Chapter 3

BASIC THEORETICAL ELEMENTS AND THEIR PERCEPTUAL MOTIVATIONS

The preceding chapter identified a number of empirical generalizations, which condition the application of consonant deletion, vowel epenthesis, and vowel deletion. These output generalizations are summarized below.

Generalization 1: Consonants want to be adjacent to a vowel, and preferably followed by a vowel.

Generalization 2: Stops want to be adjacent to a vowel, and preferably followed by a vowel.

Generalization 3: Stops that are not followed by a [+continuant] segment want to be adjacent to a vowel, and preferably followed by a vowel.

Generalization 4: Consonants that are relatively similar to a neighboring segment want to be adjacent to a vowel, and preferably followed by a vowel.

Generalization 5: Consonants that are not at the edge of a prosodic domain want to be adjacent to a vowel, and preferably followed by a vowel.

Generalization 6: Coronal stops want to be followed by a vowel.

The likelihood that a consonant deletes or triggers vowel epenthesis correlates with the degree to which it is subject to these constraints. Likewise, the likelihood that a consonant blocks vowel deletion correlates with the degree to which it would be subject to these constraints if deletion applied.

I argue that these generalizations have a perceptual motivation and follow from a general principle of perceptual salience:

(1) PRINCIPLE OF PERCEPTUAL SALIENCE:
All segments are perceptually salient.

The perceptual salience of a segment – or its degree of confusability with zero – is a function of the quantity and quality of the auditory cues that signal its presence in the speech stream. The best cues to consonants, apart from those present in the consonants themselves, are found in neighboring vowels, especially in the CV transition. It is the desirability for consonants to benefit from these vocalic cues that generalization 1 expresses. But cues may also come from other sources, and the perceptibility of a consonant without the support of an adjacent or following vowel depends on these non-(pre)vocalic cues. Generalizations 2-6 identify factors that negatively affect these cues, and consequently enhance the desirability of an adjacent vowel in order to meet the principle in (1).

I assume that consonant deletion and vowel epenthesis are motivated by the principle of perceptual salience; they apply when a consonant lacks perceptual salience and becomes more easily confusable with nothing, that is when the cues that permit a listener to detect its presence are diminished. Deletion removes such deficient segments, epenthesis provides them with the needed additional salience. Likewise, vowel deletion is blocked when it would leave a consonant with a reduced salience. Maintaining the vowel avoids removing cues that are crucial to that consonant. The link between vowel epenthesis and increase in salience has been investigated for Dutch by Donselaar et al. (1999). Dutch has an optional process of epenthesis in word-final consonant clusters, e.g. the word *film* is pronounced [film] or [filəm]. Donselaar et al. find that lexical access is significantly faster when the epenthetic vowel is present than when it is not. They argue that this is due to the increased salience or perceptibility that the vowel provides to its surrounding consonants, a finding that is supported by a phoneme-detection experiment in the last section of the paper.

I hypothesize that there is a direct relation between the perceptibility scale of consonants and the likelihood that they delete, trigger vowel epenthesis, or block vowel deletion. In other words, the likelihood that a certain consonant deletes, triggers epenthesis, or block vowel deletion correlates with the quality and quantity of the auditory cues associated to it in a given context.

I propose that the principle of perceptual salience is encoded in the grammar by means of markedness and faithfulness constraints that militate against consonants that lack auditory salience. These perceptually-motivated constraints are projected from observable phonetic properties in the course of acquisition (Hayes 1999; Steriade 1999d). The analysis is cast in Optimality Theory (Prince & Smolensky

1993; for recent overviews of the theory, see Archangeli & Langendoen 1998 and Kager 1999).

In this chapter I present the phonetic motivations that underlie the six generalizations above (3.1) and develop a constraint system that derives these generalizations and yields the desired patterns of consonant deletion, vowel epenthesis, and vowel deletion. I argue that both markedness and faithfulness constraints encode perceptual factors. I also discuss a number of issues that this perceptually-motivated analysis raises, notably the role of phonetics and perception in synchronic phonology and the treatment of variation in Optimality Theory. I end the chapter with two case studies that I use to illustrate the functioning of the constraint system I propose. Lenakel epenthesis introduces the role of markedness constraints, whereas consonant deletion in Sranan highlights that of the perceptually-based faithfulness constraints.

3.1. Perceptual motivations

I argue that the generalizations observed in patterns of consonant deletion, vowel epenthesis, and vowel deletion have a perceptual motivation: less salient consonants are more likely to delete, trigger vowel epenthesis, or block vowel deletion. The identification of consonants relies on a number of acoustic cues, which can be grouped into two categories: internal cues produced during the closure part of the consonant, and contextual cues that originate from neighboring segments. In addition, an important cue to stops is their release burst, which can be thought of as sharing characteristics of both internal and contextual cues: the burst is an inherent part of the production of stops, which relates it to internal cues, but its audibility depends on the nature of the following segment, like contextual cues. (See Wright 1996 for a summary of available cues to consonants' place and manner of articulation).

The whole system rests on the privileged status of CV transitions. Consonants are optimally salient before a vowel, and non-optimally salient in any position that lacks these transitions. Whether or not non-optimal consonants are tolerated depends on the quality and quantity of their non-CV cues and the language-specific degree of tolerance for less salient consonants. The six generalizations presented at the outset of this chapter are elucidated in terms of internal cues, contextual cues, modulation in the acoustic signal, and cue enhancement at edges of prosodic domains.

3.1.1. CV and VC transitions

Generalization 1: Consonants want to be adjacent to a vowel, and preferably followed by a vowel.

The first generalization – consonants want to be adjacent to a vowel, and preferably followed by a vowel – stems from the major role played by vocalic transitions in the perception of consonants, and the dominance of the CV transitions over the VC ones. Formant transitions from and to adjacent vowels provide optimal contextual cues to consonants because of their high amplitude and dynamic pattern which gives information about the changing configuration of the vocal tract. They provide cues to all aspects of the articulation of consonants: manner, place, and laryngeal settings. This explains why consonants want to be adjacent to at least one vowel (VC or CV). The significance of these transitions for the perception of consonants is summarized as follows by Delattre (1961/1966: 407):

> Les transitions de formants jouent, dans la perception de la parole, un rôle autrement plus important que ne le laisserait entendre le choix peu heureux du terme "transition". Au lieu d'être une phase secondaire ou négligeable, comme on l'a longtemps cru, les transitions sont à la clef même de la perception de la consonne.

There is, however, a significant difference between VC and CV transitions. An important body of research points to the privileged status of CV sequences, as opposed to VC ones (e.g. Fujimura et al. 1978; Ohala & Kawasaki 1985; Ohala 1990, 1992; Sussman et al. 1997; Dogil 1999; Joanisse 1999; Krakow 1999; Warner 1999). Everything else being equal, consonants have better contextual cues in prevocalic than in postvocalic position. The relative weakness of postvocalic cues certainly constitutes the main factor involved in one of the most firmly established generalizations in phonology: the general preference for consonants to appear in onset rather than in coda position. It also provides an explanation for the asymmetrical behavior of several deletion, weakening, debuccalization, or assimilation processes in phonology, which typically target postvocalic consonants and VC sequences.[1]

[1] The asymmetry between CV and VC could also explain statistical patterns in CVC words in English. Kessler & Treiman (1997) analyzed the distribution of phonemes in 2001 CVC English words. They found a significant connection between the vowel and the following consonant – certain vowel-coda combinations being more frequent than expected by chance – but no associations between the initial consonant and the vowel.

The perceptual advantage of CV transitions over VC ones is reflected in a number of experimental results. First, perceptual experiments have shown that when faced with contradictory transitions from the preceding and the following vowels in a VCV context, listeners mainly rely on the CV ones (Fujimura et al. 1978; Ohala 1990). Consonants are also identified much more rapidly with CV cues than VC ones (Warner 1999).

What is the source of this asymmetry? A number of differences between CV and VC sequences have been established, which all point to the enhanced perceptibility of prevocalic consonants. Ohman (1966) and Kawasaki (1982) have shown that VC formant transitions for different consonants are not as spectrally well differentiated among themselves as CV transitions. It follows that consonants are better contrasted with each other in prevocalic than in postvocalic position. We also know that the onset of a stimulus signal has a greater impact on the auditory system that its offset. It gives rise to a marked burst of activity of the auditory nerve fiber (see Wright 1996). This holds for linguistic stimuli as well, and provides a perceptual advantage to post-consonantal transitions cues: the onset of formants (those at the CV juncture) are amplified in a way that their offset (those at the VC juncture) are not. In addition, stop release bursts, an important cue to stops, occur in CV but not necessarily in VC contexts.

The auditory advantage of CV transitions seems to be reinforced by the articulatory patterns in CV vs. VC sequences. This research is reviewed by Krakow (1999) and provides consistent results.[2] First, there is more coarticulation or overlap between a consonant and a preceding vowel than between it and a following vowel. In other words, there is a more precise timing of articulatory movements in CV sequences. For example, velic lowering in [m] occurs earlier with respect to the onset of the labial constriction in postvocalic than in prevocalic position; in CV sequences both gestures are synchronized. Therefore, the nasality of the consonant spreads to the preceding vowel more than to the following one. Likewise for laterals, which involve both a tongue dorsum and a tongue tip articulation (in English): it has been observed that the tongue dorsum raises earlier with respect to the tongue tip in VC than in CV contexts. Second, prevocalic consonants have a more extreme

[2]Krakow (1999) nicely summarizes the coarticulation results. She presents her results in syllabic terms – coda vs. onset consonants – and interprets the coarticulatory differences between them as reflecting syllabic organization. Notice, however, that the data used to derive these results never contrast only in syllable structure: they can all be described in terms of prevocalic vs. postvocalic consonants and domain-internal vs. domain-edge consonants. To the extent that reference to larger domains is necessary anyway – and this is clear in numerous studies cited by Krakow – the role of the syllable becomes unclear. The syllable could be a perceptual side-effect of the articulatory organization, not its origin (see Ohala 1992).

consonantal articulation than postvocalic ones. They are produced with a tighter constriction; for example, postvocalic laterals show a weaker constriction between the tongue and the palate than prevocalic ones. Nasals are also more sonorant in postvocalic position in that they are associated with a lower velic position and longer low velic plateaus. These two articulatory properties have an increasing effect on the amount of nasal airflow, making postvocalic nasals indeed more sonorant-like or less obstruent-like than prevocalic nasals.

Increased constriction and reduced coarticulation both enhance the contrast between the consonant and the following vowel. They maximize the alternation between a closed consonantal constriction and an open vocalic articulation; they also keep the two segments more distinct by reducing the overlap between them. Although the precise perceptual effects of these articulatory properties need further investigation, one expects a correlation between the maximization of the articulatory and acoustic contrast between the consonant and the following vowel. This in turn positively affects the perception of the segments involved, since their salience is largely determined by the degree of modulation in the acoustic signal (see section 3.1.4).

3.1.2. INTERNAL CUES AND THE GREATER VULNERABILITY OF STOPS

Generalization 2: Stops want to be adjacent to a vowel, and preferably followed by a vowel.

The second generalization states that stops, more than other consonants, need an adjacent vowel, preferably a following one. The greater tendency of stops to delete, trigger epenthesis, or block deletion stems from the weakness of their non-CV cues. Consonants that lack the cues present in the CV transition have to rely more on other cues, which happen to be weaker for stops. Stops then suffer more than other consonants from not appearing in prevocalic position.

The special status of stops stems from two elements: the weakness of their internal cues and the audibility of their release burst. Stops have weak or no internal cues produced during closure. Due to the absence of oral or nasal airflow, this part of the segment is silent or associated only with low-amplitude vocal fold vibrations, and provide very weak (internal) cues.[3] The non-internal perceptual cues to stops are rather concentrated in their release burst, whose importance in the perception of

[3]Voiced stops are often not accompanied by vocal fold activity and the corresponding voicing bar, especially in postvocalic position. Periodicity in the signal therefore does not constitute a reliable cue to voiced stops (Wright 1996; Steriade 1999c).

stops has often been reported (see numerous references in Wright 1996: 5 and Clark & Yallop 1995: 282). But non-prevocalic stops do not reliably benefit from an audible release burst, as noted in the previous section, and the absence or weakness of the burst may severely reduce their salience and perceptibility. Thus the disadvantage of VC cues against CV ones is amplified in the case of stops as opposed to other consonants.

By contrast, nasals, fricatives, and liquids have relatively robust internal cues. Fricatives have frication noise, sonorants have formant structure. So they remain perceptible even in the absence of transition cues. The contrast between segments with and without internal cues (stops vs. other consonants) is not only apparent in deletion and epenthesis processes. It also affects the articulatory timing in the production of consonant clusters. Wright (1996) studied in detail the production of word-initial and word-internal consonant clusters in Tsou. He noticed that stops that lack transitional cues are produced in such a way as to maintain an audible release burst, which implies a smaller degree of overlap with the following consonant. Other consonants – those with internal cues – in the same context, however, overlap more with adjacent consonants, presumably because their internal cues are salient enough. To maintain a sufficient degree of perceptibility in the absence of flanking vowels, a stop thus tends to involve more articulatory energy.[4]

A distinction should be made, however, between strident and non-strident fricatives with respect to internal cues. Non-strident fricatives are associated with noise of low amplitude, often not detectable on normal spectrograms. Miller & Nicely (1955) show that the distinction between stops and the weak fricatives becomes unreliable in masking noise. This distinction is indeed reflected in deletion patterns, which further supports the perceptual basis of deletion processes. The historical loss of non-strident fricatives is common, but [s] and [ʃ] are generally more resistant. Non-strident fricatives may pattern with other fricatives with respect to deletion / epenthesis (the more common case in this dissertation) or with stops. The Icelandic pattern reviewed in chapter 1 provides just one example of the latter situation. I will not, however, discuss the behavior of non-strident fricatives in this dissertation, focussing only on stops.

[4]See Rhee (1998) for a discussion of the role of release in various phonological patterns.

3.1.3. The audibility of release bursts

Generalization 3: Stops that are not followed by a [+continuant] segment want to be adjacent to a vowel, and preferably followed by a vowel.

The role of the [continuancy] value of the following element on stop deletion can be related to the audibility of the release burst. There is a well-known tendency for stops to be unreleased or to lack an audible release in certain contexts. Based on Henderson & Repp (1982), we can usefully distinguish between stops with and without a release that has an observable effect in the acoustic signal. Stops without an acoustically present release actually comprise two distinct types: strictly unreleased and silently-released stops. Articulatorily unreleased stops occur before homorganic nasal or oral stops and utterance-finally. In the first case the constriction is maintained through the following consonant; utterance-finally it may be delayed. Silently-released stops are found before an oral or nasal stop with a more front articulation. When the closure of the second consonant is made before the release of the first stop, this release has no acoustic effect since the air is trapped behind the front constriction (see also Laver 1994: 359-360).

Unreleased and silently-released stops, however, are not found if the stop is followed by a segment that does not involve a complete closure in the oral cavity, since there is always an outgoing flow of air that can carry the effect of the release. Such segments correspond to the class defined by the specification [+continuant]. We can therefore establish a basic opposition between [+continuant] segments and the rest ([-continuant] segments and final position) with respect to the acoustic effect of a preceding stop release: it is necessarily present when the stop is followed by a [+continuant] segment. Since the release burst plays an important role in the perception of stops, it is advantageous to ensure that the release will not be devoid of an acoustic effect; being followed by a [+continuant] segment is one way to achieve this goal.[5]

[5]It must be noticed, however, that a release burst may be acoustically present but so weak that it is not perceived or not reliably perceived by listeners. As is made clear in Henderson & Repp (1982), a binary opposition between "released" and "unreleased" stops is insufficient and potentially misleading: the audibility of an acoustically present release is a gradual phenomenon, which ranges from inaudible to very salient, with various intermediate cases. This depends on various aspects of the segmental and prosodic context and on the articulatory timing. The basic opposition between the absence and presence of an acoustic effect of the release must be supplemented by additional factors that determine its level of perceptibility, but I do not carry out this task here.

3.1.4. CONTRAST AND MODULATION IN THE ACOUSTIC SIGNAL

Generalization 4: Consonants that are relatively similar to a neighboring segment want to be adjacent to a vowel, and preferably followed by a vowel.

The role of similarity or contrast in combinations of segments is explained by the correlation between the amount of acoustic modulation in a sound sequence and its perceptual salience (e.g. Kawasaki 1982; Ohala & Kawasaki 1985; Wright 1996; Boersma 1998). The auditory system gets rapidly "bored" or "numbed" and is little responsive to continuous stimuli. It therefore needs constant variation and the greater the modulation, the greater the salience, the more easily perceptible the segments involved are. Modulation is measured in terms of "the magnitude, rate and the number of stimulus parameters varying simultaneously" (Ohala & Kawasaki 1985: 116). Factors involved in the computation of modulation include differences in sound intensity or amplitude and variation in the spectrum. More specifically we may look at formant frequency, relative formant amplitude, overall spectral energy, and periodicity in the signal.

