated by an OT grammar, and not correspondences between input and output specifications. URs are thereby erased from the grammar completely. The problem, however, is that phonetically grounded constraints in this model do not evaluate individual linguistic expressions, e.g. |æ.təm|, for perceptual distinction and articulatory effort, but linguistic generalizations, e.g. sound inventories, are checked for an optimal dispersion of sounds in terms of perceptibility and articulatability (cf. chapter 5.5.3 above). In OT inputs are needed for the generation of a candidate set in GEN (cf. section 9.3 below). EVAL must be in the position of evaluating a rich candidate set and not optimally organized phonological systems, as McCarthy points out (2002: 225-7).

A persistent question thus remains as to what the exact nature of the input might be. Those OT researchers who do not emphasize the problematic nature of the input assume—as I do—more or less surface-based underlying or lexical representations as inputs. In my version of OT, phonetically grounded constraints evaluate the candidate set generated by GEN from the relevant UR and select the optimal candidate according to the specific[137] ranking. Hayes (1996), for example suggests that the set of phonetically grounded constraints is discovered inductively by children learning the language as they have direct access to their own perceptive and productive capabilities and are able to judge the relative complexity of percepts and articulations. They construct their URs from adult outputs, which are then used to build the constraint hierarchy for the language to be acquired (1996: 19-21). Approaches to OT which do adhere to URs no matter in what form, and no matter how surface-true these are, must face the accusation of harboring a "derivational residue" and "hidden rules" in their theoretical framework, since one level of representation, i.e. outputs, is formally linked to another level of representation, i.e. inputs, via FAITHFULNESS. Thus, one might say that outputs are *derived* from inputs (Oostendorp & Hermans 1999: 1-27; LaCharité & Paradis 2000: 211-33; cf. also figure 4.11 in chapter 4.3 above, where Goldsmith's (1995: 3) table depicting the basic tools of phonological theory has rules *and* OT constraints in one box). It is self-evident that OT officially bans serial, rule-based derivations; it does permit, however, the computation of the optimal candidate in one step in parallel processing across the whole set of universal constraints, provided that there are a vast number of

[137] With reference to the question of how specific the ranking in Eval must be cf. section 9.5 below.

potential output forms. What was formerly done step by step by a series of P-rules, must now be accomplished in one single operation. This leads straight to the problematic status of GEN in the theory, where rule behavior is supposedly masked by the assumption of GEN's unbounded creativity and almost total unconstrainedness.

9.3 GEN—THE CANDIDATE SET

As was shown in chapter 4.1.2 above, along with EVAL (cf. section 9.5 below) GEN is one of OT's two main functions which are responsible for the selection of the optimal output candidate. GEN is assumed to work the same in all languages and its sole function is to provide a large candidate set, allegedly independent of constraint requirements. Looking closely at GEN, however, its workings are dependent on at least *one* constraint: it must generate linguistically well-formed output candidates for a linguistically well-formed input (Archangeli 1997: 14). According to some researchers (e.g. Archangeli 1997: 29; Sherrard 1997: 47; McMahon 2000: 18), this fact should not be overlooked because it sheds doubt on OT's strong claim that all constraints are violable in principle; these linguists are concerned about the fact that the theory might have to include at least one inviolable 'metaconstraint'.

(9.2) Schematic organization of constraints within an OT grammar (after Sherrard 1997: 47)

inviolable constraint(s) >> ranked violable constraints >> inactive constraints

On the other hand, following Kager (1999), it can be argued that OT is capable of ruling out logically possible but crosslinguistically unattested structures through the appropriate ranking of all violable constraints. Kager states: "an empirically restrictive linguistic theory requires no inviolable constraints" (1999: 105). How this is to be accomplished in detail, however, has not yet been shown as no one has established a complete dominance hierarchy of CON for any single language or variety up to this date.

LaCharité & Paradis (1993) criticize the very fact that GEN is allowed to add and delete structure[138] in the input form selected from the LEXICON in order to produce a sufficiently rich candidate set *prior* to encountering the constraints, rather than in anticipation of the language-specific constraint ranking in EVAL. The necessary consequence of this theoretical requirement is an enormous overgeneration of candidates. Actually, a virtually infinite set of nonoptimal output candidates is created, all for the sake of one single optimal candidate. In chapter 4.1.2 above, this feature of OT was discussed under the heading of Freedom of Analysis. LaCharité & Paradis claim that Freedom of Analysis, which is an optimality-theoretic must, is an "extremely costly and unexplanatory device" (1993: 147), since it neglects commonly accepted scientific expectations on theoretical constructs, such as the economy principle (cf. discussion of Occam's Razor in chapter 1 above) and the demand for predictive power (cf. discussion of Chomsky's three adequacy criteria for analyses in chapter 1 above). It could be argued that GEN is a device, which is allowed similarly mysterious and abstract ways to the P-rules of standard GP. For the latter holds: anything can be the input to a P-rule; no input can be too abstract to be dealt with by the rules. For the former it holds: anything can come out of GEN, any conceivable, no matter how implausible, candidate can be created by GEN. Without a rich candidate set, however, the chances of retrieving the optimal candidate are minute. On the other hand, if there are no constraints on GEN whatsoever, the candidate set will indeed be infinite. LaCharité & Paradis summarize the implications of the problematic nature of GEN: "After all, why must we ever settle for an output that violates a constraint when an infinite field of candidates would make available forms that conform to all the constraints?" (1993: 148)

The rather vague and highly hypothetical character of GEN raises the further question of the ways in which GEN is allowed to manipulate the input in order to produce the candidate set. McCarthy & Prince claim that the output candidates produced by GEN come with changed, deleted, inserted, and permuted features or segments or even with altered prosodic structure (1993: 86). LaCharité & Paradis

[138] It is quite amazing to see what GEN is able to do on the syntactic level. I will cite Kager to give a brief impression of syntactic operations allowed by GEN (1999: 347):

- introducing (extended) projections conforming to X' theory
- introducing functional heads that do not appear in the input due to their lack of semantic content (e.g. the complementizer *that* and *do*-support in English)
- introducing empty elements (traces, etc.) and their coindexations with other elements
- moving lexical elements

rightfully formulate their suspicion that the many *different* operations which GEN has to perform, i.e. deletion, insertion, and permutation of features or segments, makes processing in one single step implausible (2000: 228). If at worst the candidate set produced by GEN is indeed an infinite list, psycholinguistic and computational models of language clearly will be at a loss to integrate OT into their frameworks, as Archangeli points out, "since neither [model] responds happily to infinite sets" (1997: 29). She also comments on the conventions for representing the generated output candidates in OT tableaux, namely restricting the candidates "to those which are critical to the point being made" and gives the following reason for this procedure: "the infinite set could not possibly be considered!" (1997: 12). One might indeed ask not how the linguist but—as McMahon does—how hearers and speakers manage this task of reducing the infinite set to a handful of relevant candidates since this issue undoubtedly raises problems for speech processing and production (2000: 30).

Up to now, not much in the way of new insight with respect to GEN has been achieved: "virtually all research in OT to date has glossed over the inner workings of GEN" (LaCharité & Paradis 2000: 212).

9.4. THE CONCEPT OF THE CONSTRAINT

Constraints under OT are assumed to be universal and violable statements, which are formulated in a positive, negative, or implicational format (LaCharité & Paradis 1993: 127-33).

A point of criticism that might come to mind immediately is that OT seems to be an 'old' model dressed in 'new' clothes. For example, OT's basic machinery CON (the universal set of constraints) is reminiscent of Greenberg's (1963) implicational universals of the format: if a language L has the property Q, it will also have the property P. Depending on the combination and the sequence of the universals (or the ranking of OT constraints, as one might say), a considerable amount of crosslinguistic variation with respect to these universals, e.g. a language may have SOV, SVO, or VSO[139] word order, and their implications will result in language-specific grammars, each one with a singular shape. OT's constraints express similar

[139] According to Greenberg (1963), the other three word order types are very uncommon.

notions with a different formal system. Within an OT framework, crosslinguistic variation comes about as grammars of individual languages include a LEXICON with the basic forms of morphemes listed from which input forms are constructed, and the second of the two functions EVAL (cf. section 9.5 below) which provides a specific ranking of the universal set of constraints in CON. The set of universal constraints equals the former universal principles which make up UG (Kager 1999: 4).

9.4.1 Absolute vs. nonabsolute universals

All models of generative theory previous to OT, including Principles & Parameter Theory (PPT) and Minimalism, subscribe to statements of the following kind:

> UG contains a set of absolute universals, notions and principles which do not vary from one language to the next. […] There are language-specific properties which are not fully determined by UG but which vary cross-linguistically. For these properties a range of choices is made available by UG.
>
> (Haegeman 1994: 15)

The principles of UG are not only very general and abstract statements, but also absolute, that is, inviolable principles, which are formulated as universally quantified propositions. What is specified in the proposition applies in all cases, with absolutely no exception. The logical format of absolute and inviolable universals can be stated as follows: $\forall x\ [Px \rightarrow Rx]$; that is, for all x it holds, if x is a language (i.e. has the property P), then x also has the property R (i.e. has oral vowels).

(9.3) Candidates for absolute universals in the format $\forall x\ [Px \rightarrow Rx]$

All languages have oral vowels (i.e. nonnasal vowels).
All languages have CV syllables.
All languages distinguish between verbs and nouns.
All phrases are headed.

Crosslinguistic variation is guaranteed by the different settings of parameters, i.e Haegeman's "range of choices" (e.g. all phrases of a particular language are left-headed as in English, or all phrases are right-headed as in Japanese). An important issue within such a theory has always been how to account for the exceptions, when a specific phrase, for example, is not structured the way it should be. A general answer to this kind of problem has always been that the structures visible at the surface are not necessarily the underlying structures.[140] Movement rules (syntax) and P-rules (phonology) are responsible for deriving surface-near structures from at times extremely abstract underlying structures. Within the PPT framework, a fairly recent development is the proposal of the Universal Base Hypothesis. Here it is argued that the two basic structures of phrases, either the head-initial type (English) or the head-final type (Japanese), are not the result of parametric variation. Instead, a principle of UG stipulates one of the two as the universal base pattern, in this case the head-initial phrase type. In order to generate the required phrase structure, head-final type languages, where the complement precedes its head, need a leftward movement operation ('object shift') analogous to leftward movement of the subject DP into the specifier position of the functional phrase AgrP (Haegeman & Guéron 1999: 592-3). So far as I can see, the Universal Base Hypothesis adds an extra level of abstraction to PPT.

An advantage of absolute, inviolable universals is that they reflect the intuitive notion of very general principles which all languages share, i.e. all x have the property P, and which indeed have no exceptions. Contained in these absolute universals are the parameters that account for variation among the languages and that are fixed during the acquisition process of the specific language. A disadvantage of this type of model has been seen to be in the way exceptions are accounted for and in the asymmetry between a handful of vaguely defined universal principles and a huge number of language-specific parameters (previously, rules, rule-orderings, and transformations were involved) which are designed to 'explain' the intricacies and idiosyncrasies of individual languages.