The necessity of modulation for perception is not specific to linguistic signals. Analogies with other perceptual systems are easy to find. Boersma (1998) uses a cartographical metaphor: in a country map, adjacent countries have to be represented in distinct colors if they are to be easily recognized as different entities. More generally, the production of modulations in some carrier signal can be viewed as "the essence of any communication channel" (Ohala & Kawasaki 1985: 123).

In predicting and explaining phonotactic patterns, however, modulation interacts with many other factors, in particular articulation, the way the perceptual system responds to certain properties of the acoustic signal, and the risk of confusability between different sound sequences that are acoustically similar. But we can hypothesize that, everything else being equal, sound combinations displaying a greater modulation in a given dimension are perceptually better, and are predicted to be more common, than other sequences with a smaller modulation in the same dimension. Likewise, sequences containing modulation in a larger number of dimensions are preferable to sequences with modulation in fewer dimensions. This can be transposed in featural terms, to the extent that features are associated with some acoustic contrast: a segment that contrasts in *n* features with its neighboring segments is more perceptible than a segment that contrasts in *n-1* features (again, everything else being equal). This will be the rationale of the constraint system developed below.

The role of acoustic modulation in explaining the crosslinguistic frequency of certain phonotactic patterns and combinations of segments has been investigated in particular by Kawasaki (1982) and Kawasaki-Fukumori (1992). She explored the following sequences: stop-liquid, stop-glide, stop-vowel, and vowel-stop. The hypothesis tested was whether the relative rarity of certain combinations within these groups could be motivated by acoustic/auditory constraints, in particular the lack of acoustic modulation within the sequence. The disfavored combinations are assumed to be:
- dental stop + /l/
- labial consonant + /w/
- alveolar-palatal consonant + /j/
- sequences of a labial or labialized consonant and a rounded vowel
- sequences of an alveolar/palatal or palatalized consonant and a front vowel

In addition, CV sequences are generally prefered to VC ones.

To test this hypothesis, selected CLV, CGV, CV, and VC sequences were recorded. The most influential parameter in acoustic modulation was taken to be the changes in the frequencies of the first three formants. The salience of a given sequence was approximated by the sum of the distance in frequency of these formants.

The results support the hypothesis to a large extent. Labial consonant + /w/ and alveolar-palatal consonant + /j/ clusters show little spectral modulation. This is also true of sequences of a labial or labialized consonant and a rounded vowel and sequences of an alveolar/palatal or palatalized consonant and a front vowel. The relative markedness of these combinations is therefore compatible with a perceptually-based motivation. In general, as noted in section 3.1.1, VC syllables are also spectrally closer among themselves than CV syllables, so consonants are better contrasted with each other in prevocalic than in postvocalic position, in accordance with Ohman's (1966) results.

The case of dental stop + /l/ clusters is not explained by the acoustic modulation hypothesis. In general, we observe more modulation in stop+/r/ than in stop+/l/ clusters, which is compatible with stop+/r/ sequences being less restricted crosslinguistically than stop+/l/. But if we look at stops with different points of articulation, we see that the clusters of a stop and a liquid show the least spectral change when the initial stop is bilabial and the greatest modulation in formant frequencies in /d/+liquid. This is unexpected and the modulation hypothesis clearly fails to predict the avoidance of /dl/ sequences in languages of the world. I do not have a reasonable alternative to propose and only notice that formant trajectories are not the only determinant of salience and that other perceptual factors may be involved, notably the release burst and the general dispreference for alveolar stops in nonprevocalic position (see generalization 6, in section 3.1.6).

Janson (1986), however, contests the validity of Kawasaki's generalizations concerning CV sequences, specifically the dispreference for sequences of a labial or labialized consonant and a rounded vowel and alveolar/palatal or palatalized consonant and a front vowel. By looking at a sample of five unrelated or distantly related languages, Janson actually reaches opposite conclusions: the favored sequences are alveolar consonant+front vowel and labial consonant+back rounded vowel. He suggests that these tendencies are to be explained by articulatory factors: the prefered CV sequences are those that require smaller articulatory movements. Kawasaki's generalizations, then, would hold only for /w/ and labialized consonants + rounded vowels and /j/ and palatalized consonants + front vowels. These sequences are indeed disprefered and acoustic/auditory lack of modulation is probably the relevant factor.

Janson's statistical results, however, were reanalyzed by Maddieson & Precoda (1992), who ended up with no clear trend in any direction. They found no preference or dispreference for specific CV combinations, with two salient exceptions: sequences of a glide followed by the corresponding vowel and velar consonants before high front vowels. The first probably follows from Kawasaki's modulation hypothesis, the second from articulatory considerations. What can we conclude from these results? It may well be the case that the frequency of CV sequences is relatively uninfluenced by phonetic factors of the kind Kawasaki and Janson have proposed. But this conclusion, I believe, does not extend to contexts other than CV. I would like to suggest that CV sequences, with the exception of combinations such as /wu/ and /ji/, all generally involve large spectral modulation. Their perceptibility may be beyond the level found desirable in most languages, and the distinctions in spectral change found between different CV combinations may become largely irrelevant. In other words, CV sequences are all good enough and speakers/listeners may not prize additional modulation high.

In this dissertation I am concerned with combinations of consonants, which generally show less modulation than CV sequences. I suggest that differences in amplitude and spectral variations here play a decisive role and may really determine the fate of particular sequences. It is in these less prefered segment combinations that the impact of auditory similarity is likely to reveal itself. I believe the patterns described here support this idea.

3.1.5. CUE ENHANCEMENT AT EDGES OF PROSODIC DOMAINS

Generalization 5: Consonants that are not at the edge of a prosodic domain want to be adjacent to a vowel, and preferably followed by a vowel.

The salience of consonants depends upon their position in the prosodic structure. It is by now well-established that segments at edges of prosodic constituents, from the word to the utterance, are associated with processes that enhance their salience. Specifically, edge consonants benefit from articulatory strengthening, lengthening, and reduction in the amount of overlap with the segment across the boundary, processes that are assumed to increase their perceptibility. Studies that have investigated these processes include: Oller (1973); Klatt (1975, 1976); Cooper & Danly (1981); Beckman & Edwards (1990); Wightman et al. (1992); Byrd (1994); Fougeron & Keating (1996, 1997); Gordon (1997); Keating et al. (1998); Fougeron (1999); Turk (1999); Byrd et al. (2000).

Consonants at the right and left edges behave differently; both edges benefit from cue enhancement, but through different processes. The right edge is mainly associated with segment lengthening, but is not characterized, or only marginally so, by articulatory strengthening. By contrast, the left edge involves articulatory strengthening (e.g. tighter constriction), with lengthening apparently playing a secondary role in that position. Reduction of overlap across prosodic boundaries is obviously symmetrical since it affects the final segment of the first constituent and the initial one of the following constituent. It has also been established that these effects are cumulative as we go up the prosodic hierarchy; that is, we observe more initial strengthening, final lengthening, and reduction of overlap at higher boundaries than lower ones.

There are only a handful of studies of gestural overlap between segments separated by different levels of junctures. I refer to Byrd et al. (2000) for a summary of these studies, which "suggest that phrasal position is a significant force in constraining the degree of temporal overlap between articulatory gestures."

Studies that confirm domain-final lengthening are numerous, e.g. Oller (1973), Klatt (1975, 1976), Cooper & Danly (1981), Beckman and Edwards (1990), Wightman et al. (1992), Turk (1999), and additional sources cited in the last two references. See also Edwards et al. (1990) and Beckman et al. (1992) for the articulatory mechanisms involved in final lengthening. Turk (1999) establishes that final lengthening targets predominantly the coda, that is the last consonant(s), which is lengthened in phrase-final position in her corpus by almost 200%. The preceding nucleus vowel is also lengthened, but to a much lesser extent (around 65%), while the onset of the domain-final syllable is only marginally affected (around 12%).

Stops, however, contrast with other consonants. My own analysis of a corpus very similar to that used by Turk (1999) suggests that lengthening affects stops much less than other consonants.[6] This corpus allows us to directly compare the words Duke /duk/ and Maine /men/ in phrase-final and phrase-medial position. For phrase-final /men/, we observe an increase in duration of about 155% for the coda /n/ vs. 59% for the preceding nucleus. These numbers are comparable to those provided by Turk, but they contrast dramatically to those obtained for phrase-final /duk/. In this case, the nucleus /u/ lengthens relatively more than the coda /k/: 104.5% vs. 32.2%.[7] This confirms Klatt's (1976: 1213) observation that stops tend not to lengthen as much as other consonants at phrase boundaries. This may be related to the fact that maintaining a stop closure for a longer period of time demands relatively more effort than maintaining the constriction for other consonants. In utterance-final position, Cooper & Danly (1981) found that the percentage of lengthening for alveolar and labiodental fricatives in English ranges from 79% for /v/ to 167% for /s/, that is also substantially more than what I found for stops. This is not to say that stops are not affected as much as other consonants in phrase-final position: I rather believe that the main difference for them lies in the strength and audibility of their release burst (see below) more than in their lengthening.

Wightman et al. (1992) is the most detailed study of the correlation between the amount of lengthening and the strength of the following boundary. They use seven different break indices or boundaries, with increasing strength from 0 to 6. A break index of 0 is assigned between two orthographic words where no prosodic break is perceived, the break index 6 marks sentence boundaries. Intermediate break indices can variably be related to other prosodic units cited in the literature (prosodic word, accentual phrase, intermediate phrase, intonational phrase, etc.), but no exact correspondence is established (see the discussion on p. 1710). The amount of lengthening for a segment is expressed in terms of normalized duration, which is

[6]This analysis was performed on a corpus provided by Stefanie Shattuck-Hufnagel as part of the course "Laboratory in the physiology, acoustic and perception of speech" taught at MIT by Ken Stevens, Joe Perkell, and Stefanie Shattuck-Hufnagel in the fall of 1999.

[7]It is interesting to observe, though, that the increase in the rime phrase-finally is very similar for both words: 73.8% for *Maine* and 68.1% for *Duke*. This suggests that phrase-final lengthening primarily targets the rime, and that there are compensation effects between the nucleus and the coda depending on the lengthenability of the coda consonant. The distribution of the increase in duration within the rime apparently tends to concentrate on the coda consonant, unless it is a stop. In this case, the nucleus carries most of the lengthening.

a measure of deviation from an expected value, taken to be 0.[8] They find that domain-final consonants are longer and longer as we go from a break index 0 to a break index 5. The strongest index 6 does not involve any additional lengthening with respect to the immediately preceding level. As we will see again, the absence of a contrast between the end of the utterance and the end of the immediately preceding level (standardly the Intonational Phrase or IP) is a recurrent result of the phonetic studies of edge segments.[9] The average normalized durations of consonants, depending on the level of the following break index (0 to 5), are given below. These numbers are approximations taken from the first graph in figure 4 (p. 1714). By contrast, Wightman et al. (1992) found no correlation between the duration of domain-initial consonants and the size of the preceding boundary.

(2) CONSONANT DURATION IN DOMAIN-FINAL POSITION:

Break index:	0	1	2	3	4	5	6
Normalized duration of the final consonant:	-0.5	-0.2	-0.1	0.2	0.5	0.85	0.6

Fougeron & Keating (1997) also report an effect of the phrase-final position on articulation, in an experiment involving reiterant speech with /no/ syllables: phrase-final vowels are more open than phrase-medial ones. This result was interpreted in terms of strengthening, since openness for vowels indicates a more extreme articulation. But they found no correlation between the degree of openness and the strength of the following boundary: final /o/'s above the word level are simply always quite open, irrrespective of the strength of the boundary. Thus there is no cumulative effect, unlike in final lengthening. More importantly, no similar strengthening has been reported for consonants, which most particularly concern us here.

Articulatory strengthening in initial position is a recent area of investigation, studied in particular in Pierrehumbert & Talkin (1992); Dilley et al. (1996); Fougeron & Keating (1996, 1997); Gordon (1997); Keating et al. (1998); Fougeron (1999); Byrd et al. (2000). Strengthening manifests itself differently in different classes of consonants, but it can be viewed as always resulting in a more consonant-like articulation, that is less sonorant and/or involving a tighter constriction.

[8]A negative normalized duration means that the segment is shorter than average; a positive one means that the segment is longer than average.

[9]If we interpret lengthening as a cue to prosodic boundaries, we may think that additional lengthening in the case of the utterance is unnecessary since other more salient cues are available, notably pauses.

Pierrehumbert & Talkin (1992) found that initial /h/ is more consonant-like when it is phrase-initial than when it is phrase-medial, the degree of consonantality being measured by the amount of breathiness and the corresponding degree of glottal opening. Similar results were obtained for the glottal stop. Glottalization of word-initial vowels was further investigated by Dilley et al. (1996), who found that it is more frequent at the beginning of large prosodic constituents (Intonational Phrase) than at the beginning of lower domains (Intermediate Phrase), and least likely phrase-medially. These findings are interpreted in terms of strengthening, greater gestural magnitude and increase in consonantality associated with the onset of prosodically significant domains.

Fougeron & Keating (1996, 1997), Gordon (1997), Keating et al. (1998), and Fougeron (1999) are concerned with linguopalatal contact and/or nasal flow in initial oral and/or nasal alveolar stops in various domains, from the word to the utterance. These studies consistently establish a correlation between the strength of the boundary preceding the consonant and the amount of linguopalatal contact, measured by the number of electrodes contacted on an artificial palate in EPG experiments. The identity and, to a lesser extent the number, of the prosodic domains that can be consistently distinguished by the amount of contact varies from speaker to speaker, but the general trend is invariant. As in the lengthening data presented in (2), the Utterance is not generally distinguished from the Intonational Phrase. I use the French data analyzed in Fougeron & Keating (1996) as an example. I report below for their two speakers the percentage of electrodes contacted in the production of /t/ and /n/ at the beginning of syllables (word-internal), words, accentual phrases, intonational phrases, and utterances (approximated from the graphs in figure 4).

(3) CONSONANTAL CONSTRICTION IN DOMAIN-INITIAL POSITION:
Average maxima of linguopalatal contact for /t/ and /n/ at the left edge of increasingly strong prosodic domains (from Fougeron & Keating 1996):

		Syllable	Word	AP	IP	U
Speaker 1	/n/	40	44	49	56	57
	/t/	51	54	56	60	62
Speaker 2	/n/	47	52	58	68	67
	/t/	54	55	63	69	66

Similar results are obtained for the amount of nasal airflow: nasals at the left edge of higher constituents are associated to a reduced amount of nasal airflow in comparison to nasals at the beginning of lower domains or in domain-internal position. Again, this is interpreted as an increase in consonantality. But the

correlation with boundary strength is not as good as that obtained with linguopalatal contact, which appears to be more directly influenced by the prosodic position. The maxima of nasal flow in /n/ depending on the prosodic position for speaker 1 above are given below. The underlined numbers indicate the levels that are significantly distinguished by the amount of nasal flow, the other two not following the expected trend, although this is not surprising in the case of the utterance. (The other speaker had less consistent results, which differed with the identity of the adjacent vowels; they are not shown here.)

(4) NASAL AIRFLOW IN DOMAIN-INITIAL POSITION:
Average maxima of nasal flow (in ml/sec) for /n/ at the left edge of increasingly strong prosodic domains (from Fougeron & Keating 1996):

	Syllable	Word	AP	IP	U
Speaker 1	48	<u>69</u>	<u>60</u>	<u>47</u>	59

Finally, a word should be said about lengthening in initial position of prosodic domains. Although certainly less prevalent than in constituent-final position, lengthening of initial consonants is reported in a number of studies, e.g. Oller (1973) and Pierrehumbert & Talkin (1992). In their detailed study of segmental durations at edges of prosodic domains, however, Wightman et al. (1992) found no correlation between the length of the initial consonant and the strength of the preceding boundary. Just like final strengthening, which was found to occur indistinctively in final positions above the word level, there could be a process of initial strengthening which affects all phrase-initial segments, irrespective of the level of the juncture.

The linguistic significance of these phonetic processes affecting edges of prosodic constituents – articulatory strengthening, lengthening, and reduction of overlap – is not yet entirely clear. We may think that they help with the segmentation of the signal into words and higher constituents, by signalling the presence of prosodic boundaries and providing cues to their strength (see Fougeron & Keating 1997). It seems clear that segment lengthening may be used by listeners to locate prosodic boundaries. Wightman et al. (1992) have shown that the degree of final lengthening enables listeners to distinguish at least 4 levels of prosodic domains. Strengthening and overlap reduction result in an enhanced contrast between the initial consonant and the adjacent segments. This enhancement process could also be interpreted by listeners as indicating the presence of a boundary. The amount of strengthening or contrast could even provide cues as to the strength of the boundary. Perceptual experiments are necessary, however, to assess the extent to which listeners use these phonetic variations for segmentation purposes.