In order to simplify this situation with basically two separate grammar components, OT has come up with the notion of a universal set of constraints that are all violable in principle, i.e. nonabsolute universals. Now there is only one main grammar component (CON), and exceptions are accounted for by a specific ranking of the constraints in CON. There are also some further advantages of nonabsolute principles or constraints that are worth mentioning. Absolute universals can be viewed as a special case of violable universals, but not vice versa. This is in fact

[140] Another typical answer to this type of problem is to question the validity of the data. Usually this way of resolving theoretical impasses is associated with Chomsky himself. In a recent publication he admiringly describes Galileo's methodology as follows: "[Galileo] dismissed a lot of data; he was willing to say 'Look, if the data refute the theory, the data are probably wrong.' And the data that he threw out were not minor." (2002: 98).

what is argued by OT proponents who support the strong claim that *all* constraints are violable in principle when the controversy caused by GEN is under discussion. I showed in section 9.3 above that OT is in danger of having to posit at least *one* inviolable constraint, i.e. one absolute universal, namely that GEN is resticted to producing only well-formed linguistic objects as potential output forms. Taking this as a special case, it can be argued that it is a mere coincidence that this specific constraint is always observed because its violation would take language outside the permitted range of crosslinguistic variation.

(9.4) Nonabsolute universals (after Whaley 1997: 32)

> Most languages have at least *one* coronal stop (cf. Maddieson 1984).
> Most languages have the vowel [ij] (cf. Maddieson 1984).
> Most languages employ rising intonation to signal a yes/no question.
> Most languages have adjectives.

Nonabsolute principles, i.e. OT constraints, express more or less strong tendencies. Violable universals can be ranked and even graded in the sense that the gradable concept of the preferred, unmarked, or optimal can be expanded to account for the elements of a class or category in a markedness hierarchy, e.g. coronal unmarkedness expressed in the ranking: *GESTURE/DOR >> *GESTURE/LAB >> *GESTURE/COR.

Violable universals can apply locally, i.e. a language or structure x is more preferred or optimal than another language or structure y with respect to one specific feature, e.g. CORONAL, in a specific context (cf. chapter 8 above).

A disadvantage of assuming violable universals or markedness relations (cf. section 9.6.2 below) as the core of the grammar is that they do not allow for deductive-nomological inferences or explanations (Lass 1980: 9-12; 17-20; cf. also section 9.6 below on explanatory notions). The statement, for example, that most languages have at least *one* coronal stop (Maddieson 1984) says nothing about a specific language. All that violable universals do is raise expectations. Thus, violable universals are hard to falsify (cf. the brief presentation of Popper's critical rationalism in chapter 1 above).

9.4.2 Universality and innateness of CON

Following the basic generative tenets of UG, some OT researchers assume that each constraint which is part of CON must be both universal and innate (Archangeli 1997: 15; Kager 1999: 1, 4). This is in line with all previous, strongly formalist approaches to the study of language, more or less loosely associated with Chomsky. A parallel with Natural Phonology (discussed in chapter 4.2.2 above) is obvious as well. NP accounts for the obvious differences between the speech of very young children and that of adults speaking the same language by assuming that children are born with a large number of innate and natural processes. Part of the acquisition process consists of suppressing some of these natural processes in favor of language-specific rules, which must be acquired subsequently. The task of the child faced with an innate CON then is reranking.

On the other hand, given the fact that OT's explanatory force relies to a considerable degree on language typology, it is not surprising that functionally and typologically oriented research (following Greenberg), which aims at explaining linguistic structure primarily in terms of its linguistic function, does not see any necessity in postulating an innate set of universal constraints. With their strong emphasis on phonetic grounding of phonological analyses, functionally-based OT approaches (e.g. Boersma 1998: 161-2) find it perfectly plausible to invoke perceptual and articulatory factors which determine and restrict the range of possible sound systems instead of assuming, for example, innate phonological feature hierarchies.

Given OT's general claim that language-specific constraints are disallowed, that all constraints posited must indeed be universal, there are some constraints in the OT literature that call this claim into question saying that these constraints are far too specific to be universal. This has been criticized by quite a few researchers (e.g. Archangeli 1997: 15; Kager 1999: 11; McMahon 2000: 20, 24). As a case in point, constraints which refer to specific frequencies or formant structure seem to be such dubious candidates. For example, Boersma (1998) has a perceptually-based FAITHFULNESS constraint family (*WARP) with individual constraints of the shape *WARP (F_1: [440], /300/) "do not initially classify an acoustic input of 440 Hz as a high vowel of 300 Hz" (1998: 147). Similarly, Flemming (1995) has constraints on perceptual contrasts in inventories (MINDIST), which require a specified minimal auditory distance between contrasting forms, e.g. MINDIST = F_1:2 or F_2:4, i.e. minimal distance of at least 2 on a scale on the F_1 dimension or of at least 4 on the F_2 dimension (1995: 25-33).

On the other hand constraint families as e.g. the MAXIMALITY and DEPENDENCY constraints, which I used in this investigation, seem justifiable as universals. These FAITHFULNESS constraints generally prohibit deletion or epenthesis of segments unless they are dominated by some MARKEDNESS constraint requiring just that. In the dubious situation in which some of the very specific constraints find themselves, it might indeed be much wiser to refrain from such strong claims as all constraints are innate. In this dissertation, I have tried to use only constraints whose universality is not in doubt.

9.5 EVAL—THE *-SPECIFIC RANKING

After GEN (cf. section 9.3 above), EVAL is the second of OT's two basic functions. EVAL provides a language-specific or variety-specific[141] ranking of the universal set of constraints in CON. EVAL ensures that only those output forms will surface as optimal that belong to the language or variety in question.

It can be criticized that an OT analysis conducted with a single language as data base—or even a single language variety, i.e. SAE—is bound to remain at a purely descriptive level even if 'external' evidence, for example phonetic grounding, is invoked as an explanatory notion. OT indeed seems to be largely dependent on language typology to exhibit the full range of explanatory adequacy. Only when comparing different, possibly unrelated languages, can the universality of the constraints (cf. section 9.4.2 above) be demonstrated through different rankings of one and the same set of preferably universal constraints, i.e. CON.

Nevertheless, even for single language or variety analyses, OT does provide a grammar framework that allows for the representation of contradictory forces, e.g. the conflicting demands of perception and articulation (cf. section 9.6.1 below). I am convinced that a potential for representing such conflicts is essential in a theory of language.

[141] The asterisk (*) in the heading stands for the indeterminacy as to how specific a given ranking must be assumed to be (cf. below).

A remaining issue is the question of how specific a ranking in EVAL must be to compute a particular output candidate, especially in the case of alternations. Alternations, e.g. assimilation, deletion, epenthesis, and dissimilation, are frequently optional. That is, it may be up to the speaker in a given context and a particular situation to produce an assimilated form or not do so. The possibility of choice is reflected in the corpus data I used. At times, I found the same form uttered by the same speaker, assimilated in one sentence and then a sentence or two later in its unassimilated form. This means that a variety-specific ranking does not account for all the forms but that a speech style-specific, even speaker-hearer-specific, or situation-specific ranking may be necessary, at least for cases where language users do have a choice, since output forms are solely determined by the constraint ranking. This is an unfortunate situation indeed, as Guy (1997: 135-40) and McMahon (2000: 105-15) point out. Guy finds it "implausible that human speakers would construct different grammars to encompass the facts of one language", and "Occam's Razor argues against such a general duplication of theoretical constructs" (1997: 138).

A possible reply to these objections is, however, that in view of the complete constraint hierarchy, only very minor reranking is required of speakers and hearers in the specific situation where either an assimilated or an unassimilated form is used.

9.6 EXPLANATORY NOTIONS UNDER OT

In this final section, I will discuss whether the principles of *maximization of perceptual distinction* and *minimization of articulatory effort* and concepts of *markedness* can be accepted as explanatory notions.

9.6.1 Perceptual distinction and articulatory effort

Throughout the development of generative phonology, its has frequently been argued that phonetic factors and explanatory notions deduced from these factors, as e.g. the principles of *minimization of articulatory effort* and *maximization of perceptual distinction*, are concepts which are far too fuzzy to be of any use for a formal phonological analysis as they have neither explanatory significance nor predictive power (Lass 1980; 1984; Anderson 1981). Lass and Anderson criticize that phonetic representations refer to gradient and concrete phenomena while phonological representations are categorical and abstract since their function is to reveal phonologically significant phenomena. Both acknowledge that functional criteria should not be written off as irrelevant facts but

> when a process is phonologized, it becomes in an important sense phonetically arbitrary, even though it may continue to reflect a phonetically natural content. Phonetic explanation thus serves as a sort of constraint on the entry of processes into the system, since […] rules originate in the raw material of phonetically natural variation. Once phonologized, their essential character is radically altered. (Anderson 1981: 514[142])

Lass (1984: 199-200) and Anderson (1981: 515-28) point out that what seems to be natural in one language might not be so in another, e.g. certain simplification processes. Their line of argument is that this crosslinguistic variability of phonetic principles constitutes the problem for their overall applicability.

Within the formal framework of OT, however, variation among different grammars, i.e. variation with respect to the *extent* to which certain individual principles find their application in a specific grammar, does not pose a problem for the theory. Instead, this type of crosslinguistic variability is even predicted, since OT's basic property is the general violability of each universal constraint present in the set of constraints. The language-specific ranking of constraints and the subsequent computation of the optimal candidate takes place with respect to "all linguistically relevant dimensions, not just one" (McCarthy & Prince 1994: 2). In their paper "The Emergence of the Unmarked", McCarthy & Prince discuss a common misconception about OT, which they term the "Fallacy of Perfection". Linguists adhering to this notion equate 'optimal' with 'perfect', i.e. an optimal output candidate would have to be the perfect candidate (ibid.). As a representative of this doctrin, McCarthy & Prince quote Chomsky from an unpublished manuscript:

> Since there is some least marked vowel and some least marked consonant, why isn't every word in every language exactly the same? Concretely, why aren't all words simply [ʔiʔi], [tata], or perhaps [baba]?
> (Chomsky 1994; quoted from McCarthy & Prince 1994: 2)

In fact, this issue was raised years ago by Lass in his discussion of naturalness and markedness concepts as explanatory notions (1980: 15-44; 1984: 195-200; cf also chapter 4.2 above and section 9.6.2 below).

[142] For similar statements cf. Lass (1984: 197-8) and Hyman (2001: 143).

In order to weaken Lass' and Chomsky's criticism, as an OT proponent I would argue that previous generative theory could not cope with the two principles because *maximization of perceptual distinction* and *minimization of articulatory effort* were treated as *inviolable* extralinguistic principles (Lass 1980: 17-20). According to this assumption, Lass is justified in not accepting *minimization of articulatory effort* as an explanatory concept in the strict sense of a deductive-nomological explanation:

(9.5) Deductive-nomological explanation (after Lass 1980: 9)

Why X?
X because Y.
The best explanation is X, because it could not have been otherwise (because Y).