Fougeron & Keating (1997) also suggest that initial strengthening may play a facilitating role in lexical access. It enhances the contrast between the initial segment and its neighbors. This increases the accessibility of segmental information in this position, which is welcome since initial segments are important in word recognition.

I would like to suggest a third area in which the phonetic correlates of domain-final and domain-initial positions impact the linguistic system: consonant licensing. Lengthening, increased articulatory energy, and less overlap enhance the salience of domain edges, and conspire to license more complex segments, a greater number of segments, and a wider variety of consonants in these positions. A strengthened and lengthened articulation correlates with more robust auditory cues, and those cues are not susceptible to weakening through overlap with a following segment. Stops and affricates are likely to particularly benefit from those effects, which facilitate the production of more strongly released bursts and increase their audibility through reduction of overlap. Since the burst constitutes an important element in the perception of these segments, we may think that the addition of the cues associated to it results in a radical shift upward in their perceptibility. In contrast, the effects of strengthening or lengthening may affect less radically the perceptibility of consonants other than stops and affricates, which does not so much depend on the release cues.

Since we observe a correlation between lengthening, strengthening, overlap, and the strength of the adjacent boundary, I predict that consonants are more easily licensed at edges of higher prosodic constituents than at edges of lower ones. This is indeed what we find in Hungarian degemination and the French schwa. Additional cases will be presented in chapter 5. Segments in word-internal position are not followed by any (relevant) prosodic boundary. Therefore they do not benefit at all from the advantages associated with domain edges, which explains their increased tendency to delete, trigger vowel epenthesis, and block vowel deletion.

3.1.6. CORONAL STOPS AND F2 TRANSITIONS

Generalization 6: Coronal stops want to be followed by a vowel.

Our last generalization, illustrated by deletion and assimilation in Attic Greek (chapter 1), concerns coronal stops, which contrast with other stops in being particularly disfavored in non-prevocalic position. This issue has been addressed in a recent paper by Y. Kang (1999), who provides a perceptual explanation for the specific behavior of coronal stops. I rely entirely on her treatment in this section. The Attic Greek (and Latin) pattern was used to illustrate the shortcomings and the syllabic approach to deletion and epenthesis, and Kang's explanation supports the perceptual alternative I advocate in this dissertation. This will exhaust what I have to say about the peculiarities of coronal stops.

In many languages coronal stops are more subject to deletion and assimilation than other stops in preconsonantal position. This is unexpected in view of the relative unmarkedness of coronals with respect to other places of articulation. Kang's explanation for this tendency is based on the role of F2 transitions in the perception of coronality and their distinct properties in prevocalic and postvocalic position. An important auditory cue to coronality lies in the F2 transitions. While F2 transitions from a coronal consonant to a following vowel (CV) are robust and clear, those from a vowel to a coronal (VC) are considerably weakened, almost nonexistent. There is little movement in F2 in the final 20 ms of the vowel. This acoustic fact is interpreted as the result of a weakening in the tongue body gesture, which plays a large part in shaping the F2 transition. This articulatory weakening makes coronals particularly vulnerable in (unreleased) preconsonantal position and subject to masking by the following consonant. Citing Byrd (1992) and Zsiga (1994) (see also Surprenant & Goldstein 1998), Kang notes that in $V_1C_1C_2V_2$ sequences, where C_1 is coronal, produced with extensive overlap between the two consonants, the vowel V_1 carries the cues to C_2 rather than those to the coronal C_1. What is perceived is thus $V_1C_2(C_2)V_2$. The masking of the transitions obviously affects stops more than other consonants since stops do not carry independent internal cues that could compensate for the weakness of the contextual ones.[10,11]

3.2. THEORETICAL APPARATUS

The last section established that the optimal position for a consonant is the CV context, and enumerated a number of factors that influence the perceptibility of

[10]Kang does not distinguish coronals in preobstruent vs. presonorant positions. We expect them to be more vulnerable before obstruents. Sonorants have a formant structure and may carry the needed F2 transition. But its amplitude is reduced in comparison with vowels, especially for nasals. We indeed find a three-way contrast between coronal stops in prevocalic, presononant, and preobstruent position in Attic Greek: they are systematically avoided before obstruents, only marginally so before sonorants (see note 39 in chapter 1), and not at all before vowels.

[11]Coronal stops are not weaker than other stops in all languages. They may even be the only segments allowed in preconsonantal position, in particular in Australian languages (Hamilton 1996). These languages typically contrast different coronal places of articulation and Kang argues that the presence of this phonemic contrast, primarily cued by F2, forces speakers to maintain accurate tongue body positions in the production of coronals, even in postvocalic position. The F2 transition thus remains salient, and so does the consonant. In other cases, e.g. Finnish, all stops are consistently audibly released in all positions, providing sufficient cues to coronal stops even with a weakened tongue body gesture and F2 transition.

consonants: the presence of vocalic transitions, the amount of contrast with neighboring segments, the strength of the adjacent boundary (if any), the presence of internal cues, and, for stops, the audibility of the release burst.

These phonetic factors impact the grammar by motivating both markedness and faithfulness constraints. The focus is on a family of markedness constraints against non-prevocalic consonants, that is consonants that are not in a perceptually optimal position. These constraints interact with faithfulness constraints which encode the relative perceptual impact of a modification of the input. The perceptually-motivated constraint system I propose to account for the generalizations established in the preceding chapters raises a number of issues, which have to do with the role of perception, and more generally phonetic and functional factors, in phonology (3.2.1 and 3.2.4), and the integration of variation in Optimality Theory, which is crucial in the analyses to follow (3.2.5). I suggest in particular that the inclusion of perceptually-motivated constraints in the synchronic phonological system is intimately linked to the existence of variable processes.

3.2.1. Perception in phonological theory

As with many concepts in science, perception has gone through a cycle in phonological theory. In the opposition between perception and articulation, Jakobson, Fant & Halle (1952: 12) established the primacy of the former:

> 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 feature system they developed reflects this bias toward the auditory face of speech. The *Sound Pattern of English* (1968) constituted a radical departure from this position, as the distinctive features proposed by Chomsky & Halle are primarily articulatory in nature. The articulatory orientation has been maintained in subsequent work on distinctive features and feature geometry (e.g. Clements 1985; McCarthy 1988), and even reinforced in Sagey (1986) and Halle (1995) by direct reference to articulators in the definition and organization of features.

The fundamental role played by features in phonological description and analysis cannot but influence the range of topics investigated and the way we look at

them. For example, as discussed in Hura et al. (1992), articulatory features showed a clear advantage over acoustic/auditory ones in the treatment of assimilation processes (e.g. palatalization before high-front vowels, place assimilation of nasals). As a result, these processes are typically viewed in phonology as motivated by articulatory factors. Yet more phonetically-oriented research on assimilation has shown that perception is crucial in assimilatory processes (e.g. Kohler 1990; Ohala 1990; Hura et al. 1992, who provide additional references). By contrast, patterns that do not seem to be naturally expressible in terms of the standard articulatory-based features are more likely to be overlooked or analyzed in a more ad hoc fashion. See Flemming (1995) for numerous examples. The special vulnerability of stops in deletion and epenthesis patterns may also fall into this category.

In contrast with standard phonology, however, research made by or in collaboration with phoneticians continues to stress the role of perception in shaping sound patterns. Among the influential proposals highlighting the contribution of perceptual factors, one should mention: Liljencrants & Lindblom's (1972) work on the role of perceptual distance in the configuration of vocalic systems and Lindblom's (1986, 1990) Theory of Adaptive Dispersion (see also Joanisse & Seidenberg 1998); Stevens's (1972, 1989) Quantal Theory of speech; the theory of enhancement features (Stevens, Keyser & Kawasaki 1986; Stevens & Keyser 1989; Keyser & Stevens 2001); numerous works by John Ohala (e.g. 1981, 1983, 1992, 1993, 1995, etc.), as well as Kawasaki (1982) and Kawasaki-Fukumori (1992).

The recent development of Optimality Theory, however, is associated with a renewed interest in the phonetic – in particular perceptual – motivations of phonological patterns and their direct integration into phonological analyses. Indeed, it can be argued that a "serious coming to grips with phonetic functionalism" was not workable in pre-OT non-constraint-based approaches (Hayes 1999: 244). The old idea of sound patterns being the outcome of a competition between the demands of the speaker and the hearer – maximizing articulatory ease vs. the distinctiveness of contrast – has been reappropriated in much recent work, which cite such authors as Passy (1891, cited in Boersma 1999), Zipf (1949), or Martinet (1955). This functionally-motivated phonology has been advocated particularly forcefully in work conducted at UCLA (Flemming 1995; Jun 1995; Silverman 1995; Hayes 1999; Steriade 1999a,c,d, to appear; Kirchner 1998; Fleischhacker 2000a,c), to which we may add Hamilton (1996), Côté (1997a, 1999), Padgett (1997), Boersma (1998, 1999), Hume (1999), Y. Kang (1999, 2000), Kochetov (1999), and Hume & Johnson (to appear).

The sequential approach to deletion and epenthesis processes developed here pursues the line of reseach advocated in the above-cited works. It is both motivated and constrained by direct reference to perceptual factors. It adopts more specifically the 'Licensing by cue' approach developed by Steriade (1999a,c).

In two important papers, Steriade (1999a,c) argues against the prosodic or syllabic approach to phonotactic processes, and develops an Optimality-theoretic account directly based on perceptual cues. Her hypothesis, refered to as 'Licensing by cue', is phrased as follows: "The likelihood that distinctive values of the feature F will occur in a given context is a function of the relative perceptibility of the F-contrast in that context" (Steriade 1999a: 4). In other words, retention of distinctive features in a given context correlates with the number and quality of the available perceptual cues to that feature in that context. Cues do not depend on syllable structure but on the nature of adjacent segments and boundaries. In her 1999c paper, Steriade applies this approach to laryngeal features; the 1999a one develops a more succint analysis of aspiration and place contrasts. I present here the voicing neutralization case, addressed in the first half of her 1999c paper (leaving aside issues of aspiration and ejection, dealt with in the second half). Kochetov (1999) applies Steriade's approach to palatalization; my own analysis of deletion and epenthesis can be interpreted as an extension of it to whole segments rather than features.

Obstruent devoicing and voicing neutralization have been considered classic examples of prosodically-driven feature-changing processes (e.g. Rubach 1990; Lombardi 1991, 1995, 1999; Bethin 1992; Gussmann 1992). They are described as dependent on syllabic affiliation, and typically apply in coda position. Steriade argues that the retention of distinctive voicing rather follows from the availability of possible cues to voicing in different contexts. The cues to the voicing specification of stops and the contexts where they can be found are summarized below; V_1 and V_2 correspond to the preceding and following vowel, respectively.

(5) CUES TO VOICING CONTRASTS AVAILABLE IN DIFFERENT CONTEXTS (based on Steriade 1999C: 30-31):

Cue	Context where it can be found
Closure voicing	Everywhere
Closure duration	Everywhere
V1 duration	Only after sonorant
F0 and F1 values in V1	Only after sonorant
Burst duration and amplitude	Not before obstruents
VOT value	Before sonorant
F0 and F1 values at the onset of voicing in V2	Before sonorant

We can then establish a hierarchy of contexts, from those that provide the most cues to voicing and in which voicing contrasts are best perceived, to those that provide the fewest cues and in which voicing contrasts are the least perceptible. This perceptibility scale is given below, with 'context x' → 'context y' being interpreted as context x is less favorable to the perception of voicing contrasts than context y.

(6) HIERARCHY OF CONTEXTS FOR THE PERCEPTIBILITY OF VOICING CONTRASTS (based on Steriade 1999C: 35):

O=obstruent R=sonorant #=final position

O__O , #__O → O__# → R__O → R__# → __R → R__R

This scale projects a corresponding hierarchy of markedness constraints against the preservation of voicing contrasts, of the form *αvoice/X – do not maintain a voicing contrast in context X. The constraints are universally ranked according to the perceptibility of voicing values: the lower it is in a given context X, the higher ranked the constraint *αvoice/X is.

(7) HIERARCHY OF MARKEDNESS CONSTRAINTS AGAINST THE PRESERVATION OF VOICING CONTRASTS (based on Steriade 1999C: 35):

*αvoice/O__O, #__O >> *αvoice/O__# >> *αvoice/R__O
>> *αvoice/R__# >> *αvoice/__R >> *αvoice/R__R

These markedness constraints interact with a faithfulness constraint militating for the preservation of input [voice] values: PRESERVE [voice]. The position of PRESERVE [voice] within the hierarchy of *αvoice constraints will determine the contexts in which voicing neutralization applies or not. For example, if PRESERVE [voice] is inserted between *αvoice/R__# and *αvoice/__R, voicing contrasts are maintained only before sonorants. According to Steriade, this is the pattern found in several Indo-European languages, among them Lithuanian.

Lithuanian constitutes the most transparent counterexample to the prosodic account provided by Steriade. The argument runs as follows. There is agreement that Lithuanian medial clusters are heterosyllabic, regardless of the nature of the consonants, e.g. *áuk.le*, not **áu.kle*. Distinctive voicing is preserved before sonorants but lost elsewhere, that is before obstruents and word-finally. For example, the opposition between *áukle* 'governness' and *auglingas* 'fruitful' and that between *silpnas* 'weak' and *skobnis* 'table' illustrate that stops may be voiced or voiceless before laterals and nasals. Word-finally obstruents are all voiceless, e.g. *kàd* [kat], and before another obstruent they assimilate in voicing, e.g. *dèg-ti* [kt] 'burn-INF'. In all these cases the (first) obstruent arguably appears in coda position, yet it may or may not maintain voicing contrasts. We conclude that the behavior of voicing features does not depend on the syllabic position but on the nature of the following segment. The ranking *αvoice/__O, __# >> PRESERVE [voice] >> *αvoice/__R nicely and simply accounts for the Lithuanian pattern. I refer the reader to Steriade's paper for a discussion of similar and other cases.

3.2.2. MARKEDNESS CONSTRAINTS

The evidence presented in chapters 1 and 2 supports the hypothesis that the behavior of phonological elements is shaped by their perceptibility, and applies it to segment deletion and epenthesis. I propose that the principle of perceptual salience in (1) impacts the phonology through markedness constraints that miliate against segments that are not perceptually salient. (I restrict my attention to segments but the idea and its implementation could extend to other phonological elements.) These constraints obey the general format in (8):

(8) GENERAL FORMAT OF PERCEPTIBILITY-BASED MARKEDNESS CONSTRAINTS:
S — X   A segment S appears in a context X where it is perceptually salient.

Here I consider only cases where S is a consonant. I take vocalic transitions to be crucial in a consonant's perceptibility, and I assume that consonants are maximally salient in prevocalic position, reflecting the privileged status of CV sequences. The whole architecture to be developed below rests on these observations and on a corresponding family of constraints against non-prevocalic consonants (which are necessarily in a perceptually non-optimal position). I propose the following two basic constraints, which reflect the general part of generalization 1.[12] The double arrow "↔" is used throughout to refer to adjacency, the simple arrow "→" indicates precedence.[13]

[12] These constraints were used independently by Fleischhacker (2000a,b), and the one in (9a) also by Steriade (1999d)
[13] We could also imagine a constraint C ← V "C is preceded by V", which would be posited if the preceding vowel provided better cues than the following one. This does not correspond to the general situation, but according to Steriade (1999a,c), retroflexion would be a relevant case, as she argues that it is better cued by a preceding vowel than by a following one. We might then need a constraint specific to retroflex consonants like [retroflex] ← V "a retroflex consonant is preceded by a vowel". But I do not deal at all with retroflex consonants in this dissertation.

(9) BASIC CONSTRAINTS ENFORCING ADJACENCY TO VOWELS (*Generalization 1*):
a. C ↔ V   A consonant is adjacent to a vowel.
b. C → V   A consonant is followed by a vowel.

Not all consonants are equivalent with respect to the desirability to benefit from the cues associated with an adjacent or following vowel. I integrate this fact into the system by allowing the target of these constraints – C – to be specified for any factor that affects its perceptibility: those concerned with the consonant itself (classes of consonants) and those that depend on the context (neighboring segments, adjacent boundaries[14]). More specifically, the following arguments can be specified.

(10) FACTORS AFFECTING CONSONANT PERCEPTIBILITY:
a. Class of consonants
Ex: stops, strident fricatives, nasals, coronal stops, etc.
b. Similarity with adjacent segments, expressed in terms of agreement or contrast in some feature F
Ex: agreement or contrast in place of articulation, continuancy, voicing, etc.
c. Presence of an adjacent boundary
Ex: Followed by an Intonational Phrase boundary, preceded by a Prosodic Word boundary, etc.
d. (For stops) Nature of the following element (as it affects the audibility of the release burst)

To account for generalizations 2-6, I design the constraints in (11)-(15), which are specific instantiations of the constraints in (4):

(11) CONSTRAINTS ENCODING THE SPECIAL STATUS OF STOPS (*Generalization 2*):
a. stop ↔ V   A stop is adjacent to a vowel.
b. stop → V   A stop is followed by a vowel.

(12) CONSTRAINTS ENCODING THE ROLE OF THE ELEMENT FOLLOWING A STOP (*Generalization 3*):
a. stop(¬__ [+cont]) ↔ V   A stop that is not followed by a [+continuant] segment is adjacent to a vowel.
b. stop(¬__ [+cont]) → V   A stop that is not followed by a [+continuant] segment is followed by a vowel.

[14] One could include the location of stress, which also affects salience.

(13) CONSTRAINTS ENCODING THE ROLE OF SIMILARITY (*Generalization 4*):[15]

a. C(AGREE=F) ↔ V — A consonant that agrees in some feature F with a neighboring segment is adjacent to a vowel.

b. C(AGREE=F) → V[16] — A consonant that agrees in some feature F with a neighboring segment is followed by a vowel.

(14) CONSTRAINTS ENCODING THE ROLE OF BOUNDARIES (*Generalization 5*):

a. $C|_i$ ↔ V — A consonant that is next to a boundary i is adjacent to a vowel.

b. C|i → V — A consonant that is next to a boundary i is followed by a vowel.

(15) CONSTRAINTS ENCODING THE SPECIAL STATUS OF CORONAL STOPS (*Generalization 6*):

C(cor stop) → V — A coronal stop is followed by a vowel.

In addition, for the constraints in (14) we must distinguish the preceding from the following boundaries, since they affect the phonotactics differently. This is not unexpected since, as we saw, left and right edges are not enhanced through the same mechanisms. (14) is decomposed in the two subcases below:

(16) CONSTRAINTS ENCODING THE EFFECT OF THE FOLLOWING BOUNDARY:

a. $C]_i$ ↔ V — A consonant that is followed by a boundary i is adjacent to a vowel.

b. C]i → V — A consonant that is followed by a boundary i is followed by a vowel.

[15] The role of similarity with adjacent segments is encoded in the constraints in (13) in terms of featural agreement, but it could equally well be expressed in terms of featural contrast, as in the constraints below:

(i) C(CONTRAST=F) ↔ V — A consonant that contasts only in some feature F with a neihboring segment is adjacent to a vowel.

C(CONTRAST=F) → V — A consonant that contasts only in some feature F with a neihboring segment is followed by a vowel.

I will stick to the agreement constraints in (13) in this dissertation, but I see no reason why one formulation should be prefered over the other. Agreement and contrast are really two faces of the same phenomenon. These markedness constraints being assumed to be built in the course of acquisition, it is reasonable to believe that language learners enjoy a relative degree of freedom in the formulation of these constraints.

[16] As we will see in the following chapter, this constraint is equivalent to an OCP-[F] constraint between adjacent segments.

(17) CONSTRAINTS ENCODING THE EFFECT OF THE PRECEDING BOUNDARY:

a. $_i$[C ↔ V — A consonant that is preceded by a boundary i is adjacent to a vowel.

b. $_i$[C → V — A consonant that is preceded by a boundary i is followed by a vowel.

These specifications can be freely combined to create more complex constraints. The constraints in (12) involve such a combination since they are specified for stops and the nature of the following element. The agreement and contrast specifications can also be combined with themselves, if different features are involved. Some examples follow:

(18) EXAMPLES OF CONSTRAINTS COMBINING DIFFERENT ARGUMENTS:

a. stop$]_i$ ↔ V
A stop that is followed by a boundary i is adjacent to a vowel.

b. stop(¬__ [+cont] ∧ AGREE=F) → V
A stop that is not followed by a [+continuant] segment and that agrees in a feature F with a neighboring segment is followed by a vowel.

c. $_i$[C(AGREE=F ∧ G) ↔ V
A consonant that is preceded by a boundary i and that agrees in the features F and G with a neighboring segment is adjacent to a vowel.

Within the family of constraints against non-prevocalic consonants, specific constraints may be inherently ranked. I assume that inherent ranking between two constraints is, as are the constraints themselves, based on perception and the principle of perceptual salience. I propose the condition in (19) for establishing such rankings:

(19) DOMINANCE CONDITION:

A constraint $\mathbb{C}1$ dominates a constraint $\mathbb{C}2$ if and only if the candidates that violate $\mathbb{C}1$ are, everything else being equal, equally or less perceptible than the candidates that violate $\mathbb{C}2$.

The effect of this constraint ranking is to have the less perceptible candidates eliminated before the more perceptible ones. This is what we expect from the grammar since, everything else being equal, a more perceptible candidate is always preferable to (more harmonic than) a less perceptible one. So a constraint that militates against less perceptible segments should be ranked higher than a constraint against more perceptible ones.

The ranking condition in (19) enables us in particular to establish the following dominance relation between the two constraints in (9):

(20) DOMINANCE RELATION BETWEEN THE CONSTRAINTS IN (9):
C ↔ V >> C → V

This ranking, as it will become clear later, is crucial for the analyses to follow. It is derived in the following way. Consider the following strings of segments, where § represents a pause. The consonants with a letter subscript violate both C↔V and C→V; those with a number subscript violate only C→V. No consonants may violate C↔V without simultanously violating C→V.

(21) ...V$\mathbf{C_1C_d}$CV... ...V$\mathbf{C_2C_e}$§ §$\mathbf{C_f}$CV...

Everything else being equal, I assume that consonants that lack vocalic transitions are less perceptible than consonants that benefit from transitions from at least one vowel. The letter-subscripted consonants are therefore less perceptible than the number-subscripted ones. So the consonants that violate C↔V are either equally or less perceptible than those that violate C→V. This meets the conditions in (19) for establishing the dominance relation C↔V >> C→V. This is the only possible ranking between the two constraints; the reverse order is excluded since it is not the case that the consonants that violate C→V are all equally or less perceptible than the consonants that violate C↔V. The ranking in (20) can be extended to all the constraints derived by specifying one or more of the arguments in (10): for all C_j, where C_j is any specified consonant, the ranking C_j↔V >> C_j→V necessarily holds, e.g. stop↔V >> stop→V, $C]_i$↔V >> $C]_i$→V, etc.

The rankings in (22) can be established in the same way. They follow straightforwardly from the perceptual facts described in section 3.1: stops are less perceptible than other consonants in non-prevocalic position (22a); stops that are not followed by a [+cont] segment are less perceptible than other stops (22b); consonants that are more similar to (i.e. agree in some feature F with) an adjacent segment are less perceptible than consonants that are less similar (i.e. do not agree in the same feature F) (22c-d); consonants that are adjacent to a weaker boundary i are less perceptible than consonants that are adjacent to a stronger boundary j (22e). I note the absence of boundary with the symbol Ø. Consonants that are adjacent to no boundary are the least perceptible, which establishes the ranking in (22f).

(22) INHERENT RANKINGS BETWEEN MARKEDNESS CONSTRAINTS:
a. stop → V >> C → V
stop ↔ V >> C ↔ V
b. stop(¬__ [+cont]) → V >> stop → V
stop(¬__ [+cont]) ↔ V >> stop ↔ V
c. C(AGREE=F) → V >> C → V
C(AGREE=F) ↔ V >> C ↔ V
d. C(AGREE=F∧G) → V >> C(AGREE=F) → V ; C(AGREE=G) → V
C(AGREE=F∧G) ↔ V >> C(AGREE=F) ↔ V ; C(AGREE=G) ↔ V
e. $C|_i$ → V >> $C|_j$ → V if i is a weaker boundary than j
$C|_i$ ↔ V >> $C|_j$ ↔ V if i is a weaker boundary than j
f. $C|_Ø$ → V >> $C|_i$ → V if i ≠ Ø
$C|_Ø$ ↔ V >> $C|_i$ ↔ V if i ≠ Ø

This basically exhausts the rankings that will be needed in the analyses to come. Note that these ranked constraints all are in a subset relation to one another, e.g. stops are a subset of consonants; consonants that are adjacent to a boundary j are a subset of consonants that are adjacent to a lower boundary i (including no boundary). The constraints only differ in one dimension whose effect on perceptibility is considered clear. The rankings I use never involve multidimensional comparisons of perceptibility, for example comparing stops at a boundary j and non-stops at a lower boundary i, which contrast in two dimensions with opposite effects on perceptibility. Avoiding multidimensional perceptibility comparisons allows us to escape a lot of potential difficulties and controversies, in view of the complexity involved in such comparisons. See Flemming (1995) for a similar situation. But multidimensionality is certainly an issue that should be taken up in the future.

Before leaving this section, a final word about the Sonority Sequencing Principle, which was crucially involved in the case studies in chapters 1 and 2. The phonetic nature of sonority is not yet clearly understood, nor is its relation to perception and articulation (see Clements 1990 for discussion). I take it here to be independent from the Principle of Perceptual Salience. To account for its role in consonant deletion and vowel epenthesis, I simply propose the constraint in (23), which meets our needs:

(23) SONORITY SEQUENCING PRINCIPLE (SSP):
Sonority maxima correspond to sonority peaks.

3.2.3. FAITHFULNESS CONSTRAINTS

The markedness constraints against non-prevocalic consonants interact with faithfulness constraints to yield the attested patterns. Since I deal here only with epenthesis and deletion, I use the following two basic constraints (from McCarthy & Prince 1995):

(24) BASIC FAITHFULNESS CONSTRAINTS:
a. MAX — Do not delete
b. DEP — Do not epenthesize

It has been noticed several times, however, that these general faithfulness constraints do not allow us to reduce the set of optimal candidates to the desired singleton (Lamontagne 1996; Steriade 1999d; Wilson 2000). The problem is easy to see. I illustrate it first with a hypothetical case of consonant deletion, and discuss epenthesis later. Suppose an input of the form /VC_1C_2V/ and a grammar G characterized by the two constraint rankings C→V >> MAX and DEP >> MAX. This grammar yields obligatory deletion of one of the two consonants, to ensure that all consonants in the output are followed by a vowel. But it cannot determine *which* consonant to delete. As illustrated in the tableau below, the outputs [VC_1V] and [VC_2V] are equivalent with respect to G. Here and in the rest of this dissertation I use thick lines between columns to indicate that the constraint at the left dominates that at the right, e.g. between DEP and MAX in (25). Thin lines between two constraints indicate ranking indeterminacy between them, e.g. between C→V and DEP.

(25) FAILURE TO IDENTIFY THE CORRECT DELETION SITE:

/VC_1C_2V/	C → V	DEP	MAX
a. VC_1C_2V	* !		
b. VC_1VC_2V		* !	
c. → VC_1V			*
d. → VC_2V			*

G then needs to be augmented to be able to pick between candidates c. and d. I propose that this is done by using context-sensitive faithfulness constraints, whose ranking is perceptually motivated and determined by considerations of relative perceptibility of constrasts. This corresponds to the partial adoption of Steriade's (1999b,d, 2000b, to appear) new approach to correspondence, based on a linguistic component called the P-map. Other proposals that are meant to solve this problem

include Relativized Contiguity (Lamontagne 1996) and targeted constraints (Wilson 2000), which I will review in turn.

Lamontagne proposes that the choice between VC_1V and VC_2V is to be made by contiguity constraints which demand that any sequence of segments contiguous in the input/output be contiguous in the output/input. He defines two general types of contiguity constraints, called DOMAIN-CONTIGUITY (D-CONTIG) and JUNCTURE-CONTIGUITY (J-CONTIG), which evaluate contiguity between segments within a domain and across adjacent domains, respectively, where domains correspond to prosodic units like the syllable, the foot, the Prosodic word, etc. D-CONTIG penalizes the existence of segments that are contiguous within a constituent in the output, but are not contiguous in the input. J-CONTIG penalizes the existence of segments that are contiguous across a boundary in the output, but are not contiguous in the input. The ranking between these two constraints determines which consonant to delete or where to epenthesize.

Consider the same /$V_aC_1C_2V_b$/ input and the two possible outputs [$V_a.C_1V_b$] and [$V_a.C_2V_b$], syllabified as indicated by the dot. The [$V_a.C1V_b$] output violates D-CONTIG(syllable): C1 and V_b are contiguous within a syllable in the output, but they are not contiguous in the input. But the same output does not violate J-CONTIG(syllable), since V_a and C1, which are contiguous across a syllable boundary in the output, are also contiguous in the input. The candidate [$V_a.C_2V_b$] is the mirror image of [$V_a.C_1V_b$]. It violates J-CONTIG(syllable) (since V_a and C_2 are contiguous across a syllable boundary in the output but they are not contiguous in the input) but not D-CONTIG(syllable). Which of [$V_a.C_1V_b$] and [$V_a.C_2V_b$] turns out to be optimal depends on the language-specific ranking between J-CONTIG(syllable) and D-CONTIG(syllable). If D-CONTIG(syllable) dominates J-CONTIG(syllable), [$V_a.C2V_b$] wins out and it is the first consonant that deletes. Diola Fogny instantiates this ranking, e.g. /let-ku-jaw/ → [lekujaw] 'they won't go'. If J-CONTIG(syllable) outranks D-CONTIG(syllable), [$V_a.C_1V_b$] is selected. As an example of this ranking, Lamontagne cites Wiyot (Teeter 1964), e.g. /pucarag+lolisw-/ → [pucaragorišw-] 'whistle a tune' (where /g/ corresponds to /ɣ/ in Teeter's transcription).

Lamontagne's solution works; the problem I see with it is that it considers the deletion of C_1 and C_2 equally likely. In fact they are not; Wilson (2000) and Steriade (1999b) note that it is typically the first consonant that deletes, as in Diola Fogny, and both relate this fact to the better cues associated with prevocalic consonants, hence their higher perceptibility and greater resistance (see section 3.1.1). Wilson claims that known exceptions to this pattern – that is deletion of the second (prevocalic) consonant – involve independent factors, in particular a preference for keeping stem consonants over affixal ones, or less sonorous consonants (which form better onsets) over more sonorous ones. Turkish (Keyser & Clements 1983) is given as an illustration of morphologically-based deletion, Pāli (Hankamer & Aissen 1974) as one of sonority-based deletion.

As for Wiyot, the evidence it provides is unclear. Teeter (1964: 26) does suggest that illicit combinations of two consonants across morpheme boundaries are repaired by deletion of the second element. Supporting data, however, are scarce. Teeter cites one exception to his generalization: when /h/ is followed by a consonant with which it cannot combine, it is the /h/ that deletes. Interestingly, all but one of the examples I have found of deletion of the prevocalic consonant in /...VC+CV.../ also involve /h/ in /C+h/ sequences. One may wonder, then, whether it is not the deletion of the laryngeal consonant that is favored, irrespective of its position. Deletion of a prevocalic consonant other than /h/ was only found in the example cited above (/pucarag+lolisw-/ → [pucaragorišw-] 'whistle a tune'), on which I cannot comment.[17]

Granting the unconclusiveness of the Wiyot case, the theory should predict that, everything else being equal, it is the postvocalic consonant rather than the prevocalic one that deletes in a VCCV sequence. Both Wilson (2000) and Steriade (1999b,d, 2000b) accomplish this. Wilson derives this result by introducing a new type of markedness constraints, called targeted constraints, whose main novelty is to restrict the candidates that are being compared by these constraints to a set of forms that are considered similar enough, according to a similarity criterion. Similarity here is defined in terms of perceptual confusability. Formally, a targeted constraint →C is defined in terms of a specific statement of absolute markedness and a similarity criterion. For any two candidates a and b, the targeted constraint →C prefers a over b iff a is less marked than b according to the absolute markedness statement and a is considered sufficiently similar to b.