Deductive-nomological explanations follow a certain schema. There is a fact that needs to be explained (E = explanandum), e.g. nasal place assimilation in [nb] → [mb], and a conjunction of certain conditions (C), e.g. the sequence [nb], and a set of general laws (L), e.g. [mb] is easier than [nb] (Lass 1980: 10). Assuming a hypothetical percentage of 90% assimilated forms, he rightfully points to the particular instances, i.e. 10%, that are not assimilated (null strategy) or where a different solution is preferred, e.g. [bb] (multiple strategy; 1984: 199-200). The 'law' is merely a statistical or probabilistic statement which neither explains the single instance nor does it predict the occurrence or nonoccurrence of the single instance.

However, the principles which are encoded in an OT grammar as constraints are not absolute laws where one instance that does not observe the law annuls its status as a law. These principles are violable in the sense that a high or low ranking of a given constraint will result in different individual grammars. As violable constraints under OT, these functional principles are well-suited for a phonological analysis of alternational phenomena, as I tried to show in the preceding chapters. The interaction of particular substantive considerations which are not language-specific with phonological factors that are language-specific is depicted in their formalization as conflicting FAITHFULNESS and MARKEDNESS constraints which serve to enrich phonological theory and bring the goal of explanatory adequacy one step closer.

9.6.2 Markedness

The introduction of markedness (MT) theory by Chomsky & Halle (1968: 400-35) and its subsequent development by Kean (1975) served the important purpose of incorporating phonetic substance into the formal system of generative phonology; furthermore, markedness was introduced to weaken the propositional strength of very general and abstract statements which characterize generative theory, i.e. the principles of UG discussed in section 9.4.1 above.

However, markedness is a controversial notion and moreover has received quite a few different definitions, amomg them the following: a marked segment is less natural, more complex, and not as stable as an unmarked segment; it is harder to articulate and perceptually more salient; it appears in fewer languages, in fewer positions, and later in language acquisition. For the sake of convenience, I will summarize further criteria for markedness presented in this dissertation:

(7.21) Diagnostics for markedness

- Implication: marked features or segments imply unmarked ones (e.g. in inventories).
- Frequency: segments with unmarked features are more frequent than marked ones.
- Alternations: unmarked features are more prone to assimilation, deletion, epenthesis, and dissimilation than marked ones.

One notable and still valid objection to MT is provided by Lass (1980: 42-4), who argues that markedness is based mainly on the intuitive judgement of the linguists and not on empirical facts. Since what is common and expected is 'unmarked' and 'natural' and the very commonness of the common is a sign for its superiority, the uncommon is less stable and will eventually be replaced by the common and unmarked state, therefore being more natural and *optimal* than before (1980: 43). In this line of argumentation Lass sees an expression of "the blinding tautology that nature tends toward the natural" (ibid.). There is clearly a problem concerning the predicative power of a theory which evaluates language states in terms of cost and *optimization*, when a system goes from an unmarked to a marked state. This was the case, as Lass points out, with the Germanic i-umlaut in which back vowels fronted before a following /i/ or /j/ (1980: 133). A further problem can be seen in the fact that not all systems arrive at the ultimately unmarked, simple, and expected stage ([tata], [baba], etc.; cf. also quote from Chomsky in section 9.6.1 above).

One of Schachter's two marking conventions (1969: 346) which try to deal with the fact that vowels become nasalized when adjacent to a nasal consonant illustrates the problematic nature of the theory: unmarked feature values assimilate to marked feature values. One simply wonders why assimilation should work this way rather than in the opposite direction. Lass states further:

> The problem is […], there ought to be consequences assignable to marked states: instability, difficulty in learning, etc. And there is no real evidence that any such consequences exist. (1984: 198)

Anderson claims that the various problems in connection with notions of markedness, e.g. its circularity, do not arise because an adequate set of marking conventions has not yet been discovered. Instead he states:

> [P]honetic content in a phonological system can only be analyzed relative to the other properties of the system. If this is true, it is simply not possible to embody this role in a comprehensive and universal way in the definition of the notation in the way foreseen by markedness theory. (1985: 334)

Unless there is some empirical foundation for markedness, it will remain a circular notion in Lass' (1980: 43) sense even in OT. In this sense, I agree with Anderson (1981: 536) that an understanding or even an explanation of the language facts is only successful if both the so-called external factors, i.e. phonetics, governing linguistic variation and the so-called internal factors governing the systematization of language phenomena in a grammar are recognized. The incorporation of phonetic content into phonological analysis, however, poses a problem for Anderson because he must have envisioned markedness as an *inviolable* principle as his generative framework still lacks the means of expressing the relative aspects of markedness.

Under OT, the concept of markedness is formalized in a very specific way: as a whole family of MARKEDNESS constraints which makes up a large part of the sum total of constraints. The version of OT which I propose aims at supplying an empirical basis for these constraints by grounding markedness, i.e. what constitutes a marked element or an unmarked element, in phonetic facts. In addition, OT stresses the relational aspect of potentially conflicting forces in natural languages by introducing *violable* MARKEDNESS constraints, which are capable of expressing the fact that certain elements are marked or unmarked only in comparison to other elements in the system. Since MARKEDNESS constraints demand simple structures and FAITHFULNESS constraints require that inputs are always realized in accordance with their lexical specifications, regardless of their complexity, the two kinds of constraints frequently impose conflicting demands on outputs. These conflicts must be negotiated in specific ways by language-particular constraint systems, the resolution of which in individual grammars is one of OT's most important self-acclaimed goals. Within such a model, markedness thus escapes the threat of circularity.

From a more distanced and philosophical perspective, Derrida proposes a general critique of the notion of markedness as such (1977: 162-254). He rejects the structuralist concept of markedness of binary oppositions by exposing markedness as a conventional setting-up of "hierarchical ethical-ontological value-oppositions around an ideal and unfindable limit" (1977: 236). Moreover, the conventions have the effect of subordinating these values to each other and thus establishing *asymmetrical* relations between them, e.g. *standard vs. parasite* (cf. standard vs. nonstandard varieties of speech), *simple vs. complex* (cf. feature specifications of segments: simple = optimal), *essential vs. accidental* (cf. common vs. uncommon phenomena; Lass 1980: 43), *normal vs. abnormal* (cf. natural vs. unnatural phenomena; Lass 1980: 43), and the ultimate opposition *positive vs. negative*. In the perpetuation of these asymmetrical relations along with the value-judgements attached to them, Derrida rightfully sees a strategical return to

> an origin or to a "priority" held to be simple, intact, normal, pure, standard, self-identical, in order *then* to think in terms of derivation, complication, deterioration, accident, […]. And this [the hierarchization of value-oppositions] is not just *one* metaphysical gesture among others, it is *the* metaphysical exigency, that which has been the most constant, most profound and most potent. (Derrida 1977: 236)

CHAPTER 10

Conclusions

10.1 CORONALS

This investigation has shown that coronals are indeed very special sound segments. There is abundant evidence from various fields of phonetics which clearly establishes coronals as a class of consonants appropriate for phonological analysis.

As has also been shown, the set of coronals is stable across varieties of English unlike other consonant types, e.g. labials and dorsals, which are subject to a greater or lesser degree of variation. From this observation, I was able to deduce that coronals have an important function to fulfill: they exhibit stability in inventories, but they simultaneously display flexibility in alternations when it is required by the contradictory forces of perception and production.

These generalizations seem to hold not only for varieties of English, but also crosslinguistically. For various articulatory and perceptual reasons, inventories usually have a relatively large number of coronals in them; onset and coda obstruent clusters contain at least one coronal (Greenberg 1978: 268-9)[143]. Therefore, coronals are regarded as the unmarked class of consonants. Moreover, these unmarked coronals are targeted by alternations such as assimilation while noncoronals trigger them.

Within the alternations assimilation, deletion, epenthesis, and dissimilation, one can surmise the following: the coronal segments that are subjected to

[143] Greenberg's exact formulations are:

Universal 38: Every language with final clusters contains at least one cluster with a final obstruent in the dental-alveolar region.

Universal 39: Every language with initial clusters contains at least one cluster with an initial consonant in the dental-alveolar region.

assimilation asymmetrically are also the segments that are deleted, that are frequently inserted in epenthesis, and that are affected by dissimilation.

10.2 OT

Irrespective of the various points of criticism discussed in chapter 9 above, I think that OT is a promising theory of grammar. Just like its predecessors—various models of GT (e.g. PPT, Minimalism)—OT demands high standards with respect to explanatory adequacy and universality. Like all other models of GT, OT aspires to be a grammar of universal scope (UG).

Considering Chomsky's requirements on adequacy once more, it is questionable whether OT does in fact meet his condition of explanatory adequacy for a theory or whether OT remains at the level of descriptive adequacy:

> The standard way to express that distinction is to take a descriptively adequate theory to be a true theory of an attained state, whereas an explanatorily adequate theory is a true theory of the initial state. So, in this view there is a sharp distinction between the initial state, the topic of Universal Grammar, and the attained states, the actual languages. (Chomsky 2002: 130-1)

OT does supply a descriptively adequate analysis of the surface phenomena of the investigated languages; however, the question still remains whether it succeeds in providing an explanatorily adequate account of Universal Grammar through CON (the universal set of constraints).

It seems to me that using an OT approach which integrates facts from speech perception and production and which thereby grounds phonological theory in phonetics is a very promising one. OT's machinery of inherently conflicting FAITHFULNESS and MARKEDNESS constraints which reflect the substance of communication events at large is well-suited to both of OT's goals, i.e. to provide typologically well-founded analyses of language as such, as well as accounts of a specific language or variety.

REFERENCES

Aitchison, Jean. [2]1991. *Language Change: Progress or Decay?* Cambridge: Cambridge University Press.

Aitchison, Jean. [2]1994. *Words in the Mind: An Introduction to the Mental Lexicon*. Oxford: Blackwell.

Algeo, John. 1998. America Is Ruining the English Language. In: Laurie Bauer & Peter Trudgill (eds.). *Language Myths*. Harmondsworth: Penguin. 176-82.

Anderson, John M., Colin J. Ewen & Jørgen Staun. 1985. Phonological Structure: Segmental, Suprasegmental, and Extrasegmental. *Phonology* 2: 203-24.

Anderson, Stephen R. 1974. *The Organization of Phonology*. New York, NY: Academic Press.

Anderson, Stephen R. 1981. Why Phonology Isn't 'Natural'. *Linguistic Inquiry* 12: 493-539.

Anderson, Stephen R. 1985. *Phonology in the Twentieth Century: Theories of Rules and Theories of Representations*. Chicago, IL: University of Chicago Press.

Archangeli, Diana. 1988. Aspects of Underspecification Theory. *Phonology* 5: 183-207.

Archangeli, Diana. 1997. Optimality Theory: An Introduction to Linguistics in the 1990s. In: Diana Archangeli & D. Terence Langendoen (eds.). *Optimality Theory: An Overview*. Malden, MA: Blackwell. 1-32.

Archangeli, Diana & D. Terence Langendoen (eds.). 1997. *Optimality Theory: An Overview*. Malden, MA: Blackwell.

Archangeli, Diana & Douglas Pulleyblank. 1994. *Grounded Phonology*. Cambridge, MA: MIT Press.