A more concrete example will make this system clearer. Take again our hypothetical VC_1C_2V case and assume the targeted constraint →NoWeak-C, which militates against segmental root nodes in the output (the absolute markedness statement corresponds to *Struc(Rt)). Wilson states that consonants in preconsonantal position are perceptually weak (on which we agree), that is they are difficult to distinguish from Ø. Prevocalic consonants, however, are associated with

[17]There is a class of inalienable nouns that may appear to involve the deletion of a prevocalic consonant in possessed forms (pp. 80-81), e.g. *bápt* 'teeth' but *kh́ápt* 'your teeth', containing a second person possessive prefix *kh-*. All the unpossessed forms of the words in this class, however, begin with /b.../, which is most probably not part of the base but also a prefix.

strong cues. The constraint →NoWeak-C only compares candidates that are perceptually comparable, i.e. VC_1C_2V and VC_2V, but crucially not VC_1C_2V and VC_1V. In this limited competition, VC_2V fares better on *Struc(Rt) and wins. The crucial consequence of the targeted constraint is to evacuate the candidate VC_1V, which is in the end what we aim at.

Wilson's proposal crucially relies on perceptual salience and auditory similarity, which are I believe the relevant factors. It is C_1 that deletes because it is perceptually weaker than C_2 (recall the comparison between consonants in CV and VC contexts in section 3.1.1). But my main concern about targeted constraints is the dichotomized split they impose between the comparable and non-comparable candidates. How are we to define and determine the levels of acceptable similarity, acknowledging the gradient nature of perceptibility? This issue has immediate empirical consequences. Take a more complex three-consonant cluster $VC_1C_2C_3V$. Under simplification, it is typically C_2 that deletes, which is the consonant that does not benefit from any vocalic transitions. C_3 is the perceptually strongest consonant (everything else being equal), C_1 being in an intermediate situation between C_2 and C_3. We may safely assume that $VC_1C_2C_3V$ and VC_1C_3V are comparable under →NoWeak-C, and that VC_1C_2V is excluded from the comparison. But what about VC_2C_3V? Should it be considered similar enough to $VC_1C_2C_3V$? The answer is no if we want VC_1C_3V to end up as the only optimal candidate; because if we include VC_2C_3V in the comparison, both VC_1C_3V and VC_2C_3V will fare equally. But is there a motivation for this exclusion, other than the desire to get the correct result?

Consider now a case where C_2 cannot delete for some independent reason; for example, it has to surface because of its morphological status. C_1 would then be more likely to delete than C_3. Unfortunately, I do not have a specific pattern at hand, but suppose that there exists a language in which C_1 deletes if the deletion of C_2 is ruled out by some independent higher-ranked constraint. Such a case does not seem to me to be at all implausible. If both VC_2C_3V and VC_1C_2V are excluded by the targeted constraint, we find again the initial problem and the grammar cannot choose between deleting C_1 and deleting C_3. In this language, the targeted constraint should consider the intermediate candidate VC_2C_3V if we are to derive the correct output.

I do not believe that it is fatal for Wilson's proposal that the set of similar enough candidates is grammar-specific; indeed, this may be the expected situation. But I think that the dichotomy involved in the similarity criterion of targeted constraints is at odds with the inherent relativity of perceptibility. Rather than deciding whether or not a candidate is to be included in the evaluation of a constraint, grammars should encode the relative likelihood that consonants in different positions delete. This can be done quite naturally in a framework such as Optimality Theory. Determining which consonant will ultimately be dropped then follows from interactions with other constraints.

This is precisely what Steriade's (1999b,d, 2000b, to appear) approach to faithfulness constraints achieves. Steriade proposes that faithfulness or correspondence constraints are projected from, and their ranking determined by, a grammatical component, called the P-map. The P-map is a set of statements about perceived distinctiveness differences between different contrasts in different contexts. For example, the P-map may tell us that the contrast between [t] and [d] is better perceived before a vowel than before a consonant (same contrast in different positions), or that the contrast between [t] and [n] is better perceived than the contrast between [t] and [d] word-finally (different contrasts in the same environment). The contrast and the context may covary and the P-map can also claim that the contrast between Ø and [ə] after a consonant word-finally is better perceived than the contrast between [t] and [d] after a vowel word-finally (examples from Steriade 2000b). These comparisons are derived from statements about the absolute distinctiveness or perceptibility of contrasts. Each contrast x-y/__K (contrast between x and y in context K) is associated with a specific distinctiveness index and projects a corresponding faithfulness constraint of the form CORRESP.(x-y/__K). If it can be determined from the P-map that a contrast x-y/__K is more perceptible than a contrast w-z/__Q, then for any correspondence constraint, CORRESP.(x-y/__K) dominates CORRESP.(w-z/__Q).

Let us go back to our VC_1C_2V example again. We have determined that in this context C_2 is perceptually more salient than C_1 (everything else being equal). In other words, the contrast between C and Ø in the context C__V is more distinctive or perceptible than the contrast between C and Ø in the context V__C. Translated in terms of the correspondence constraint MAX-C, this comparison derives the ranking MAX-C/C__V >> MAX-C/V__C. This ranking determines that, everything else being equal, deletion of a postvocalic consonant is always favored over that of a prevocalic one. That is, VC_1C_2V is reduced to VC_2V and not VC_1V, as shown in the tableau. This is the result we intended to derive.

(26) GETTING THE DELETION SITE WITH CONTEXT-SENSITIVE FAITHFULNESS:

/VC_1C_2V/	C → V	MAX-C/C__V	MAX-C/V__C
a. VC_1C_2V	* !		
b. VC_1V		* !	
c. → VC_2V			*

To account for the simplification of three-consonant clusters $VC_1C_2C_3V$, we need to extend the ranking of MAX-C constraints to include the constraint against deletion of interconsonantal consonants MAX-C/C__C. Such consonants are less perceptible than consonants that benefit from vocalic transitions. Again, the contrast between C and Ø in the context C__C is less distinctive than the contrast between Ø and a C that is adjacent to a vowel. Consequently, MAX-C/C__C is ranked lower than the constraints against deletion of pre- and post-vocalic consonants:

(27) RANKING OF CONTEXT-SENSITIVE MAX CONSTRAINTS:
MAX-C/C__V >> MAX-C/V__C >> MAX-C/C__C

This ranking ensures that if nothing prevents it, C_2 is the consonant that deletes in $VC_1C_2C_3V$ sequences. But it also follows from it that if deletion of C_2 is ruled out by some independent constraint, it is C_1 that deletes, not C_3 (provided the appropriate ranking of the markedness constraint that motivates deletion, say C↔V, above MAX-C/V__C). This situation is illustrated in the tableau below. Let us have a three consonant-cluster in the input and two unviolable constraints: C↔V demanding that every consonant be adjacent to a vowel, and KEEPC_2, which could be any constraint that prevents the deletion of C_2, presumably for morphological reasons. In a grammar without KEEPC_2, it is easy to see that the optimal candidate is VC_1C_3V, given the inherent and perceptually-motivated ranking of the MAX-C constraints. The addition of the high-ranked constraint KEEPC_2 rules out this candidate, and the winner automatically becomes VC_2C_3V.

(28) DELETING THE LEAST PERCEPTIBLE CONSONANT POSSIBLE:

/$VC_1C_2C_3V$/	KEEPC_2	C↔V	MAX-C/C__V	MAX-C/V__C	MAX-C/C__C
a. $VC_1C_2C_3V$		* !			
b. VC_1C_3V	* !				*
c. → VC_2C_3V				*	
d. VC_1C_2V			* !		

This approach to correspondence is perfectly coherent with the basic intuition behind faithfulness constraints: the idea that the input should be modified minimally.

The innovation here is to define what counts as minimal in terms of perceptual distinctiveness. The relative ranking of a faithfulness constraint correlates with the extent to which its violation would perceptually disrupt the input. The ranking in (27) follows from the fact that deleting an interconsonantal consonant has a smaller perceptual impact or is less disruptive than deleting a postvocalic consonant; likewise for postvocalic vs. prevocalic consonants. This approach, however, requires a change in the way we view inputs. Inputs have standardly been considered abstract unpronounceable entities. But if we evaluate faithfulness in terms of perceptual modification, we have to define inputs as elements that are, at least potentially, perceivable, that is, basically, as potential outputs. The consequences of this shift for phonology are not clear to me at this point. It is obvious that this issue deserves a more elaborate discussion, but I can only hope that it will be taken up in the future.

The reasoning that has led to the ranking in (27) can be extended to variables other than the vocalic context of consonants, and can motivate similar rankings. Given two constraints MAX-C_1 and MAX-C_2, MAX-C_1 >> MAX-C_2 iff the contrast between C_2 and Ø is less perceptible than the contrast between C_1 and Ø, in other words if C_2 itself is less perceptible (everything else being equal) than C_1. Section 3.1 identified a number of factors that increase or decrease the perceptibility of consonants. One of them was the presence of adjacent vowels, hence the ranking in (27). Other variables include the nature of the consonant (stops having weaker internal cues than other consonants), the continuancy value of the segment following stops, the amount of contrast with adjacent segments, and the presence of adjacent boundaries. These factors motivated the existence of markedness constraints against non-prevocalic consonants; they motivate faithfulness constraints in the same fashion. The constraints and the rankings that can be derived are given in (29), together with the generalization that they encode:

(29) PERCEPTIBILIY-BASED FAITHFULNESS CONSTRAINTS:

a. *Generalization 1:*
MAX-C/__V >> MAX-C/V__ >> MAX-C
MAX-C/__V Do not delete a consonant that is followed by a vowel.
MAX-C/V__ Do not delete a consonant that is preceded by a vowel.

b. *Generalization 2:*
MAX-C(-stop) >> MAX-C
MAX-C(-stop)
Do not delete a consonant that is not a stop.

c. *Generalization 3:*
MAX-stop/__[+cont] >> MAX-stop
MAX-stop/__[+cont]
Do not delete a stop that is followed by a [+continuant] segment.

d. *Generalization 4:*
MAX-C/CONTRAST=F >> MAX-C (where F is any feature)
MAX-C/CONTRAST=F
Do not delete a consonant that contrasts in some feature F with an adjacent segment.

e. *Generalization 5:*
MAX-C$|_i$ >> MAX-C (where i is any prosodic boundary)
MAX-C$|_i$
Do not delete a consonant that is adjacent to a prosodic boundary i.

Each ranking identifies a factor that affects the salience of consonants. In the general case consonants are endowed with enhancing factors and are correspondingly associated with specific higher-ranked MAX constraints, which dominate the general MAX-C. These include:
1) Consonants that are adjacent to a vowel (29a);
2) Consonants other than stops (29b). Note that I use +/-stop here in a purely descriptive fashion, and do not consider "stop" to be a phonological feature in the strict sense;[18]
3) Stops that are followed by a [+continuant] segment (29c);
4) Consonants that contrast in some feature F with an adjacent segment (29di);
5) Consonants that are adjacent to a prosodic boundary (29e).
The constraints in (29a) and (29b) will be illustrated (and supported) in the analysis of consonant deletion in Sranan in section 3.4 and Québec French in chapter 4. Those in (29c) will be used in the formal accounts developed in chapter 4.

The ranking of faithfulness constraints according to the principle of minimal perceptual disruption or modification of the input also applies to constraints other than MAX-C, in particular DEP-V. Epenthesis is indeed less disruptive in certain

[18]Consonants other than stops could be more formally refered to as: "consonants that bear a positive "+" specification for some manner feature". Stops are [-sonorant], [-continuant], [-approximant], [-vocoid], i.e. they are negatively specified for all manner features, whereas all other consonants have at least one "+" specification for one or more of these features. This is the formulation I used in the original (official) version of this dissertation, but I adopt a more descriptive and straightforward formulation here.

contexts than in others, but the effect of the segmental and prosodic context does not appear to be as clear and systematic as with consonant deletion. In a /VC_1C_2V/ sequence, there is only one possible site for vowel epenthesis (if the motivation is to have every consonant adjacent to a vowel): [VC_1V̲C_2V]. Consider now a three-consonant sequence /V$C_1C_2C_3$V/, not tolerated on the surface. There are two possible outputs: [VC_1V̲C_2C_3V] and [VC_1C_2V̲C_3V]. Each of them is widely attested crosslinguistically, and the choice between them seems to be largely independent from perceptual factors, unlike consonant deletion. The famous contrast between different Arabic dialects (Broselow 1980, 1992; Selkirk 1981; Itô 1986, 1989; Lamontagne 1996; Zawaydeh 1997, among others) illustrates this variation in epenthesis sites: given an underlying three-consonant sequence, Cairene Arabic inserts an epenthetic [i] between the second and third consonants, whereas Iraqi inserts it between the first and second (30). In other languages, epenthesis systematically targets morphemic boundaries, e.g. French (chapter 2) and Chukchi (Kenstowicz 1994b).

(30) VOWEL EPENTHESIS IN CAIRENE AND IRAQI ARABIC:

a. Cairene	/ʔul+t+l+u/	→	[ʔultilu]	'I said to him'
b. Iraqi	/gil+t+l+a/	→	[gilitla]	'I said to her'

The factors underlying the distinction between Cairene and Iraqi are not entirely clear and I will not attempt to enlighten the issue. The contrast has been accounted for with directional syllabification (Itô 1986, 1989), reanalyzed in terms of alignment in Optimality-theoretic terms (Mester & Padgett 1993). Broselow (1992) proposed an alternative analysis, which links the location of epenthesis to the moraic or nonmoraic status of stray consonants, building on Selkirk's (1981) proposal based on the distinction between onsets and codas. I will simply adopt the alignment strategy when the issue arises.

This is not to say that perceptual factors are always irrelevant to the choice of the epenthesis site. Fleischhacker (2000a[19],b,c) conducted a crosslinguistic study of epenthesis in word-initial consonant clusters, in particular in loanword adaptation. I focus here only on two-consonant sequences. Some languages systematically insert the vowel in the same location, either before the two consonants (/CC/ → [V̲CC], e.g. Iraqi Arabic) or inside the cluster (/CC/ → [CV̲C], e.g. Korean). But in an interesting subset of languages, e.g. Egyptian Arabic and Sinhalese (see

[19]Fleischhacker (2000a) is a revised version of her M.A. thesis (2000c), which contains expanded discussion of the cross-linguistic data and results from an additional experiment, while omitting certain details of the experimental portion of the M.A. (Fleischhacker p.c.). I have had only access to this revised version.

Fleischhacker 2000a for additional languages), this choice is determined by the nature of the cluster: initial epenthesis (prothesis) with sibilant+stop (ST) clusters but medial epenthesis (anaptyxis) in stop+sonorant (TR) clusters.[20] No languages display the opposite pattern. What is also found are languages that use prothesis with ST clusters but leave TR clusters intact (e.g. Haitian, Catalan), and languages that allow initial ST clusters but break TR ones with anaptyxis (e.g. Lakhota, Central Yup'ik). What we observe, then, is a clear tendency to favor anaptyxis with stop+sonorant sequences and prothesis with sibilant+stop ones.

Fleischhacker's explanation for this contrast relies on perception and the idea of minimal disruption of the perceptual properties of the input: "the epenthesis site is chosen to maximize auditory similarity between the non-epenthesized input and the ouput" (2000a: 4); in other words, "epenthetic vowels are located exactly where they are least auditorily obtrusive" (p.14). Fleischhacker explains that the stop-sonorant juncture is acoustically similar to a stop-vowel one because both are characterized by a rapid increase in amplitude and onset of formant structure. The epenthetic vowel appears in a location corresponding to a vowel-like portion of the input, where we find no contrast in sonorancy. The sibilant-stop juncture lacks those vowel-like properties and anaptyxis there would constitute a major modification of the input. Prothesis is a better alternative, to the extent that "the output string corresponding to the input is not interrupted by an inserted element" (p.16). Fleischhacker provides experimental support for this perceptually-based hypothesis: əST was judged more similar to ST than SəT by a group of English speakers, while TəR was judged more similar to TR than əTR. She concludes that an inserted vowel is less perceptible, i.e. more confusable with Ø, in the context T__R, and more perceptible between a sibilant and a stop S__T. Word-initial epenthesis (before an obstruent) appears to form an intermediate case between S__T and T__R in terms of the auditory obstrusiveness of the process.[21] This hierarchy of perceptibility of the vowel is reflected in the following ranking of DEP-V constraints:

(31) RANKING OF CONTEXT-SENSITIVE DEP CONSTRAINTS:
DEP-V/S__T >> DEP-V/#__ >> DEP-V/T__R

[20]The behavior of sibilant+sonorant sequences is more variable and depends in particular on the sonority level of the sonorant; I omit these cases and refer the reader to Fleischhacker (2000a) for discussion.

[21]For the position of the word-initial context with respect to auditory similarity and the corresponding ranking in (31), I follow Fleischhacker (2000b). Fleischhacker (2000a) does not compare the context #__ with T__R and S__T, and does not use the coresponding constraint DEP-V/#__; she obtains the expected results by means of faithfulness constraints independent from the ranking in (31). For purposes of expository simplicity, I use the approach exposed in Fleischhacker (2000b).