Avery, Peter & Keren Rice. 1989. Segment Structure and Coronal Underspecification. *Phonology* 6: 179-200.

Bailey, Richard W. 1991. *Images of English: A Cultural History of the Language*. Cambridge: Cambridge University Press.

Baron, Dennis. 1990. *The English-Only Question: An Official Language for Americans*. New Haven, CT: Yale University Press.

Baudouin de Courtenay, Jan. 1895. An Attempt at a Theory of Phonetic Alternations. In: Edward Stankiewicz (ed.). 1972. *Selected Writings of Baudouin de Courtenay*. Bloomington, IN: Indiana University Press. 144-212.

Bauer, Laurie. 1988. What Is Lenition? *Journal of Linguistics* 24: 381-92.

Bauer, Laurie. 1994. *Watching English Change: An Introduction to the Study of Linguistic Change in Standard Englishes in the Twentieth Century*. London: Longman.

Bauer, Laurie, John M. Dienhart, Hans H. Hartvigson & Leif K. Jakobsen. 1980. *American English Pronunciation*. Copenhagen: Gyldendal.

Beckman, Jill N. 1999. Positional Faithfulness : *An Optimality Theoretic Treatment of Phonological Asymmetries*. New York, NY: Garland.

Bird, Steven & Ewan Klein. 1990. Phonological Events. *Journal of Linguistics* 26: 33-56.

Bloomfield, Leonard. 1933. *Language*. London: Allen & Unwin.

Bloomfield, Leonard. 1933. *Language*. New York, NY: Holt.

Blumstein, Sheila E. 1986. On Acoustic Invariance in Speech. In: Joseph S. Perkell & Dennis H. Klatt (eds.). *Invariance and Variability in Speech Processes*. Hillsdale, NJ: Erlbaum. 178-93.

Boersma, Paul. 1998. *Functional Phonology: Formalizing the Interactions Between Articulatory and Perceptual Drives*. PhD dissertation (Ms.). http://www.fon.hum.uva.nl/paul/diss/.

Booij, Geert & Jerzy Rubach. 1984. Morphological and Prosodic Domains in Lexical Phonology. *Phonology* 1: 1-27.

Bos, Egbert. 1992. William of Ockham. In: Marcelo Dascal, Dietfried Gerhardus, Kuno Lorenz & Georg Meggle (eds.). *Philosophy of Language*. Berlin: de Gruyter. 296-308.

Browman, Catherine P. & Louis M. Goldstein. 1985. Dynamic Modeling of Phonetic Structure. In: Victoria A. Fromkin (ed.). *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*. Orlando, FL: Academic Press. 35-53.

Browman, Catherine P. & Louis M. Goldstein. 1986. Towards an Articulatory Phonology. *Phonology* 3: 219-52.

Browman, Catherine P. & Louis M. Goldstein. 1989. Articulatory Gestures as Phonological Units. *Phonology* 6: 201-51.

Browman, Catherine P. & Louis M. Goldstein. 1990. Tiers in Articulatory Phonology, with Some Implications for Casual Speech. In: John Kingston & Mary E. Beckman (eds.). *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge: Cambridge University Press. 341-76.

Browman, Catherine P. & Louis M. Goldstein. 1992. Articulatory Phonology: An Overview. *Phonetica* 49: 155-80.

Bryson, Bill. 1995. *Made in America*. London: Minerva.

Bühler, Karl. [1934] 1982. *Sprachtheorie*. Stuttgart: UTB Gustav Fischer Verlag.

Bybee, Joan L. 1994. A View of Phonology from a Cognitive and Functional Perspective. *Cognitive Linguistics* 5: 285-305.

Byrd, Dani. 1992. Perception of Assimilation in Consonant Clusters: A Gestural Model. *Phonetica* 49: 1-24.

Byrd, Dani. 1993. 54.000 American Stops. *UCLA Working Papers in Phonetics* 83: 97-115.

Byrd, Dani. 1996. Influences on Articulatory Timing in Consonant Sequences. *Journal of Phonetics* 24: 209-44.

Byrd, Dani. 2003. Frontiers and Challenges in Articulatory Phonology. (Ms.). http://www-rcf.usc.edu/~dbyrd/byrd_CV.html.

Cameron, Deborah. 1995. *Verbal Hygiene*. London: Routledge.

Carr, Philip. 1990. *Linguistic Realities: An Autonomist Metatheory for the Generative Enterprise*. Cambridge: Cambridge University Press.

Carr, Philip. 1993. *Phonology*. Basingstoke: Macmillan.

Carver, Craig M. 1987. *American Regional Dialects: A Word Geography*. Ann Arbor, MI: University of Michigan Press.

Cassidy, Frederic G. 1983. Geographical Variation of English in the United States. In: Richard W. Bailey & Manfred Görlach (eds.). *English as a World Language*. Ann Arbor, MI: University of Michigan Press. 177-209.

Cassimjee, Farida & Charles W. Kisseberth. 1998. Tonal Variation Across Emakhuwa Dialects. (Ms.). ROA-289. Rutgers Optimality Archive: http://roa.rutgers.edu/.

Cassimjee, Farida & Charles W. Kisseberth. 1999. A Conspiracy Argument for Optimality Theory: Emakhuwa Dialectology. (Ms.). ROA-331. Rutgers Optimality Archive: http://roa.rutgers.edu/.

Catford, John C. 1977. *Fundamental Problems in Phonetics*. Edinburgh: Edinburgh University Press.

Catford, John C. 1988. *A Practical Introduction to Phonetics*. Oxford: Clarendon Press.

Chafe, Wallace L., John W. du Bois & Sandra Thompson. 1991. Towards a New Corpus of Spoken American English. In: Karin Aijmer & Bengt Altenberg (eds.). *English Corpus Linguistics*. Harlow: Longman. 64-82.

Chambers, John K. & Peter Trudgill. [2]1998. *Dialectology*. Cambridge: Cambridge University Press.

Cheshire, Jenny. 1991. The UK and the USA. In: Jenny Cheshire (ed.). *English Around the World: Sociolinguistic Perspectives*. Cambridge: Cambridge University Press. 13-34.

Cho, Young-Mee Yu. 1990. *Parameters of Consonantal Assimilation*. PhD dissertation (Ms.). Stanford, CA: Stanford University.

Cho, Young-Mee Yu. 1991. On the Universality of the Coronal Articulator. In: Carole Paradis & Jean-François Prunet (eds.). *Phonetics and Phonology* (vol. 2). *The Special Status of Coronals: Internal and External Evidence*. San Diego, CA: Academic Press. 159-79.

Chomsky, Noam. 1964. *Current Issues in Linguistic Theory*. The Hague: Mouton.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1970. Remarks on Nominalization. In: Roderick A. Jacobs & Peter S. Rosenbaum (eds.). *Readings in English Transformational Grammar*. Waltham, MA: Blaisdell. 184-221.

Chomsky, Noam. 1983. On Cognitive Structures and their Development: A Reply to Piaget. In: Massimo Piattelli-Palmarini (ed.). Language and Learning: *The Debate between Jean Piaget and Noam Chomsky*. London: Routledge. 35-54.

Chomsky, Noam. 1986. Changing Perspectives on Knowledge and Use of Knowledge. *Leuvense Bijdragen* 75: 1-71.

Chomsky, Noam. 1988a. *Language and Problems of Knowledge: The Managua Lectures*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1988b. Language and Problems of Knowledge. In: Aloysius P. Martinich (ed.). *The Philosophy of Language*. Oxford: Oxford University Press. 509-27.

Chomsky, Noam, Adriana Belletti & Luigi Rizzi. 2002. An Interview on Minimalism. In: *On Nature and Language*. Cambridge: Cambridge University Press. 92-161.

Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. Cambridge, MA: MIT Press.

Clark, John & Colin Yallop. [2]1995. *An Introduction to Phonetics and Phonology*. Oxford: Blackwell.

Clark, Thomas L. 1983. Linguistic Atlases: Archival Data for Language Study in the Western United States. *Journal of English Linguistics* 16: 13-7.

Clements, George N. 1985. The Geometry of Phonological Features. *Phonology* 2: 225-52.

Clements, George N. 1990. The Role of the Sonority Cycle in Core Syllabification. In: John Kingston & Mary E. Beckman (eds.). *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge: Cambridge University Press. 283-333.

Clements, George N. & Elisabeth V. Hume. 1995. The Internal Organization of Speech Sounds. In: John A. Goldsmith (ed.). *The Handbook of Phonological Theory*. Oxford: Blackwell. 245-306.

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

Cole, Jennifer & Charles Kisseberth. 1994. An Optimal Domains Theory of Harmony. (Ms.). ROA-22. Rutgers Optimality Archive: http://roa.rutgers.edu/.

Comrie, Bernard. 1987. Introduction. In: Bernard Comrie (ed.). *The World's Major Languages*. London: Croom Helm. 1-29.

Crowley, Tony. 1991. *Proper English? Readings in Language, History and Cultural Identity*. London: Routledge.

Cruttenden, Alan (ed.). [1962] ^{6}2001. *Gimson's Pronunciation of English*. London: Arnold.

Crystal, David. ^{2}1997. *The Cambridge Encyclopedia of Language*. Cambridge: Cambridge University Press.

Crystal, David. ^{5}2003. *A Dictionary of Linguistics and Phonetics*. Oxford: Blackwell.

Cutler, Anne. 1992. Psychology and the Segment. In: Gerard J. Docherty & D. Robert Ladd (eds.). *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*. Cambridge: Cambridge University Press. 290-5.

Davis, Stuart. 1995. Emphasis Spread in Arabic and Grounded Phonology. *Linguistic Inquiry* 26: 465-98.

Decamp, David. 1971. The Pronunciation of English in San Francisco. In: Juanita V. Williamson & Virginia M. Burke (eds.). *A Various Language: Perspectives on American Dialect*. New York, NY: Holt, Rinehart & Winston. 549-69.

Dekkers, Joost, Frank van der Leeuw & Jeroen van de Weijer (eds.). 2000. *Optimality Theory: Phonology, Syntax and Acquisition*. Oxford: Oxford University Press.

Dell, Gary S. 1995. Speaking and Misspeaking. In: Lila R. Gleitman & Mark Liberman (eds.). ^{2}1995. *An Invitation to Cognitive Science* (vol. 1): *Language*. Cambridge, MA: MIT Press. 183-208.

Denes, Peter B. 1963. On the Statistics of Spoken English. *Journal of the Acoustical Society of America* 35: 892-904.

Dillard, Joey L. 1992. *A History of American English*. London: Longman.

Donegan, Patricia J. & David Stampe. 1979. The Study of Natural Phonology. In: Daniel A. Dinnsen (ed.). *Current Approaches to Phonological Theory*. Bloomington, IN: Indiana University Press. 126-73.

Dressler, Wolfgang U. 1984. Explaining Natural Phonology. *Phonology* 1: 29-51.

Dressler, Wolfgang U. 1985. On the Predictiveness of Natural Morphology. *Journal of Linguistics* 21: 321-37.

Farnetani, Edda. 1997. Coarticulation and Connected Speech Processes. In: William J. Hardcastle & John Laver (eds.). *The Handbook of Phonetic Sciences*. Oxford: Blackwell. 371-404.

Fery, Caroline. 2002. *The Syllable in Optimality Theory*. Cambridge: Cambridge University Press. (dec. 2002)

Firth, John R. 1948. Sounds and Prosodies. In: Valerie B. Makkai (ed.). 1972. *Phonological Theory*. New York, NY: Holt, Rinehart & Winston. 252-63.

Flemming, Edward S. 1997. Phonetic Detail in Phonology: Towards a Unified Account of Assimilation and Coarticulation. (Ms.). http://www.stanford.edu/~flemming/#papers/ (12/17/01).

Flemming, Edward S. 2001. Scalar and Categorical Phenomena in a Unified Model of Phonetics and Phonology. *Phonology* 18: 7-44.

Flemming, Edward S. [1995] 2002. *Auditory Representations in Phonology*. New York, NY: Routledge.

Flexner, Stuart B. [2]1993. *Random House Unabridged Dictionary*. New York, NY: Random House.

Foley, James. 1972. Rule Precursors and Phonological Change by Meta-Rule. In: Robert P. Stockwell & Ronald K. S. Macaulay (eds.). *Linguistic Change and Generative Theory*. Bloomington, IN: Indiana University Press. 96-100.

Fromkin, Victoria A. 1971. The Non-Anomalous Nature of Anomalous Utterances. *Language* 47: 27-52.

Fromkin, Victoria, Robert Rodman & Nina Hyams. [7]2003. *An Introduction to Language*. Boston, MA: Thomson & Heinle.

Giegerich, Heinz J. 1992. *English Phonology: An Introduction*. Cambridge: Cambridge University Press.

Gleason, Henry A. 1961. *An Introduction to Descriptive Linguistics*. New York, NY: Holt, Rinehart & Winston.

Goldsmith, John A. 1976. *Autosegmental Phonology*. Bloomington, IN: Indiana University Linguistics Club.

Goldsmith, John A. 1979. The Aims of Autosegmental Phonology. In: Daniel A. Dinnsen (ed.). *Current Approaches to Phonological Theory*. Bloomington, IN: Indiana University Press. 202-22.

Goldsmith, John A. 1990. *Autosegmental and Metrical Phonology*. Oxford: Blackwell.

Goldsmith, John A. 1993. Harmonic Phonology. In: John A. Goldsmith (ed.). *The Last Phonological Rule: Reflections on Constraints and Derivations*. Chicago, IL: University of Chicago Press. 21-60.

Goldsmith, John A. 1995. Phonological Theory. In: John A. Goldsmith (ed.). *The Handbook of Phonological Theory*. Oxford: Blackwell. 1-23.

Goldstein, Louis M.. 1994. Possible Articulatory Bases for the Class of Guttural Consonants. In: Patricia Keating (ed.). *Papers in Laboratory Phonology III*. Cambridge: Cambridge University Press. 234-41.

Golston, Chris & Harry van der Hulst. 1999. Stricture is Structure. In: Ben Hermans & Marc Van Oostendorp (eds.). *The Derivational Residue in Phonological Optimality Theory*. Amsterdam: Benjamins. 153-73.

Golston, Chris & Richard Wiese. 1995. Zero Morphology and Constraint Interaction: Subtraction and Epenthesis in German Dialects. (Ms.). ROA-100. Rutgers Optimality Archive: http://roa.rutgers.edu/.

Gove, Philip B. 1961. *Webster's Third New International Dictionary*. Springfield, MA: Merriam.

Greenberg, Joseph H. 1963. Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In: Joseph H. Greenberg (ed.). [2]1966. *Universals of Language*. Cambridge, MA: MIT Press. 73-113.

Greenberg, Joseph H. 1976. *Language Universals with Special Reference to Feature Hierarchies*. The Hague: Mouton.

Greenberg, Joseph H. 1978. Some Generalizations Concerning Initial and Final Consonant Clusters. In: Joseph H Greenberg (ed.). *Universals of Human Language* (vol. 2): *Phonology*. Stanford, CA: Stanford University Press. 243-79.

Guy, Gregory R. 1980. Variation in the Group and the Individual: The Case of Final Stop Deletion. In: William Labov (ed.). 1980. *Locating Language in Time and Space*. New York, NY: Academic Press. 1-36.

Guy, Gregory R. 1997. Competence, Performance, and the Generative Grammar of Variation. In: Frans Hinskens, Roeland van der Hout & W. Leo Wetzels (eds.). *Variation, Change, and Phonological Theory*. Amsterdam: Benjamins. 125-43.

Haegeman, Liliane. [2]1994. *Introduction to Government and Binding Theory*. Oxford: Blackwell.

Haegeman, Liliane & Jacqueline Guéron. 1999. *English Grammar: A Generative Perspective*. Oxford: Blackwell.

Hale, Mark & Charles Reiss. 1998. Formal and Empirical Arguments Concerning Phonological Acquisition. *Linguistic Inquiry* 29: 656-83.

Hall, T. Alan. 1997. *The Phonology of Coronals*. Amsterdam: Benjamins.

Halle, Morris. 1961. On the Role of Simplicity in Linguistic Descriptions. In: Roman Jakobson (ed.). *Structure of Language and its Mathematical Aspects*. Providence, RI: American Mathematical Society. 89-94.

Halle, Morris. 1962. Phonology in Generative Grammar. In: Jerry A. Fodor & Jerrold J. Katz (eds.). 1964. *The Structure of Language: Readings in the Philosophy of Language*. Englewood Cliffs, NJ: Prentice-Hall. 334-52.

Halle, Morris. 1992. Phonological Features. In: William Bright (ed.). *International Encyclopedia of Linguistics* (vol. 3). Oxford: Oxford University Press. 207-12.

Halle, Morris. 1995. Feature Geometry and Feature Spreading. *Linguistic Inquiry* 26: 1-46.

Halle, Morris & George N. Clements. 1983. *Problem Book in Phonology*. Cambridge, MA: MIT Press.

Halliday, Michael A. K., Angus McIntosh & Peter Strevens. 1968. The Users and Uses of Language. In: Joshua A. Fishman (ed.). *Readings in the Sociology of Language*. The Hague: Mouton. 139-69.

Hammond, Michael. 1995. There Is No Lexicon. (Ms.). ROA-43. Rutgers Optimality Archive: http://roa.rutgers.edu/.

Hammond, Michael. 1997. Optimality Theory and Prosody. In: Diana Archangeli & D. Terence Langendoen (eds.). *Optimality Theory: An Overview*. Malden, MA: Blackwell. 33-58.

Hammond, Michael. 1999. *The Phonology of English: A Prosodic Optimality-Theoretic Approach*. Oxford: Oxford University Press.

Hardcastle, William J. 1976. *Physiology of Speech Production: An Introduction for Speech Scientists*. London: Academic Press.

Harris, John. 1994. *English Sound Structure*. Oxford: Blackwell.

Harris, Roy. 1990. The Dialect Myth. In: Jerold A. Edmondson, Crawford Feagin & Peter Mühlhäusler (eds.). *Development and Diversity: Language Variation Across Time and Space*. Arlington, TX: University of Texas. 3-20.

Harris, Zellig S. 1951. *Methods in Structural Linguistics*. Chicago, IL: University of Chicago Press.

Hartley, Laura C. & Dennis R. Preston. 1999. The Names of US English: Valley Girl, Cowboy, Yankee, Normal, Nasnal and Ignorant. In: Tony Bex & Richard

J. Watts (eds.). *Standard English: The Widening Debate*. London: Routledge. 207-38.

Hayes, Bruce. 1995. A Phonetically-Driven, Optimality-Theoretic Account of Postnasal Voicing. (Ms.). ROA-93. Rutgers Optimality Archive: http://roa.rutgers.edu/.

Hayes, Bruce. 1996. Phonetically Driven Phonology: The Role of Optimality Theory and Inductive Grounding. (Ms.). ROA-158. Rutgers Optimality Archive: http://roa.rutgers.edu/.

Hayes, Bruce. 1999. Phonological Acquisition in Optimality Theory: The Early Stages. (Ms.). ROA-327. Rutgers Optimality Archive: http://roa.rutgers.edu/.

Helt, Marie E. 2001. A Multi-Dimensional Comparison of British and American Spoken English. In: Susan Conrad & Douglas Biber (eds.). *Variation in English: Multi-Dimensional Studies*. London: Longman. 171-83.

Hermans, Ben & Marc Van Oostendorp (eds.). 1999. *The Derivational Residue in Phonological Optimality Theory*. Amsterdam: Benjamins.

Hill, Archibald A. 1962. Various Kinds of Phonemes. *Studies in Linguistics* 16: 3-10.

Hoenigswald, Henry M. 1966. A Proposal for the Study of Folk-Linguistics. In: William F. Bright (ed.). *Sociolinguistics: Proceedings of the UCLA Sociolinguistics Conference 1964*. The Hague: Mouton. 16-20.

Holton, David B. 1995. Assimilation and Dissimilation of Sundanese Liquids. In: Jill N. Beckman, Suzanne Urbanczyk & Laura Walsh-Dickey (eds.). *Papers in Optimality Theory*. Amherst, MA: University of Massachusetts Press. 167-80.

Hooper, Joan B. 1972. The Syllable in Phonological Theory. *Language* 48: 525-40.

Hooper, Joan B. 1976. *An Introduction to Natural Generative Phonology*. New York, NY: Academic Press.

Hooper, Joan B. 1979. Substantive Principles in Natural Generative Phonology. In: Daniel A. Dinnsen (ed.). *Current Approaches to Phonological Theory*. Bloomington, IN: Indiana University Press. 106-25.

Hume, Elisabeth V. & Keith Johnson. 2001. A Model of the Interplay of Speech Perception and Phonology. In: Elisabeth V. Hume & Keith Johnson (eds.). *The Role of Speech Perception in Phonology*. San Diego, CA: Academic Press. 3-26.

Hura, Susan L., Björn Lindblom & Randy L. Diehl. 1992. On the Role of Perception in Shaping Phonological Assimilation Rules. *Language and Speech* 35: 59-72.

Hyman, Larry M. 2001. The Limits of Phonetic Determinism in Phonology: *NC Revisited. In: Elisabeth V. Hume & Keith Johnson (eds.). *The Role of Speech Perception in Phonology*. San Diego, CA: Academic Press. 141-85.

International Phonetic Association. 1999. *Handbook of the IPA: A Guide to the Use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

Itō, Junko. 1989. A Prosodic Theory of Epenthesis. *Natural Language and Linguistic Theory* 7: 217-59.

Itō, Junko, Armin Mester & Jaye Padgett. 1995. Licensing and Underspecification in Optimality Theory. *Linguistic Inquiry* 26: 571-613.

Jackendoff, Ray. 1975. Morphological and Semantic Regularities in the Lexicon. *Language* 51: 639-71.

Jakobson, Roman. [1941] 1968. *Child Language, Aphasia and Phonological Universals*. The Hague: Mouton.