Patterns with anaptyxis in TR clusters and prothesis in ST ones follow directly from this ranking, epenthesis being motivated by the high ranking of the markedness constraint C↔V. The Lakhota/Central Yup'ik case – anaptyxis in TR but ST allowed – derives straightforwardly from C↔V being ranked above DEP-V/T__R but below DEP-V/#__: only the least obtrusive instances of epenthesis are tolerated. The Haitian/Catalan case – prothesis in ST but TR allowed – appears more problematic, but could be understood in terms of the markedness of ST vs. TR sequences. TR clusters display a contrast in sonorancy absent from ST ones. I suggest that this makes the latter more marked, subject to the constraint C(AGREE=[son])↔V (13), while TR clusters are only affected by the general and lower-ranked C↔V (22c). The ranking in (32) yields the Haitian/Catalan pattern. Prothesis in ST clusters follows from the ranking C(AGREE=[son])↔V >> DEP-V/#__ while the ranking DEP-V/T__R >> C↔V yields the absence of anaptyxis in TR sequences.

(32) RANKING YIELDING PROTHESIS IN ST AND NO EPENTHESIS IN TR:
C(AGREE=[son]) ↔ V >> DEP-V/#__ >> DEP-V/T__R >> C ↔ V

As for patterns with systematic anaptyxis or prothesis, Fleischhacker assumes that they arise from independent requirements, possibly a preference for consonants being followed (rather than preceded) by a vowel (systematic anaptyxis), or a CONTIGUITY constraint (systematic prothesis).

We may briefly venture beyond initial epenthesis, to which Fleischhacker's study is restricted, and reflect on the observed tendency in several languages to epenthesize next to a sonorant but leave obstruent sequences intact. I cite three examples: Winnebago, Irish, and Upper Chehalis. In Winnebago (Miner 1979; Hale & White Eagle 1980), all sequences of an obstruent followed by a sonorant are broken by an epenthetic vowel, either a copy of the following vowel or a slight intrusive schwa. In the second case, the obstruent also becomes voiced. The copy type of epenthesis is known as Dorsey's Law, and is illustrated in the example in (33), from Hale & White Eagle (1980), which also shows the absence of epenthesis in the [kʃ] sequence.

(33) DORSEY'S LAW IN WINNEBAGO:
/ha+ra+ki+ʃ+ru+dʒik-ʃa̧na̧/ → [harakiʃu̲rudʒikʃa̧na̧] 'pull taut, 2ND'

Irish (Carnie 1994; Ní Chiosáin 1996, 1999; Green 1997) displays epenthesis between any sequence of a sonorant followed by a voiced obstruent (34a), while

clusters composed of a sonorant and a voiceless obstruent (34b) or two obstruents (34c) surface intact.

(34) VOWEL EPENTHESIS IN IRISH:

a.	/gorm/	→	[gorə̲m]	'blue'
b.	/kork/	→	[kork]	'Cork (place name)'
c.	/ʃaxt/	→	[ʃaxt]	'seven'

In Upper Chehalis (a Tsamosan Salish language), Rowicka (2000) proposes a rule of schwa epenthesis that applies specifically in sequences composed of a consonant and a sonorant (or a glottal stop), while the language tolerates long clusters of obstruents. The exact contexts for schwa epenthesis, however, are not clearly defined in the paper.

I believe these cases of asymmetry between clusters containing a sonorant and clusters composed only of obstruents can be understood in terms of the perceptual account of epenthesis proposed by Fleischhacker. Epenthesis applies only in clusters where it is not disruptive, leaving intact some marked clusters in which epenthesis would be too salient. This is a particularly welcome result as this asymmetry has remained puzzling. Alderete (1995) has analyzed the Winnebago case in terms of the Syllable Contact Law, which requires sonority to fall across syllable boundaries, but such an analysis cannot extend to the Irish and Upper Chehalis cases. In Irish, the fact that epenthesis is restricted to apply before voiced obstruents is consistent with the perceptual explanation since it is expected that vowel epenthesis will be less obtrusive in the context of voiced segments, which share with vowels the presence of low frequency energy associated with voicing. The fact that voicing favors epenthesis is also independently noticed in Fleischhacker (2000a: 15-16).

In this long section, I have argued for the adoption of perceptually-motivated faithfulness constraints, whose ranking reflects the degree of disruption of the auditory properties of the input. Deletion of less perceptible consonants or vowel epenthesis in a context where the vowel remains relatively non-salient leads to the violation of lower-ranked faithfulness constraints. This approach to correspondence constraints is obviously in keeping with what I have proposed for markedness constraints. In fact, one may be struck by the resemblance between the rankings of the MAX-C constraints in (29) and those of the markedness constraints in (20) and (22), which are the mirror image of one other. Consider in this respect the rankings of MAX-C and markedness constraints in (35), extracted from (20), (22), and (29).

The rankings in (35a-e) express the generalization that consonants that are less perceptible should be avoided more than consonants that are more perceptible. Those in (35f-j) encode the fact that the deletion of consonants that are more perceptible is less easily tolerated than the deletion of consonants that are less perceptible. The correspondence between the two series obviously follows from the fact that they are motivated by the same perceptual factors, and they both result in less perceptible consonants being less likely to surface than more perceptible ones.

(35) EQUIVALENCE BETWEEN MARKEDNESS AND MAX-C CONSTRAINTS:

Markedness constraints	MAX-C constraints
a. C ↔ V >> C → V	f. MAX-C/C__V >> MAX-C/V__C >> MAX-C
Common motivation: prevocalic consonants are most perceptible, postvocalic ones are less perceptible, those that are not adjacent to any vowel are least perceptible	
b. stop → V >> C → V stop ↔ V >> C ↔ V	g. MAX-C(-stop) >> MAX-C
Common motivation: stops are less perceptible than other consonants (in non-prevocalic position)	
c. stop(¬__ [+cont]) → V >> C → V stop(¬__ [+cont]) ↔ V >> C ↔ V	h. MAX-stop/__[+cont] >> MAX-stop
Common motivation: stops that are followed by a [+continuant] segment are more perceptible than other stops	
d. C(AGREE=F) → V >> C → V C(AGREE=F) ↔ V >> C ↔ V	i. MAX-C/CONTRAST=F >> MAX-C
Common motivation: consonants that agree/contrast in some feature F with an adjacent segment are less/more perceptible than consonants that do not.	
e. $C\|_i$ → V >> $C\|_j$ → V $C\|_i$ ↔ V >> $C\|_j$ ↔ V if i is a weaker boundary than j	j. MAX-$C\|_i$ >> MAX-C
Common motivation: consonants that are adjacent to a prosodic boundary are more perceptible than consonants that are not.	

One may worry about the redundancy present in this system. For example, is it necessary to integrate the effect of adjacent vowels (a and f), manner of articulation (b and g), the continuancy value of segments following stops (c and h), contrast/similarity (d and i), or the prosodic boundary (e and j) in both markedness and MAX-C constraints? I believe so, this system being both empirically adequate and maximally coherent. On the one hand, doing away with the context-specific MAX-C and DEP-V constraints yields an empirically inadequate system, which cannot

derive the correct outputs, because it cannot predict which consonant deletes and where epenthesis occurs. This is exactly the reason why the perceptual faithfulness constraints were proposed. On the other hand, failing to incorporate the perceptual motivations into the markedness constraints leads to a theory that seems at best incoherent. This conclusion arises when we consider the existence of multiple strategies to eliminate perceptually weak consonants. Consonant deletion and vowel epenthesis are frequent ones; metathesis is also a possible solution, as illustrated by the Lithuanian and Singapore English cases presented in the appendix to this chapter. In addition, vowel deletion may be blocked to satisfy perceptual requirements. All these processes are subject to the same factors (the presence of adjacent consonants, the perceptual weakness of stops, the strengthening effects of prosodic boundaries and contrast, etc.), and several of them may coexist in the same grammar (e.g. vowel deletion, vowel epenthesis, and consonant deletion in French; see chapters 2 and 4). Perceptually-motivated markedness constraints serve to provide a unified motivation for these different processes. Without such markedness constraints, the perceptual factors would have to be incorporated into each of the faithfulness constraints as well as the constraint motivating vowel deletion. We would then need our constraint ranking to encode, for example, the fact that epenthesis is more easily tolerated next to stops than next to other consonants. This appears inconsistent with the finding above that epenthesis is more likely next to a sonorant. The former generalization stems from the marked nature of stops lacking an adjacent vowel, the latter from the preference for less obtrusive epenthesis. Incorporating both of them into the ranking of DEP-V constraints would require it to meet potentially conflicting requirements: maximizing similarity between input and output and "saving" weak consonants. These requirements are better kept apart and dealt with by separate faithfulness and markedness constraints, as in the ranking in (32) above for the Haitian/Catalan pattern of initial epenthesis. The conclusion that both markedness and faithfulness constraints need to be context-specific is also reached by Kang (1998); see also Zoll (1998) who argues that positional markedness constraints are a necessary component of the grammar.

• Note on the P-map and the "Too-many-solutions problem"

Before closing this section, I should add a few comments concerning the scope of Steriade's proposal regarding perceptually-motivated constraints, and my position with respect to it. First, note that the main motivation behind Steriade's new approach to correspondence is not so much to solve the problem of which consonant to delete or where to insert a vowel in cluster simplification, although this is obviously a welcome result of it, but to develop a theory that better predicts the range of repair strategies that are available to a given phonotactic constraint. The idea is easy to grasp: in current versions of OT, any phonotactic constraint can be met by the use of any possible repair strategy, depending on the ranking of the various faithfulness constraints. For example, suppose that a grammar disallows voiced obstruents word-finally. In principle, an input of the form /tab/ could be modified in a number of different ways to conform to this phonotactic requirement: devoicing [tap], nasalization [tam], approximantization [taw], epenthesis [tabə], deletion [ta], metathesis [bat], etc. Since the faithfulness constraints that prevent these processes are ranked freely, we expect to find languages that instantiate each of these solutions, depending on which of the faithfulness constraint is ranked lowest:

(36) PREFERED OUTPUT DEPENDING ON THE LOWEST-RANKED FAITHFULNESS CONSTRAINT:

Phonotactic constraint: no word-final voiced obstruents

Input: /tab/

a.	[tap]	if the lowest faithfulness constraint is	IDENT-[voice]
b.	[tam]		IDENT-[nasal] / [son]
c.	[taw]		IDENT-[approximant]
d.	[tabə]		DEP-V
e.	[ta]		MAX-C
f.	[bat]		LINEARITY

Steriade's observation, however, is that only devoicing (36a) is attested as a response to a constraint againt final voiced obstruents. This is completely unexpected in the current state of the theory and she refers to this situation as the Too-Many-Solutions Problem. Her answer to it is the P-map and the correspondence contraints its projects. The claim is that only devoicing is attested because it involves the smallest modification of the input. That is, the pair [tab]-[tap] is perceptually more similar than any other input-output pair in which the output conforms to the phonotactics: [tab]-[tam], [tab]-[tabə], [tab]-[ta], etc.

To show this, however, we have to compare the distinctiveness of contrasts that differ over multiple dimensions. For example, to conclude that the pair [tab]-[tap] is more similar than the pair [tab]-[tabə], we have to determine that the contrast between [b] and [p] in the context [a]__# is less distinctive than the contrast between Ø and [ə] in the context [b]__#. From this comparison we derive the following constraint ranking: DEP-V/C__# >> IDENT-[voice]/V__#.

This is clearly a more complex case than the one used to solve the consonant deletion problem above and which resulted in the ranking in (27), extended to those

in (29). These rankings are based on comparisons which involve the same contrast (C vs. Ø) in different contexts, or different contrasts (e.g. C vs. stops) in the same context. What we know about the acoustics and the perception of consonants allows us to establish with a reasonable degree of confidence a hierarchy of distinctiveness among different contexts or contrasts, when the other variable is held constant. The idea was not to compare different repair strategies, that is consonant deletion vs. something else, but rather the same process in different situations. In contrast, the voicing problem just described requires that we compare different contrasts in different contexts, a much more complicated task, the goal being to establish a hierarchy among distinct repair strategies.

We will not have to perform multidimensional comparisons in this dissertation, nor establish perceptually-motivated rankings between different types of faithfulness constraints. In fact, unlike in the voicing case, there is no single process designated as the optimal repair for phonotactic constraints against perceptually weak consonants: both consonant deletion and vowel epenthesis are widely attested, and it does not seem that DEP and MAX should be ranked in the way IDENT-voice and DEP were ranked above. Yet in her discussion of the various solutions to final voiced obstruents, Steriade (1999d) cites work by Fleischhacker (2000c), who compares consonant deletion and vowel epenthesis as strategies to avoid consonant clusters. In a psycholinguistic experiment, English speakers had to judge whether *hef* or *heftə* sounds more similar to a reference term *heft*. The form involving consonant deletion, *hef*, was rated as more similar to *heft* than the form with an epenthetic vowel *heftə*. This leads to the prediction that final clusters of this type should always be repaired by deletion rather than epenthesis, given the corresponding fixed ranking DEP-ə/C__# >> MAX-C/C__# that can be derived from the similarity judgments. This prediction is contradicted by numerous cases of epenthesis, from which I conclude that either Fleischhacker's result cannot be generalized or that auditory similarity is irrelevant in choosing between epenthesis and deletion in the avoidance of consonant clusters.[22] It remains to be seen to what extent this conclusion weakens Steriade's proposal for the voicing case. I leave this issue open and remain agnostic on whether and to what extent multidimensional comparisons between different repairs should be performed and determine the ranking between distinct faithfulness constraints. In the mean time, it should be clear that I adopt the idea of constraint ranking based on comparisons of distinctiveness of

[22]In section 7 on cluster simplification, Steriade suggests that "the choice between V insertion and C deletion might remain free in resolving a size-of-cluster violation", on which I agree. But this claim can be contrasted with the results of Fleischhacker's study just presented, from which Steriade derives the ranking DEP(ə vs.Ø) >> MAX(C vs. Ø). This ranking could be taken to suggest that deletion should be favored over epenthesis in cluster reduction, and it is not clear to me why Steriade does not make this inference.

contrast only for a given repair, in order to determine what segment or portion of the string will be affected, and not to choose between repairs.

3.2.4. LIMITING THE ROLE OF PHONETIC GROUNDING

The perception-based approach developed here implies a view of the relationship between phonetics and phonology by which the former directly constrains the latter. This functionalist orientation in phonological theory has become prominent in recent years; Hayes (1999), for example, claims that "virtually all of segmental phonology (...) is driven by considerations of articulatory ease and perceptual distinctness". This view has not met with unanimity, and several researchers remain sceptical of the integration of functional, notably phonetic, factors in synchronic grammars (e.g. Ohala 1997; Hyman, to appear; Hale & Reiss 2000; Hansson 2000). These authors rather believe that phonetic determinism is only relevant in sound change and acquisition, but that synchronic grammars are formal systems which are subject to different principles. To the extent that synchronic processes are phonetically natural, this is considered a result of history and the acquisition process, not a property of phonological systems constrained by phonetic determinism.

Hyman (to appear) and Hale & Reiss (2000) in particular point to the existence of synchronic phenomena that are phonetically unnatural. Sound patterns interact with independent factors, such as borrowings, analogy, restructuring, and the result may be unnatural on articulatory or perceptual grounds. Yu (2000), for instance, describes a process of voicing in coda position found in Lezgian, which is quite unexpected from the point of view of universal phonetics. The existence of such processes leads to the inclusion of an arbitrary component in the grammar, that is one that is not functionally motivated. But once the necessity of an arbitrary grammatical component is acknowledged, conceptual economy argues for a view of grammar that comprises *only* arbitrary processes. As Hale & Reiss (2000) put it:

> [A grammar that has an arbitrary component and a nonarbitrary one] is empirically nondistinct from the theory we propose (...), which posits that *all* grammatical computations are arbitrary with respect to phonetic substance. (...) Since [we] must adopt a model which allows arbitrary phenomena (...), the addition to the theory of a special subcomponent to account for alleged "non-arbitrary" phenomena violates Occam's Razor. [their emphasis] (p. 161)

> Phonology is not and should not be grounded in phonetics since the facts which phonetic grounding is meant to explain can be derived without reference to phonology. Duplication of the principles of acoustics and acquisition constitutes a violation of Occam's razor and thus must be avoided. (p. 162)

As is often the case, I suggest that the solution lies neither in the all-phonetic approach nor in the all-arbitrary one. I see no reason why acknowledging the existence of phonetically unnatural processes should lead one to completely exclude phonetic grounding from phonology. Importantly, the conceptual economy argument brought by Hale & Reiss to evacuate phonetics from synchronic grammar seems to hold only if ones assumes, as they apparently do, that constraints are innate. I do not make such an assumption, but rather believe that constraints are built by language learners in the course of acquisition. What may be innate is only a constraint-building mechanism. Under this view, it seems difficult to consider formal phonology and acquisition to be two completely separate components of language, as is done by Hale & Reiss.

I argue that perception plays a direct role in the application of deletion and epenthesis processes. I also believe that grammars have to accomodate arbitrary phenomena. An obvious question, then, is: What is the division of labor between the arbitrary and functionally-motivated components of grammars, specifically phonology? I see two plausible options at this point, whose value will be determined by further research. First, notice that almost all the patterns examined in this dissertation and brought in support of the perceptual approach are variable ones. These include: consonant deletion in Hungarian, English, Icelandic, Catalan, Marais-Vendéen, and Québec French, as well as vowel epenthesis in French and Picard, and consonant deletion and vowel epenthesis in Basque (some of these cases will be examined in the following chapters). It could be that the role of functional motivations is synchronically limited to variable phenomena, in which direct comparisons between forms with different perceptual and articulatory properties can be made. The phonetic motivation, however, could be lost when processes become categorical. Under this view, final obstruent devoicing, for instance, could be considered an arbitrary process for kids learning German or Russian, but schwa insertion in French would be directly constrained by perception.[23]

Alternatively, phonetically-motivated constraints in phonology could be viewed as default ones, that is constraints that are more readily available to learners

[23]Note that variable phenomena cannot be dismissed from synchronic grammars as change in progress. The French schwa has been variable for centuries.

in the process of grammar building. Arbitrary constraints would only emerge as a fall-back option when required by data that are not amenable to a functional account. It is not implausible to think that functional constraints would be constructed more easily than arbitrary ones since the former are grounded in and constrained by physical reality, whereas the latter are completely dependent on language-specific and process-specific data. Interestingly, this view of grammar can be tested psycholinguistically. We expect default elements to be acquired earlier than more marked ones. If the proposed split between the functional and arbitrary components of grammar is correct, we expect that children will generally master functionally-motivated processes before arbitrary ones. This remains to be investigated.

This discussion makes it clear that I am not claiming that all segmental phonology is phonetically-driven; I am only arguing for the existence of perceptually-based constraints in phonology. These constraints could have a more or less limited role in the grammar, depending on the correct division of labor between the arbitrary and non-arbitrary components. If functional constraints are limited to variable processes, their role in the grammar may be rather reduced; if they correspond to default options, much of phonoloy may be functionally-motivated, with the arbitrary part playing a subsidiary role.

3.2.5. Variation in Optimality Theory

As mentioned in the previous section, variation and frequency/likelihood are omnipresent in the processes investigated in this dissertation. This requires that we spend some time discussing the treatment of these aspects in phonological theory, particularly in Optimality Theory.

Variation has been a neglected area of phonological theory. Optional rules have been used to express non-categorical processes, but notions of frequency/likelihood or preference have been to a large extent relegated to the sociolinguistic domain. Yet a large portion of phonological variability is driven by the same factors that underlie categorical processes. I believe one of the major advantages of Optimality Theory over previous rule-based approaches is precisely its ability to model variation and derive hierarchies of frequency or gradient well-formedness.

Categorical phenomena are straightforwardly derived in OT by strict constraint ranking. Optionality is standardly handled by constraint ties (although these are excluded under the most constrained version of the theory), but this approach is too restrictive to account for all cases of variation. See e.g. Anttila (1997),

Côté (1999), and Auger (2000) for patterns that cannot be accounted for with tied constraints. A more powerful solution becomes available if we adopt Anttila's (1997) view of grammars as partial orders.[24] This approach abandons the assumption that all constraints are ranked (possibly tied) with respect to all others, and allows constraint rankings to remain underdetermined. A grammar may then be compatible with many different full or total rankings. These distinct rankings may, in turn, yield different outputs (for a given input). This is how variation (and optionality) is generated by the system.[25]

An additional assumption of Anttila is that frequency of use or the relative well-formedness of a given output should reflect the probability that it be generated by the grammar. This probability corresponds to the proportion of the possible rankings that yield this output. The following abstract example illustrates the mechanism. Suppose three constraints A, B, C, and a grammar consisting in the unique ranking A >> B. Three possible total orders of the constraints A, B, C are compatible with this grammar: A>>B>>C, A>>C>>B, and C>>A>>B. Suppose that for some input I the first ranking yields an output O_1, and the last two a different output O_2. This grammar then predicts variation / optionality between O_1 and O_2. In addition, it is expected that O_1, which is generated by one ranking out of three, will surface one third of the time, while O_2 will be used two thirds of the time.

I adopt Anttila's view of grammars as partial orders, as well as the relation between the frequency/likelihood of a form and the probability that it be selected by the constraint ranking. This relation, however, will not be interpreted in a strict fashion. That is, I will not expect these probabilities to be equal to actual frequencies of use, but only to reflect hierarchies of frequency or likelihood. If an output O_2 is generated by more rankings than an output O_1, I will not go much further than the prediction that O_2 is prefered to, or more likely than, O_1. The reasons for this loosening are twofold. First, in most cases I do not know the actual frequencies of use, which makes it impossible to test the stricter version of Anttila's theory. Second, actual frequencies are usually influenced by non-grammatical factors, which lead to deviations with respect to what is expected from the constraint system alone. I expect, however, that the order of preference of the forms is preserved.

---

[24] Reynolds's (1994) floating constraints can be viewed as a sub-case of Anttila's partial orders.

[25] See Boersma (1998), Boersma & Hayes (1999), and Hayes (2000) for different approaches to variation in Optimality Theory, which I will not consider here.

3.3. APPLICATIONS

3.3.1. LENAKEL VOWEL EPENTHESIS

Vowel epenthesis in Lenakel is a good example to provide a first illustration of the functioning of the constraint system I propose. It specifically highlights the role of the markedness constraints. This process displays several of the factors identified as relevant – contrast, edge effects, adjacent vowels – and also shows a certain amount of variation. Yet the pattern is relatively simple and immune from independent intricacies.

The Lenakel epenthesis pattern can be described as follows (Lynch 1978; Blevins 1995; Kager 1999). An epenthetic vowel [ɨ] or [ə], depending on the preceding consonant, is automatically inserted in sequences of two consonants word-initially (37a-b) and finally (37c-d), and in clusters of three consonants word-internally (37e-f). The epenthetic vowel (underlined in the examples below) is inserted between the second and third consonant word-internally, and between the two consonants at word edges.[26]

(37) OBLIGATORY VOWEL EPENTHESIS IN LENAKEL:

a.	/t-n-ep-kɨn/	→	[tɨ̲nébgən] - [dɨ̲nébgən]	'you will eat it'
b.	/t-r-ep-ol/	→	[tɨ̲rébɔl]	'he will then do it'
c.	/r-ɨm-ɨgn/	→	[ɾɨmə́ŋə̲n]	'he was afraid'
d.	/n-əm-əpk/	→	[nɨmə́bə̲kʰ]	'you (sg.) took it'
e.	/ɨs-ɨt-pn-aan/	→	[ə̀sɨdbə̲nán]	'don't go up there'
f.	/k-ar-pkom/	→	[kɑrbə̲́gɔm]	'they are heavy'

There is one exception to this pattern: glide+consonant sequences are tolerated word-finally:[27]

---

[26]I adapt Lynch's (1978) transcription in the following way, in conformity with the IPA: [y] is replaced by [j]; [ř] is described as a flap and is replaced by [ɾ]; [v] is described as a high central glide noted [ɨ̯] and this is the symbol I adopt.

[27]In fact, Lynch (1978: 15) describes this exception as follows: "when two consonants come together at the beginning or the end of a word, [ɨ] is inserted between them provided that neither is a glide". This characterization is met in principle in four different cases, the combinations C+G or G+C word-initially or word-finally. In fact I have found on the surface only the word-final G+C combination, illustrated in (38). Some combinations were not found in the data provided, especially initial G+C clusters. Interestingly, Bell & Hooper (1978: 11) claim that these are unattested crosslinguistically. Others merged into a single consonant by independent processes which I disregard here: glides becoming secondary articulations (i) or /h/ deleting while devoicing the adjacent consonant (ii). Note that Lynch includes /h/ in the set of underlying

(38) NO EPENTHESIS IN /G+C/ CLUSTERS WORD-FINALLY:

a.	/pwapwauk/	→	[pʷɒbʷɒ́wkʰ]	'butterfly'
b.	/aik/	→	[ajkʰ]	'to swim'

In addition to the obligatory cases of epenthesis in (37), [ɨ]/[ə] is optionally inserted between any two consonants word-internally (39).[28] Insertion becomes obligatory, however, between two identical consonants across a morpheme-boundary (40).[29]

(39) OPTIONAL EPENTHESIS IN INTERNAL /CC/ CLUSTERS:

a.	/r-am-alfa/	→	[ɾɑmálfɑ] / [ɾɑmálɨ̲fɑ]	'he is lazy'
b.	/nɨmr-n/	→	[nɨ́mɾɨn] / [nɨ́mə̲ɾɨn]	'his eyes'

(40) OBLIGATORY EPENTHESIS BETWEEN IDENTICAL CONSONANTS:

a.	/i-ak-kɨn/	→	[yɑgə̲́gən]	'I eat it'
b.	/t-r-rai/	→	[tɨɾɨ̲ɾɑy] / [dɨɾɨ̲ɾɑy]	'he will write'

I analyze these facts in the following way. Consonants in Lenakel must surface with an adjacent vowel. This follows from a high-ranked general C↔V constraint. This constraint applies exceptionlessly word-internally and word-initially. However, it is relaxed for word-final consonants that are preceded by a glide. I interpret the latter condition as a requirement that the consonant contrasts in the feature [vocoid] with an adjacent segment. Consonants that agree in this feature

---

glides, along with /w/ and /ɨ̯/; [j] is assumed to only surface as a reflex of /i/ in certain positions. In the case of /C+ɨ̯/, normal epenthesis applies, contrary to Lynch's generalization (iii).

(i)	/amnuumw/	→	[ɑmnúmʷ]	'to drink'
	/t-i-ɨs-ɨ̯a-aan/	→	[tʲèsɨ̯eán] / [dʲèsɨ̯eán]	'I won't come'
(ii)	/rho/	→	[ɾ̥ó]	'he hit it'
	/r-am-awh/	→	[ɾɑmáw̥]	'she is weaving'
(iii)	/m-ɨ̯ɨn/	→	[mə̀ɨ̯ɨn]	'and-go'
	/r-ɨ̯a/	→	[rɨ̯ɨa]	'3s-come'

Even if glide-containing sequences other than final G+C sequences turned out to be attested, it would not be a problem for the analysis sketched here.

[28]As long as the first consonant is not a glide and the following vowel is unstressed. I leave these additional conditions aside for the purposes of this illustration.

[29]When both consonants are coronals deletion of the first consonant occurs rather than epenthesis. Certain verbal prefixes, however, like /t/ and /r/ in (40b), cannot delete. When they are followed by an identical consonant, like the /r/ in the same example, then the general epenthesis rule applies. I leave a unified analysis of coronal deletion and vowel epenthesis for future research.

with their neighboring consonants invariably trigger epenthesis.[30] To account for these generalizations I design the following markedness constraints:

(41) RELEVANT MARKEDNESS CONSTRAINTS:

a. C|ø ↔ V
A PW-internal consonant (which is adjacent to no prosodic boundary) is adjacent to a vowel.
b. $_{PW}$[C ↔ V
A consonant that is preceded by a PW boundary is adjacent to a vowel.
c. C]$_{PW}$ (AGREE=[vocoid]) ↔ V
A consonant that is followed by a PW boundary and that agrees in [vocoid] with a neighboring segment is adjacent to a vowel.
d. C]$_{PW}$ ↔ V
A consonant that is followed by a PW boundary is adjacent to a vowel.

By the dominance condition in (19), we can establish the inherent rankings in (42) between these constraints; the reader may also refer to the rankings in (22).

(42) INHERENT RANKINGS BETWEEN THE MARKEDNESS CONSTRAINTS IN (41):

a. C|ø ↔ V >> $_{PW}$[C ↔ V
b. C|ø ↔ V >> C]$_{PW}$ ↔ V
c. C]$_{PW}$ (AGREE=[vocoid]) ↔ V >> C]$_{PW}$ ↔ V

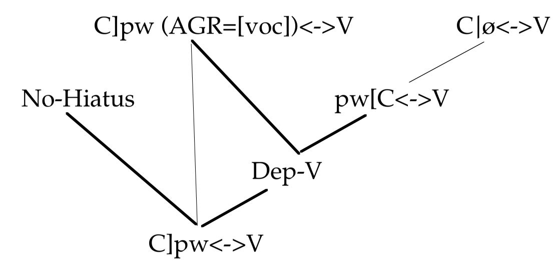
Our task is now to rank DEP-V within this web of markedness constraints. The three constraints in (41a-c) are unviolated in the language and must dominate all constraints against vowel epenthesis. But DEP-V outranks C]$_{PW}$↔V, since epenthesis does not apply word-finally in the clusters that are not subject to the higher-ranked C]$_{PW}$(AGREE=[vocoid])↔V. This mini-grammar is given in graphic form in (43) and illustrated in the tableau in (44), with examples from (37) and (38). In this and all following graphics thick lines are used to indicate language-specific rankings determined on the basis of the available data, whereas thin lines indicate fixed inherent rankings.

The issue of the site of epenthesis obviously arises here. In internal three-consonant clusters, the vowel is inserted between the second and the third consonant, while it always occurs between the two consonants at edges. I disregard this issue in this first step and consider only the candidates with the correct

[30]I assume that the final consonant is a non-glide. If glide+glide sequences are tolerated as well, the generalization would be that it is agreement in [-vocoid] specifically rather than [vocoid] that systematically triggers epenthesis.

placement of the epenthetic vowel. This problem will be addressed below. Finally, in /aik/, the last example in the tableau, I assume that the faithful candidate [aik] is excluded by a constraint against hiatus, which must at least dominate C]$_{PW}$↔V. I disregard the rules of alternation between high vowels and glides.

(43) PARTIAL GRAMMAR OF LENAKEL I:



(44) EPENTHESIS AND NON-EPENTHESIS IN LENAKEL:

a. /t-r-ep-ol/	NO-HIATUS	C]$_{PW}$ (AGR=[voc])↔V	C|ø↔V	$_{PW}$[C↔V	DEP-V	C]$_{PW}$ ↔ V
→ tɨ̠rɛbɔl					*	
trɛbɔl				(t) !		
b. /n-əm-əpk/						
→ nɨmə́bə̠kʰ					*	
nɨmə́pkʰ		(kʰ) !				*
c./k-arp-kom/						
→ kɑrbə̠́gɔm					*	
kɑrbgɔm			(b) !			
d. /aik/						
ajVkʰ					* !	
→ ajkʰ						*
aikʰ	* !					

In (44a) the faithful candidate [trɛbɔl] (disregarding vowel quality and intervocalic voicing) violates $_{PW}$[C↔V, which requires every word-initial consonant to be adjacent to a vowel. The epenthesized candiate [tɨ̠rɛbɔl] violates DEP-V and is the winning output since DEP-V is ranked lower than $_{PW}$[C↔V. The situation in (44c) is similar, except that the markedness constraint violated by the faithful candidate is C|ø↔V rather than $_{PW}$[C↔V. (44b,d) contain underlying word-final two-consonant clusters. In (b) vowel insertion applies, in (d) it does not. The difference between these two cases lies in the nature of the cluster. The two segments in the sequence [pkʰ] (44b) share the same value for the feature [vocoid]. The final [kʰ] agrees in [vocoid] with the preceding consonant and is not adjacent to a vowel, in violation of the higher-ranked constraint $C]_{PW}$ (AGREE=[vocoid])↔V, which dominates DEP-V. Unlike [pkʰ], the sequence [jkʰ] (44d) displays a contrast in the feature [vocoid] and only yields a violation of the general lower-ranked constraint $C]_{PW}$↔V.

Let us now look at word-internal two-consonant sequences. We have seen that epenthesis in such medial clusters is optional in the general case, but obligatory between two identical consonants. The relevant constraints to deal with these facts are given in (45), and the derivable inherent rankings that involve them in (46).

(45) ADDITIONAL MARKEDNESS CONSTRAINTS:
- a. C|ø (AGREE=∀F) → V
  A word-internal consonant (that is next to no prosodic boundary) and that agrees in all features with an adjacent segment is followed by a vowel.
- b. C|ø → V
  A word-internal consonant (that is next to no prosodic boundary) is followed by a vowel.