Jakobson, Roman. 1949. The Sound Laws of Child Language and their Place in General Phonology. In: Linda R. Waugh & Monique Monville-Burston (eds.).

1990. *On Language: Roman Jakobson*. Cambridge, MA: Harvard University Press. 294-304.

Jakobson, Roman. 1949. On the Identification of Phonemic Entities. In: Valerie B. Makkai (ed.). 1972. *Phonological Theory: Evolution and Current Practice*. New York, NY: Holt, Rinehart & Winston. 318-22.

Jakobson, Roman. 1979. Quest for the Ultimate Constituents. In: Linda R. Waugh & Monique Monville-Burston (eds.). 1990. *On Language: Roman Jakobson*. Cambridge, MA: Harvard University Press. 259-93.

Jakobson, Roman, Gunnar Fant & Morris Halle. 1951. *Preliminaries to Speech Analysis: The Distinctive Features and their Correlates*. Cambridge, MA: MIT Press.

Jakobson, Roman & Morris Halle. 1956. *Fundamentals of Language*. The Hague: Mouton.

Jespersen, Otto. 1922. 1922. *Language: Its Nature, Development and Origin*. London: Allen & Unwin.

Jones, Daniel. 1950. *The Phoneme: Its Nature and Use*. Cambridge: Heffer.

Jones, Daniel. [1918] [9]1960. *An Outline of English Phonetics*. Cambridge: Cambridge University Press.

Joshi, S. D. & Paul Kiparsky. 1979. Siddha and Asiddha in Pāninian Phonology. In: Daniel A. Dinnsen (ed.). *Current Approaches to Phonolgical Theory*. Bloomington, IN: Indiana University Press. 223-50.

Jun, Jongho. 1996. Place Assimilation as the Result of Conflicting Perceptual and Articulatory Constraints. In: José Camacho, Lisa Choueiri & Maki Watanabe (eds.). *The Proceedings of the Fourteenth West Coast Conference on Formal Linguistics*. Stanford, CA: CSLI Publications. 221-37.

Kager, René. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.

Kahn, Daniel. 1976. *Syllable-Based Generalizations in English Phonology*. Bloomington, IN: Indiana University Linguistics Club.

Kaisse, Ellen M. & Patricia A. Shaw. 1985. On the Theory of Lexical Phonology. *Phonology* 2: 1-30.

Kean, Mary-Louise. 1975. *The Theory of Markedness in Generative Grammar*. Bloomington, IN: Indiana University Linguistics Club.

Keating, Patricia A. 1985. Universal Phonetics and the Organization of Grammars. In: Victoria A. Fromkin (ed.). *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*. Orlando, FL: Academic Press. 115-32.

Keating, Patricia A. 1988. The Phonology–Phonetics Interface. In: Frederick J. Newmeyer (ed.). *Linguistics: The Cambridge Survey I: Linguistic Theory: Foundations*. Cambridge: Cambridge University Press. 281-302.

Keating, Patricia A. 1991. Coronal Places of Articulation. In: Carole Paradis & Jean-François Prunet (eds.). *Phonetics and Phonology* (vol. 2). *The Special Status of Coronals: Internal and External Evidence*. San Diego, CA: Harcourt Brace Jovanovich. 29-48.

Kenstowicz, Michael. 1981. Functional Explanations in Generative Phonology. In: Didier L. Goyvaerts (ed.). *Phonology in the 1980's*. Ghent: E. Story-Scientia. 431-44.

Kenstowicz, Michael. 1994. *Phonology in Generative Grammar*. Oxford: Blackwell.

Kenstowicz, Michael & Charles W. Kisseberth. 1979. *Generative Phonology: Description and Theory*. Orlando, FL: Academic Press.

Kenyon, John S. [12]1994. *American Pronunciation*. Ann Arbor, MI: Wahr.

Kenyon, John S. & Thomas A. Knott. [4]1953. *A Pronouncing Dictionary of American English*. Springfield, MA: Merriam.

Kiparsky, Paul. 1968. Linguistic Universals and Linguistic Change. In: Emmon Bach & Robert T. Harms (eds.). *Universals in Linguistic Theory*. New York, NY: Holt, Rinehart & Winston. 171-202.

Kiparsky, Paul. 1973. "Elsewhere" in Phonology. In: Stephen R. Anderson & Paul Kiparsky (eds.). *A Festschrift for Morris Halle*. New York, NY: Holt, Rinehart & Winston. 93-106.

Kiparsky, Paul. 1982. From Cyclic Phonology to Lexical Phonology. In: Harry van der Hulst & Norval Smith (eds.). *The Structure of Phonological Representations* (vol. 1). Dordrecht: Foris. 131-75.

Kiparsky, Paul. 1985. Some Consequences of Lexical Phonology. *Phonology* 2: 85-138.

Kiparsky, Paul. 1995. The Phonological Basis of Sound Change. In: John A. Goldsmith (ed.). *The Handbook of Phonological Theory*. Oxford: Blackwell. 640-70.

Kirchner, Robert M. 1998. *An Effort-Based Approach to Consonant Lenition*. PhD dissertation (Ms.). ROA-276. Rutgers Optimality Archive: http://roa.rutgers.edu/.

Kisseberth, Charles W. 1970. On the Functional Unity of Phonological Rules. *Linguistic Inquiry* 1: 291-306.

Kohler, Klaus J. 1991. The Phonetics/Phonology Issue in the Study of Articulatory Reduction. *Phonetica* 48: 180-92.

Kohler, Klaus J. 1992. Gestural Reorganization in Connected Speech: A Functional Viewpoint on 'Articulatory Phonology'. *Phonetica* 49: 205-11.

Krapp, George P. 1925. *The English Language in America* (2 vols.). New York, NY: Ungar.

Labov, William. 1972a. *Language in the Inner City: Studies in the Black English Vernacular*. Philadelphia, PA: University of Pennsylvania Press.

Labov, William. 1972b. *Sociolinguistic Patterns*. Philadelphia, PA: University of Pennsylvania Press.

Labov, William. 1975. *What is a Linguistic Fact*? Lisse: Peter de Ridder Press.

Labov, William. 1991. The Three Dialects of English. In: Penelope Eckert (ed.). *New Ways of Analyzing Sound Change: Quantitative Analysis of Linguistic Structure*. New York, NY: Academic Press. 1-44.

Labov, William. 1994. *Principles of Linguistic Change: Internal Factors*. Cambridge, MA: Blackwell.

Labov, William, Sharon Ash & Charles Boberg. 1997. *A National Map of the Regional Dialects of American English*. 15 July 1997 <http://www.ling.upenn.edu/phono_atlas/NationalMap/NatMap1.html>.

LaCharité, Darlene & Carole Paradis. 1993. The Emergence of Constraints in Generative Phonology and a Comparison of Three Current Constraint-Based Models. *Canadian Journal of Linguistics* 38: 127-53.

LaCharité, Darlene & Carole Paradis. 2000. Derivational Residue: Hidden Rules in Optimality Theory. In: Joost Dekkers, Frank van der Leeuw & Jeroen van de Weijer (eds.). *Optimality Theory: Phonology, Syntax and Acquisition*. Oxford: Oxford University Press. 211-33.

Ladefoged, Peter. 1971. *Preliminaries to Linguistic Phonetics*. Chicago: University of Chicago Press.

Ladefoged, Peter. 1980. What Are Linguistic Sounds Made of? *Language* 56: 485-502.

Ladefoged, Peter. 1997. Linguistic Phonetic Descriptions. In: William J. Hardcastle & John Laver (eds.). *The Handbook of Phonetic Sciences*. Oxford: Blackwell. 589-618.

Ladefoged, Peter & Ian Maddieson. 1988. Phonological Features for Places of Articulation. In: Larry M. Hyman & Charles N. Li (eds.). *Language, Speech and Mind: Studies in Honour of Victoria A. Fromkin*. London: Routledge. 49-61.

Ladefoged, Peter & Ian Maddieson. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell.

Lass, Roger. 1975. How Intrinsic is Content? Markedness, Sound Change, and 'Family Universals'. In: Didier L. Goyvaerts & Geoffrey K. Pullum (eds.). *Essays on the Sound Pattern of English*. Ghent: Story-Scientia. 475-504.

Lass, Roger. 1980. *On Explaining Language Change*. Cambridge: Cambridge University Press.

Lass, Roger. 1984. *Phonology: An Introduction to Basic Concepts*. Cambridge: Cambridge University Press.

Laver, John. 1994. *Principles of Phonetics*. Cambridge: Cambridge University Press.

Leben, William R. [1973] 1980. *Suprasegmental Phonology*. New York, NY: Garland.

Lehmann, Winfred P. 1953. A Note on the Change of American English /t/. *American Speech* 28: 271-5.

Lenk, Uta. 1998. *Marking Discourse Coherence: Functions of Discourse Markers in Spoken English*. Tübingen: Narr.

Lesser, Ruth & Lesley Milroy. 1993. *Linguistics and Aphasia: Psycholinguistic and Pragmatic Aspects of Intervention*. London: Longman.

Liberman, Mark & Alan Prince. 1977. On Stress and Liguistic Rhythm. *Linguistic Inquiry* 8: 249-336.

Lindau, Mona & Peter Ladefoged. 1986. Variability of Feature Specifications. In: Joseph S. Perkell & Dennis H. Klatt (eds.). *Invariance and Variability in Speech Processes*. Hillsdale, NJ: Erlbaum. 426-49.

Lindblom, Björn. 1986. Phonetic Universals in Vowel Systems. In: John J. Ohala & Jeri J. Jaeger (eds.). *Experimental Phonology*. Orlando, FL: Academic Press. 13-44.

Lindblom, Björn. 1992. Phonological Units as Adaptive Emergents of Lexical Development. In: Charles A. Ferguson, Lise Menn & Carol Stoel-Gammon (eds.). *Phonological Development: Models, Research, Implications*. Timonium, MD: York Press. 131-63.

Lindblom, Björn & Ian Maddieson. 1988. Phonetic Universals in Consonant Systems. In: Larry M. Hyman & Charles N. Li (eds.). *Language, Speech and Mind: Studies in Honour of Victoria A. Fromkin*. London: Routledge. 62-78.

Lippi-Green, Rosina. 1997. *English with an Accent: Language, Ideology and Discrimination in the United States*. London: Routledge.

Lombardi, Linda. 1999. Positional Faithfulness and Voicing Assimilation in Optimality Theory. *Natural Language & Linguistic Theory* 17: 267-302.

Lombardi, Linda. 2001a. Introduction. In: Linda Lombardi (ed.). *Segmental Phonology in Optimality Theory: Constraints and Representations*. Cambridge: Cambridge University Press. 1-9.

Lombardi, Linda. 2001b. Why Place and Voice are Different: Constraint-Specific Alternations in Optimality Theory. In: Linda Lombardi (ed.). *Segmental Phonology in Optimality Theory: Constraints and Representations*. Cambridge: Cambridge University Press. 13-45.

Lombardi, Linda 2003. Coronal Epenthesis and Markedness. (Ms.). ROA-579. Rutgers Optimality Archive: http://roa.rutgers.edu/.

Maddieson, Ian. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.

Martinet, André. [2]1964. *Economie des Changements Phonétiques*. Berne: A. Francke.

Martinet, André. 1968. Phonetics and Linguistic Evolution. In: Bertil Malmberg (ed.). *Manual of Phonetics*. Amsterdam: North-Holland Publishing. 464-87.

McArthur, Tom. 1998. *The English Languages*. Cambridge: Cambridge University Press.

McCarthy, John J. 1984. Prosodic Organization in Morphology. In: Mark Aronoff & Richard T. Oerle (eds.). *Language Sound Structure*. Cambridge, MA: MIT Press. 299-317.

McCarthy, John J. 1988. Feature Geometry and Dependency: A Review. *Phonetica* 43: 84-108.

McCarthy, John J. 1993. A Case of Surface Constraint Violation. *Canadian Journal of Linguistics* 38: 169-95.

McCarthy, John J. 1997. Process-Specific Constraints in Optimality Theory. *Linguistic Inquiry* 28: 231-51.

McCarthy, John J. 2002. *A Thematic Guide to Optimality Theory*. Cambridge: Cambridge University Press.

McCarthy, John J. & Alan Prince. 1993. Generalized Alignment. In: Geert Booij & Jaap van Marle (eds.). *Yearbook of Morphology 1993*. Dordrecht: Kluwer. 79-153.

McCarthy, John J. & Alan Prince. 1994. The Emergence of the Unmarked: Optimality in Prosodic Morphology. (Ms.). ROA-13. Rutgers Optimality Archive: http://roa.rutgers.edu/.

McCarthy, John J. & Alan Prince. 1995a. Prosodic Morphology. In: John A. Goldsmith (ed.). *The Handbook of Phonological Theory*. Oxford: Blackwell. 318-66.

McCarthy, John J. & Alan Prince. 1995b. Faithfulness and Reduplicative Identity. (Ms.). ROA-60. Rutgers Optimality Archive: http://roa.rutgers.edu/.

McCarthy, John J. & Alison Taub. 1992. Review of Carole Paradis & Jean-François Prunet (eds.). 1991. The Special Status of Coronals. *Phonology* 9: 363-70.

McEnery, Tony & Andrew Wilson. 1996. *Corpus Linguistics*. Edinburgh: Edinburgh University Press.

McMahon, April. 2000. *Change, Chance, and Optimality*. Oxford: Oxford University Press.

Mencken, Henry L. [4]1963. *The American Language*. New York, NY: Knopf.

Mester, Armin & Junko Itō. 1989. Feature Predictability and Underspecification. *Language* 65: 258-93.

Milroy, James. 1992. *Linguistic Variation and Change*. Oxford: Blackwell.

Milroy, James. 1998. Children Can't Speak or Write Properly Any More. In: Laurie Bauer & Peter Trudgill (eds.). *Language Myths*. Harmondsworth: Penguin. 58-65.

Milroy, Lesley. 1999. Standard English and Language Ideology in Britain and the United States. In: Tony Bex & Richard J. Watts (eds.). *Standard English: The Widening Debate*. London: Routledge. 173-206.

Milroy, James & Lesley Milroy. [3]1999. *Authority in Language: Investigating Standard English*. London: Routledge.

Mohanan, Karuvannur P. 1986. *The Theory of Lexical Phonology*. Dordrecht: Reidel.

Mohanan, Karuvannur P. 1991. On the Bases of Radical Underspecification. *Natural Language and Linguistic Theory* 9: 285-325.

Mohanan, Karuvannur P. 1993. Fields of Attraction in Phonology. In: John A. Goldsmith (ed.). *The last Phonological Rule: Reflections on Constraints and Derivations*. Chicago, IL: University of Chicago Press. 61-116.

Moonwomon, Birch. 1992. *Sound Change in San Francisco English*. PhD dissertation (Ms.). Berkeley, CA: University of California.

Moren, Bruce. 1999. *Distinctiveness, Coercion, and Sonority: A Unified Theory of Weight*. (Ph.D. Diss.). ROA-346. Rutgers Optimality Archive: http://roa.rutgers.edu/.

Murray, Lindley. [1795] 1968. *English Grammar*. Menston: Scolar Press.

Murray, Stephen O. 1998. *American Sociolinguistics: Theorists and Theory Groups*. Amsterdam: Benjamins.

Nespor, Marina & Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris.

Neu, Helene. 1980. Ranking of Constraints on /t,d/ Deletion in American English: A Statistical Analysis. In: William Labov (ed.). 1980. *Locating Language in Time and Space*. New York, NY: Academic Press. 37-54.

Newman, Rochelle, James Sawusch & Paul Luce. 2000. Underspecification and Phoneme Frequency in Speech Perception. In: Michael B. Broe Janet B. Pierrehumbert (eds.). *Papers in Laboratory Phonology V: Acquisition and the Lexicon*. Cambridge: Cambridge University Press. 299-312.

Nolan, Francis. 1992. The Descriptive Role of Segments: Evidence from Assimilation. In: Gerard J. Docherty & D. Robert Ladd (eds.). *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*. Cambridge: Cambridge University Press. 261-80.

Odden, David. 1995. Tone: African Languages. In: John A. Goldsmith (ed.). *The Handbook of Phonological Theory*. Oxford: Blackwell. 444-75.

O'Grady, William, John Archibald, Mark Aronoff & Janie Rees-Miller. [4]2001. *Contemporary Linguistics: An Introduction*. New York: St. Martin's Press.

Ohala, John J. 1990. The Phonetics and Phonology of Aspects of Assimilation. In: John Kingston & Mary E. Beckman (eds.). *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge: Cambridge University Press. 258-75.

Ohala, John J. 1995. The Perceptual Basis of Some Sound Patterns. In: Bruce Connell & Amalia Arvaniti (eds.). *Papers in Laboratory Phonology IV*. Cambridge: Cambridge University Press. 87-92.

Ohala, John J. 1997. The Relation Between Phonetics and Phonology. In: William J. Hardcastle & John Laver (eds.). *The Handbook of Phonetic Sciences*. Oxford: Blackwell. 674-94.

Ohala, John J. & Haruko Kawasaki. 1984. Prosodic Phonology and Phonetics. *Phonology* 1: 113-27.

Oostendorp, Marc van & Ben Hermans. 1999. Optimality Theory and Derivational Effects. In: Ben Hermans & Marc van Oostendorp (eds.). *The Derivational Residue in Phonological Optimality Theory*. Amsterdam: Benjamins. 1-27.

Oswald, Victor A. 1943. 'Voiced T'—A Misnomer. *American Speech* 18: 18-25.

Padgett, Jaye. 1995a. *Stricture in Feature Geometry*. Stanford, CA: CSLI Publications.

Padgett, Jaye. 1995b. Feature Classes. In: Jill N. Beckman, Suzanne Urbanczyk & Laura Walsh-Dickey (eds.). *Papers in Optimality Theory*. Amherst, MA: University of Massachusetts Press. 385-420.

Padgett, Jaye. 2002. Constraint Conjunction versus Grounded Constraint Subhierarchies in Optimality Theory. (Ms.). ROA-530. Rutgers Optimality Archive: http://roa.rutgers.edu/.

Paradis, Carole. 1988. On Constraints and Repair Strategies. *The Linguistic Review* 6: 71-97.

Paradis, Carole & Jean-François Prunet. 1989. On Coronal Transparency. *Phonology* 6: 317-48.

Paradis, Carole & Jean-François Prunet. 1991. Introduction: Asymmetry and Visibility in Consonant Articulations. In: Carole Paradis & Jean-François Prunet (eds.). *Phonetics and Phonology* (vol. 2). *The Special Status of Coronals: Internal and External Evidence*. San Diego, CA: Academic Press. 1-28.

Passy, Paul. 1890. *Etude sur les Changements Phonétiques et leurs Caractères Généraux*. Paris: Firmin-Didot.

Pater, Joe. 1995. Austronesian Nasal Substitution and Other NC Effects. (Ms.). ROA-92. Rutgers Optimality Archive: http://roa.rutgers.edu/.

Pater, Joe. 1998. From Phonological Typology to the Development of Receptive and Productive Phonological Competence: Applications of Minimal Violations (Ms.). ROA-296. Rutgers Optimality Archive: http://roa.rutgers.edu/.

Pike, Kenneth L. 1943. *Phonetics: A Critical Analysis of Phonetic Theory and a Technic for the Practical Description of Sounds*. Ann Arbor, MI: University of Michigan Press.

Pike, Kenneth L. 1947. *Phonemics: A Technique for Reducing Languages to Writing*. Ann Arbor, MI: University of Michigan Press.

Popper, Karl R. [2]1968. *The Logic of Scientific Discovery*. London: Hutchinson.

Popper, Karl R. [5]1973. *Logik der Forschung*. Tübingen: Mohr.

Postal, Paul M. 1968. *Aspects of Phonological Theory*. New York, NY: Harper & Row.

Preston, Dennis R. 1993. Folk Dialectology. In: Dennis R. Preston (ed.). *American Dialect Research*. Amsterdam: Benjamins. 333-77.

Preston, Dennis R. 1998. They Speak Really Bad English Down South and in New York City. In: Laurie Bauer & Peter Trudgill (eds.). *Language Myths*. Harmondsworth: Penguin. 139-49.

Prince, Alan & Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. New Brunswick: Rutgers University. (Ms.).

Pukui, Mary K. & Samuel H. Elbert. [6]1986. *Hawaiian Dictionary*. Honolulu, HI: University of Hawai'i Press.

Pulleyblank, Douglas. 1997. Optimality Theory and Features. In: Diana Archangeli & D. Terence Langendoen (eds.). *Optimality Theory: An Overview*. Malden, MA: Blackwell. 59-101.

Quine, Willard V. O. 1953. Two Dogmas of Empiricism. In: *From a Logical Point of View*. Cambridge, MA: Harvard University Press. 20-46.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Roach, Peter. 1983. *English Phonetics and Phonology*. Cambridge: Cambridge University Press.

Robins, Robert H. 1957. Aspects of Prosodic Analysis. In: Valerie B. Makkai (ed.). 1972. *Phonological Theory*. New York, NY: Holt, Rinehart & Winston. 264-74.

Romaine, Suzanne (ed.). 1998. *The Cambridge History of the English Language. Volume IV: 1776-1997*. Cambridge: Cambridge University Press.

Rose, Yvan. 2001. *Headedness and Prosodic Licensing in the L1 Acquisition of Phonology*. (Ph.D. Diss.). ROA-475. Rutgers Optimality Archive: http://roa.rutgers.edu/.

Russell, Kevin. 1995. Morphemes and Candidates in Optimality Theory. (Ms.). ROA-44. Rutgers Optimality Archive: http://roa.rutgers.edu/.

Sagey, Elisabeth. 1986. *The Representation of Features in Non-Linear Phonology: The Articulator Node Hierarchy*. New York, NY: Garland.

Sapir, Edward. 1921. *Language: An Introduction to the Study of Speech*. San Diego, CA: Harcourt Brace & Company.

Saussure, Ferdinand de. 1959. *Course in General Linguistics*. New York, NY: Philosophical Library.

Schachter, Paul. 1969. Natural Assimilation Rules in Akan. *International Journal of American Linguistics* 35: 342-355.

Schane, Sanford A. 1971. The Phoneme Revisited. *Language* 47: 503-21.

Schane, Sanford A. 1972. Natural Rules in Phonology. In: Robert P. Stockwell & Ronald K. S. Macaulay (eds.). *Linguistic Change and Generative Theory*. Bloomington, IN: Indiana University Press. 199-229.

Schiffrin, Deborah. 1994. *Approaches to Discourse*. Oxford: Blackwell.

Schilling-Estes, Natalie. 1995. Extending Our Understanding of the /z/ → [d] Rule. *American Speech* 70: 291-302.

Schneider, Edgar W. 1996. Research Trends in the Study of American English. In: Edgar W. Schneider (ed.). *Varieties of English Around the World: Focus on the USA*. Amsterdam: Benjamins. 1-12.

Schütze, Carson T. 1996. *The Empirical Base of Linguistics: Grammaticality Judgements and Linguistic Methodology*. Chicago, IL: University of Chicago Press.

Selkirk, Elisabeth O. 1982. The Syllable. In: John A. Goldsmith (ed.). 1999. *Phonological Theory: The Essential Readings*. Malden, MA: Blackwell. 328-50.

Selkirk, Elisabeth O. 1984a. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.

Selkirk, Elisabeth O. 1984b. On the Major Class Features and Syllable Theory. In: Mark Aronoff & Richard T. Oehrle (eds.). *Language Sound Structure*. Cambridge, MA: MIT Press. 107-36.

Sharf, Donald J. 1960. Distinctiveness of 'Voiced T' Words. *American Speech* 35: 105-9.

Sherrard, Nicholas. 1997. Questions of Priorities. In: Iggy Roca (ed.). *Derivations and Constraints in Phonology*. Oxford: Clarendon Press. 43-89.

Shibatani, Masayoshi. 1973. The Role of Surface Phonetic Constraints in Generative Phonlogy. *Language* 49: 87-106.

Sievers, Eduard. [1881] [5]1901. *Grundzüge der Phonetik*. Leipzig: Breitkopf & Härtel.

Singh, Rajendra. 1988. In Defense of External Evidence. *Canadian Journal of Linguistics* 33: 329-43.

Slater, Rosalie J. (ed.). [1828] [4]1985. *Noah Webster's First Edition of an American Dictionary of the English Language*. San Francisco, CA: Foundation for American Christian Education.

Smolensky, Paul. 1995. On the Structure of the Constraint Component CON of UG. (Ms.). ROA-86. Rutgers Optimality Archive: http://roa.rutgers.edu/.

Smolensky, Paul. 1996a. The Initial State and 'Richness of the Base' in Optimality Theory. (Ms.). ROA-154. Rutgers Optimality Archive: http://roa.rutgers.edu/.

Smolensky, Paul. 1996b. On the Comprehension/Production Dilemma in Child Language. *Linguistc Inquiry* 27: 720-31.

Soames, Laura. [2]1899. *Introduction to English, French and German Phonetics*. London: Swan Sonnenschein.

Spencer, Andrew. 1996. *Phonology: Theory and Description*. Oxford: Blackwell.

Stampe, David. 1973. On Chapter Nine. In: Michael J. Kenstowicz & Charles W. Kisseberth (eds.). *Issues in Phonological Theory*. The Hague: Mouton. 44-52.

Stampe, David. 1979. *A Dissertation on Natural Phonology*. New York, NY: Garland.

Stanley, Richard. 1967. Redundancy Rules in Phonology. *Language* 43: 393-436.

Steriade, Donca. 1995. Underspecification and Markedness. In: John A. Goldsmith (ed.). *The Handbook of Phonological Theory*. Oxford: Blackwell. 114-74.

Steriade, Donca. 2000. Paradigm Uniformity and the Phonetics–Phonology Boundary. In: Michael B. Broe Janet B. Pierrehumbert (eds.). *Papers in Laboratory Phonology V: Acquisition and the Lexicon*. Cambridge: Cambridge University Press. 313-33.

Stevens, Kenneth N. 1985. Evidence for the Role of Acoustic Boundaries in the Perception of Speech Sounds. In: Victoria A. Fromkin (ed.). *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*. Orlando, FL: Academic Press. 243-55.

Stevens, Kenneth N. 1994. Phonetic Evidence for Hierarchies of Features. In: Patricia Keating (ed.). *Papers in Laboratory Phonology III*. Cambridge: Cambridge University Press. 242-57.

Stevens, Kenneth N. 1997. Articulatory–Acoustic–Auditory Relationships. In: William J. Hardcastle & John Laver (eds.). *The Handbook of Phonetic Sciences*. Oxford: Blackwell. 462-506.

Stevens, Kenneth N. 1998. *Acoustic Phonetics*. Cambridge, MA: MIT Press.

Stevens, Kenneth N. & Sheila E. Blumstein 1981. The Search for Invariant Acoustic Correlates of Phonetic Features. In: Peter D. Eimas & Joanne L. Miller (eds.). *Perspectives on the Study of Speech*. Hillsdale, NJ: Erlbaum. 1-38.

Stevens, Kenneth N. & Samuel J. Keyser. 1989. Primary Features and their Enhancement in Consonants. *Language* 65: 81-106.

Suzuki, Keiichirō. 1998. *A Typological Investigation of Dissimilation*. PhD dissertation (Ms.). ROA-281. Rutgers Optimality Archive: http://roa.rutgers.edu/.

Sweet, Henry. 1888. *A History of English Sounds*. Oxford: Clarendon Press.

Taylor, John R. [2]1995. *Linguistic Categorization: Prototypes in Linguistic Theory*. Oxford: Oxford University Press.

Tesar, Bruce & Paul Smolensky. 1993. The Learnability of Optimality Theory: An Algorithm and Some Basic Complexity Results. (Ms.). ROA-2. Rutgers Optimality Archive: http://roa.rutgers.edu/.

Tesar, Bruce & Paul Smolensky. 2000. *Learnability in Optimality Theory*. Cambridge, MA: MIT Press.

Thomas, Linda & Shân Wareing (eds.). 1999. *Language, Power and Society: An Introduction*. London: Routledge.

Tognini-Bonelli, Elena. 2001. *Corpus Linguistics at Work*. Amsterdam: Benjamins.

Toon, Thomas E. 1983. Variation in Contemporary American English. In: Richard W. Bailey & Manfred Görlach (eds.). *English as a World Language*. Ann Arbor, MI: University of Michigan Press. 210-50.

Trager, George L. 1930. The Pronunciation of 'Short *A*' in American Standard English. *American Speech* 5: 396-400.

Trager, George L. 1942. The Phoneme 'T': A Study in Theory and Method. *American Speech* 17: 144-8.

Trubetzkoy, Nikolaj S. [1939] ^{4}1958. *Grundzüge der Phonologie*. Göttingen: Vandenhoeck & Ruprecht.

Trudgill, Peter. ^{3}1995. *Sociolinguistics: An Introduction to Language and Society*. Harmondsworth: Penguin.

Trudgill, Peter. 1999. Standard English: What it Isn't. In: Tony Bex & Richard J. Watts (eds.). *Standard English: The Widening Debate*. London: Routledge. 117-28.

Trudgill, Peter & Jean Hannah. ^{3}1994. *International English: A Guide to the Varieties of Standard English*. London: Arnold.

Twaddell, W. Freeman. 1935. On Defining the Phoneme. In: Martin Joos (ed.). ^{2}1958. *Readings in Linguistics: The Development of Descriptive Linguistics in America Since 1925*. New York, NY: American Council of Learned Societies. 55-80.

Wang, William S-Y. & John Crawford. 1960. Frequency Studies of English Consonants. *Language and Speech* 3: 131-9.

Webster, Noah. [1789] 1951. *Dissertations on the English Language*. Gainesville, FL: Scholars' Facsimiles & Reprints.

Wells, John C. 1982. *Accents of English* (3 vols.). Cambridge: Cambridge University Press.

Wells, John C. 1990. *The Longman Pronouncing Dictionary*. London: Longman.

Werker, Janet F. ^{2}1995. Exploring Developmental Changes in Cross-Language Speech Perception. In: Lila Gleitman & Mark Liberman. *An Invitation to Cognitive Science: Language*. 87-106.

Westbury, John R. & Patricia A. Keating. 1986. On the Naturalness of Stop Consonant Voicing. *Journal of Linguistics* 22: 145-66.

Whaley, Lindsay J. 1997. *Introduction to Typology: The Unity and Diversity of Language*. London: Sage.

Whitney, William D. [1868] 1971. *Language and the Study of Language*. New York, NY: AMS Press.

Whitney, William D. ^{2}1889. *Sanskrit Grammar: Including both the Classical Language, and the Older Dialects, of Veda and Brahmana*. Cambridge, MA: Harvard University Press.

Wilson, Kenneth G. 1993. *The Columbia Guide to Standard American English*. New York, NY: Columbia University Press. (http://www.bartleby.com/68/35/1935.html)

Wolfram, Walt. 1991. *Dialects and American English*. Englewood Cliffs, NJ: Prentice Hall.

Wolfram, Walt, Carolyn T. Adger & Donna Christian. 1999. *Dialects in Schools and Communities*. Mahwah, NJ: Erlbaum.

Wolfram, Walt & Ralph W. Fasold. 1974. *The Study of Social Dialects in American English*. Englewood Cliffs, NJ: Prentice Hall.

Wolfram, Walt & Robert Johnson. 1982. *Phonological Analysis: Focus on American English*. Washington, DC: Center for Applied Linguistics.

Wolfram, Walt & Natalie Schilling-Estes. 1998. *American English: Dialects and Variation*. Oxford: Blackwell.

Yip, Moira. 1988. The Obligatory Contour Principle and Phonological Rules: A Loss of Identity. *Linguistic Inquiry* 19: 65-100.

Yip, Moira. 1989. Feature Geometry and Cooccurrence Restrictions. *Phonology* 6: 349-74.

Yip, Moira. 1991. Coronals, Consonant Clusters, and the Coda Condition. In: Carole Paradis & Jean-François Prunet (eds.). *Phonetics and Phonology* (vol. 2). *The Special Status of Coronals: Internal and External Evidence*. San Diego, CA: Harcourt Brace Jovanovich. 61-78.

Zipf, George K. 1949. *Human Behavior and the Principle of Least Effort*. Cambridge, MA: Addison-Wesley.

Zwicky, Arnold M. 1972a. Note on a Phonological Hierarchy in English. In: Robert P. Stockwell & Ronald K. S. Macaulay (eds.). *Linguistic Change and Generative Theory*. Bloomington, IN: Indiana University Press. 275-301.

Zwicky, Arnold M. 1972b. The Strategy of Generative Phonology. In: Wolfgang U. Dressler & František V. Mareš (eds.). 1975. *Phonologica 1972*. München: Fink. 151-68.