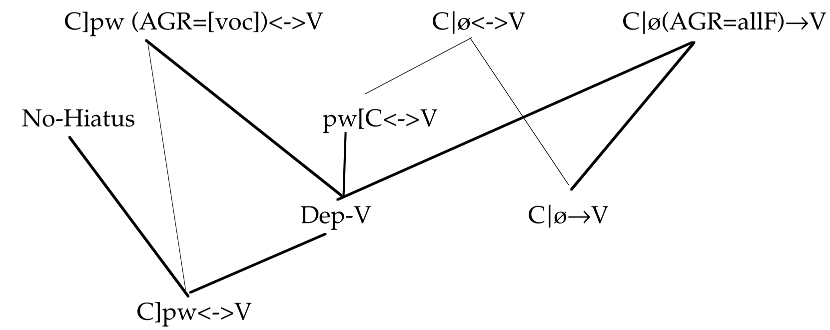
(46) ADDITIONAL INHERENT RANKINGS:
- a. C|ø (AGREE=∀F) → V >> C|ø → V
- b. C|ø ↔ V >> C|ø → V

C|ø (AGREE=∀F)→V is violated in cases of two identical consonants word-internally. This constraint is undominated in Lenakel and forces epenthesis. The ranking between DEP-V and the lower-ranked C|ø→V remains undetermined, since we find variation between forms that violate C|ø→V ([...VCCV...]) and forms that violate DEP-V ([...VCV̲CV...]). This is illustrated in the tableau below with forms from (39) and (40). The mini-grammar in (43) is augmented as in (48).

(47) EPENTHESIS AND NON-EPENTHESIS IN WORD-INTERNAL CC CLUSTERS:

a. /r-am-alfa/	C|ø↔V	C|ø(AGR=∀F)→V	DEP-V	C|ø→V
→ ɾamál̲i̲fa			*	
→ ɾamálfa				(l)
b. /i-ak-kɨn/				
→ yaga̲̲ə̲gən			*	
yaggən		(g) !		*

(48) PARTIAL GRAMMAR OF LENAKEL II:



Let us now consider the issue of the site of epenthesis. I assume that the word-internal placement of epenthesis between the second and third consonants in three-consonant clusters is due to an alignment constraint requiring every consonant to align with the left edge of the prosodic word (49a), which dominates the corresponding constraint favoring alignment to the right (49b). These constraints are evaluated gradiently in terms of the number of segments that intervene between a consonant and the edge.

(49) ALIGNMENT CONSTRAINTS DETERMINING THE LOCUS OF EPENTHESIS:
- a. ALIGN-L (C,PW): A consonant aligns with the left edge of a PW.
- b. ALIGN-R (C,PW): A consonant aligns with the right edge of a PW.
- c. ALIGN-L (C,PW) >> ALIGN-R (C,PW)

(50) DETERMINING THE LOCUS OF EPENTHESIS WORD-INTERNALLY:

/k-ar-pkom/	C|ø↔V	DEP-V	ALIGN-L (C,PW)	ALIGN-R (C,PW)
→ karbə̲gɔm		*	0+2+3+5+7=17	0+2+4+5+7=18
karə̲bgɔm		*	0+2+4+5+7=18 !	0+2+3+5+7=17
karbgɔm	* !		0+2+3+4+6=15	0+2+3+4+6=15

At word edges epenthesis is always medial. Medial epenthesis (i.e. between the two consonants) is correctly predicted by the alignment constraints word-initially, but not word-finally, where we rather expect final epenthesis. Given an initial #CC sequence, left-alignment is better achieved in #CV̲C than in #V̲CC, which is what we find in Lenakel. The opposite holds with final CC# inputs: CCV̲# satisfies left-alignment better than CV̲C#. Yet it is the latter output that surfaces in Lenakel. As discussed in Blevins (1995), this is a problem for the directionality approach to the location of epenthesis, which carries over to the alignment one. This pattern – medial epenthesis at both edges, irrespective of the prefered site word-internally – is not exceptional and is also found for example in Chukchi.

A somewhat unexpected but welcome result of the system of markedness constraints we have developed is that they automatically derive the Lenakel/Chukchi pattern of epenthesis in edge clusters, without the need for additional constraints. This follows from the observation, encoded in the ranking, that consonants are more easily tolerated at edges than domain-medially, everything else being equal. Epenthesis takes advantage of this and preferably applies in a way that puts the consonants at an edge rather than medially. The mini-grammar in (48), with the constraint C|ø→V playing the crucial role, yields the desired result, as shown in the tableau below, which concludes our first case study.

(51) DETERMINING THE LOCUS OF EPENTHESIS AT WORD EDGES:

a. /t-r-ep-ol/	C]PW (AGR=[voc])↔V	C|ø↔V	PW[C↔V	DEP-V	C|ø→V
→ ti̱rɛbɔl				*	
i̱trɛbɔl				*	(t) !
trɛbɔl			(t) !		
b. /n-əm-əpk/					
→ nɨmə́bə̱kʰ				*	
nɨmə́bkə̱				*	(b) !
nɨmə́bkʰ	(kʰ) !				

3.3.2. SRANAN CONSONANT DELETION

Alber & Plag (1999) discuss vowel deletion and consonant epenthesis in the formation of Sranan, an English-based creole language spoken in Surinam. Consonant clusters in the source language were extensively simplified in Sranan, usually by deletion, except word-finally, were we often find vowel epenthesis (paragoge). I am interested here in word-internal consonant deletion. It applies quite systematically to sequences of two consonants composed of obstruents and nasals. Liquids that are not intervocalic are subject to more varied and partly unpredictable processes: deletion, metathesis with an adjacent consonant or vowel, epenthesis, preservation. I focus here on clusters that do not involve liquids. Consider the data in (52) to (54).

(52) SRANAN ADAPTATIONS OF STOP+FRICATIVE AND STOP+NASAL CLUSTERS:

	English word	*Sranan adaptation*
a.	curtsey	kosi
b.	goodmorrow	kumara
c.	goodnight	kuneti

(53) SRANAN ADAPTATIONS OF FRICATIVE+STOP AND NASAL+STOP CLUSTERS:

a.	master	masra, masera
b.	nasty	nasi
c.	sister	sisa
d.	softly	safri
e.	remember	memre, memere
f.	something [mθ]	sani

(54) SRANAN ADAPTATIONS OF STOP+STOP CLUSTERS:

a.	doctor	datra
b.	sit down	sidon

In (52) we have English forms containing stop+fricative (a) and stop+nasal (b-c) clusters. In all cases only the second consonant is retained in Sranan. (53) shows examples of fricative+stop (a-d) and nasal+stop (e-f) sequences. Here it is the first consonant that shows up in the adapted form. The generalization is that stops preferentially delete over non-stops. It has been noticed in the discussion of faithfulness constraints, however, that in VCCV sequences, it is typically the first consonant that deletes. This generalization can be observed in clusters composed of two stops, in which case it is the second stop that is retained (54). This deletion pattern shows that the tendency to delete the first consonant in an intervocalic two-consonant cluster can be overriden by conflicting factors, here the stop or non-stop nature of the consonants.[31]

Alber & Plag do not extract these generalizations from the data. They notice variation in the position of the deleted consonant, but cannot account for it and simply leave the issue open. This pattern, however, receives a natural and simple

[31] I suspect that the position of stress is relevant in the data in (52)-(54), but the data in the paper do not allow us to test this hypothesis. It could be that retention of the postvocalic rather than the prevocalic consonant occurs only in the context v́ccv̆, where the stable postvocalic consonant is adjacent to a stressed vowel, while the deleted stop is followed by an unstressed one. Adding the effect of stress to the analysis would not be problematic. The cues present in the transition to or from a stressed vowel are better than those to or from an unstressed one, since stressed vowels are generally associated with higher amplitude. This contrast could be easily integrated into our markedness and faithfulness constraints.

explanation in the framework developed here. The distinctions in (52)-(54) follow straightforwardly from the perceptually-motivated faithfulness constraints in (29a-b), repeated below. The deletion of postvocalic consonants is prefered over that of prevocalic ones, due to the better cues present in the CV transition. The deletion of stops is also more likely than that of non-stops because of the weakness of their internal cues.

(29) RELEVANT FAITHFULNESS CONSTRAINTS IN SRANAN:
a. MAX-C/__V >> MAX-C/V__ >> MAX-C
MAX-C/__V Do not delete a consonant that is followed by a vowel.
MAX-C/V__ Do not delete a consonant that is preceded by a vowel.
b. MAX-C(-stop) >> MAX-C
MAX-C(-stop) Do not delete a consonant that is not a stop.

By assuming the simple ranking in (55), we derive the data in (52)-(54), as shown in the tableau in (56). This ranking interacts with the constraint C→V, which is taken to motivate medial consonant deletion in Sranan. To account for the data in (52)-(54) C→V must at least dominate MAX-C/__V.

(55) RANKING BETWEEN THE FAITHFULNESS CONSTRAINTS:
MAX-C(-stop) >> MAX-C/__V >> MAX-C/V__

(56) CONSONANT DELETION IN SRANAN:

a. Eng. *goodnight*	C→V	MAX-C(-stop)	MAX-C/__V	MAX-C/V__
kudneti	(d) !			
→ kuneti			*	
kudeti		* !		*
b. Eng. *sister*				
sista	(s) !			
→ sisa			*	
sita		* !		*
c. Eng. *sit down*				
sitdon	(t) !			
→ sidon				*
siton			* !	

In all these examples the faithful output (in terms of the size of number of consonants, irrespective of other phonological processes) violates C→V and one of the consonants deletes. When the cluster contains a stop and a non-stop (56a-b), the

stop deletes, whether it appears in cluster-initial or cluster-final position, due to the high-ranking of MAX-C(-stop), which prohibits the deletion of non-stops. In clusters composed of two stops (56c), the first one is dropped since the constraint against the deletion of prevocalic consonants MAX-C/__V dominates that against the deletion of postvocalic ones MAX-C/V__.

3.4. CONCLUSIONS

In this chapter I have introduced the theoretical apparatus designed to account for the empirical generalizations presented in chapters 1 and 2 concerning deletion and epenthesis. The constraint system developed in section 3.2 rests on the perceptual motivations that underlie these generalizations, as explained in section 3.1. Both faithfulness and markedness constraints are taken to encode the notion of perceptibility and the desirability for segments to be perceptible. Faithfulness constraints ensure that consonant deletion targets the auditorily weakest consonants, and vowel insertion maximizes auditory similarity between input and output. Markedness constraints establish a correlation between the degree of perceptibility of consonants and their relative markedness. This theoretical orientation raises the more general issue of the role of perception, phonetic grounding, and other functional motivations in phonology, and I have argued for a mixed view of grammars as comprising both functionally-motivated and arbitrary processes, although the exact domains of these two components remain to be identified. Additionally, the treatment of variation in Optimality Theory, seen as a major advantage of this theoretical approach, has been addressed, as most patterns analyzed in the remainder of this dissertation are variable ones. Finally the constraint system was illustrated in the analysis of two simple cases of consonant deletion in Sranan and vowel epenthesis in Lenakel, which highlight the role of perceptually-motivated faithfulness and markedness constraints, respectively. The functioning of the constraint system will be more fully appreciated in the following two chapters, which expand on the role of contrast and edge effects in deletion and epenthesis.

APPENDIX:
ADDITIONAL PATTERNS SHOWING THE SPECIAL STATUS OF STOPS

I provide here additional patterns that exclusively or more specifically target stops. I present these patterns to further illustrate the special status of these consonants and their increased vulnerability in the absence of adjacent vowels. But I will not refer to them in the rest of the dissertation. Other cases are also described or mentioned in Steriade (1999d, to appear), among them Colloquial Latin (Niedermann 1953) and Dihovo Macedonian (Groen 1977).

There is one case of consonant deletion (Farsi) and, more interestingly, two cases of metathesis. Metathesis has not been mentioned as a possible repair strategy for complex consonant clusters. It is indeed marginal in comparison with deletion and epenthesis, but the Lithuanian and Singapore English examples clearly show how metathesis can be used productively to avoid stops in perceptually weak positions. These two cases were discussed in Côté (1997a). The Lithuanian one is analyzed in the same terms but independently by Steriade (to appear).

A. <u>Metathesis in Lithuanian</u>

In Lithuanian, verbs that end in a fricative-stop cluster undergo metathesis when followed by a consonant-initial suffix (Kenstowicz 1971; Ambrazas 1985: 60; Mathiassen 1996: 26):

(1) STOP-FRICATIVE METATHESIS IN LITHUANIAN:

	URs	*+Vowel*		*+Consonant*	
/-sk/	/dresk-/	dreskia	'he/they tear(s)'	dreksti	'to tear'
/-zg/	/mezg-/	mẽzga	'he/they knot(s)'	mègzdamas	'knotting'
/-ʒg/	/dʒerʒg-/	dʒerʒgia	'he/they scrape(s)'	dʒergʒti	'to scrape'

I interpret this process in the following way. When the last stop of the stem precedes a vowel, it benefits from the strong contextual cues present in the transition to the vowel. If the last stop preceded a consonant, it would find itself in an inter-consonantal weak position. Metathesis of the stop and the fricative then allows both consonants to be sufficiently salient. On the one hand, the stop is strengthened by now being in post-vocalic position. On the other hand, fricatives remain perceptually salient even in inter-consonantal position.

B. <u>Metathesis in Singapore English</u>

In Singapore English (Mohanan 1992), final /-sp/ metathesizes to /-ps/. For example, *crisp* is pronounced [krips], *grasp* [grɑ:ps]. As in Lithuanian above, this process allows both consonants to remain acoustically salient: /p/ gains vocalic transitions from the preceding vowel, while /s/ is strong enough by itself.

C. <u>Consonant deletion in Farsi</u>

Colloquial Farsi (Darzi 1991; Mahootian 1997) productively simplifies certain consonant clusters, in particular word-finally. We can distinguish three distinct deletion processes:

1. Deletion of /ʔ/ and /h/. This occurs in numerous positions, especially in clusters but also word-finally after a vowel and even intervocalically. I disregard these cases of deletion, which involve a restricted class of glottal consonants.

2. Deletion of /r/ after an obstruent word-finally, e.g. /fekr/ → [fek] 'thought. I suspect this process is motivated by the SSP.

3. Deletion of stops in C__C and C__## contexts. This is what interests me here. Mahootian (1997) states that stop deletion applies (optionally) to /t/ after a coronal fricative /s, ʃ/ (2) and /d/ after /n/ (3).

(2) /t/ DELETION AFTER A CORONAL FRICATIVE IN FARSI:

a.	/dæst/	[dæs]	'hand'
b.	/dæstgire/	[dæsgire]	'handle'
c.	/dæstgah/	[dæsgah]	'equipment'
d.	/bist/	[bis]	'twenty'
e.	/rastgu/	[rasgu]	'truthful'
f.	/moʃt/	[moʃ]	'fist'
g.	/ængoʃtnema/	[ængoʃnema]	'notorious'

(3) /d/ DELETION AFTER /n/:

a.	/qænd/	[qæn]	'sugar'
b.	/kond/	[kon]	'slow'
c.	/mund-ænd/	[mundæn]	'they stayed'
d.	/mi-neveʃt-ænd/	[mineveʃtæn]	'they were writing'
e.	/tʃænd-ta/	[tʃænta]	'how many'
f.	/bolænd-qæd/	[bolænqæd]	'tall'

But according to Darzi (1991), the process extends at least to /d/ after /z/ (4), /t/ preceded by non-coronal fricatives (5), as well as stops at places of articulation other than coronal (6).

(4) /d/ DELETION AFTER /z/:

a. /mozd/	[moz]	'wage'
b. /dozd/	[doz]	'thief'

(5) /t/ DELETION AFTER A NON-CORONAL FRICATIVE:

a. /hæft/	[hæf]	'seven'
b. /gereft/	[geref]	'(he) got'
c. /loxt/	[lox]	'naked'
d. /saxt/	[sax]	'(he) built'

(6) NON-CORONAL STOP DELETION:

a. /xoʃk/	[xoʃ]	'dry'

First, the process appears to be restricted to stops. No cases of fricative or nasal deletion are reported, except in the isolated example /tʃeʃm/ 'eye', pronounced [tʃeʃ] (Mahootian 1997: 336). Final /m/ does not delete in other similar words – e.g. /pæsʃm/ 'wool' – or after other consonants – e.g. /esm/ 'name', /elm/ 'science', /hokm/ 'order' – even if the SSP is violated, as in the last two examples.

Stop deletion, however, is clearly dependent on contrast between the stop and the preceding consonant. But Darzi and Mahootian differ on the amount of contrast that is necessary to block deletion. According to Mahootian, only coronal stops that are homorganic with the preceding consonant delete. So a contrast in place of articulation prevents simplification.[32] In addition, stops are dropped only after consonants that contrast minimally in manner of articulation: nasals, which contrast only in [sonorant], and fricatives, which contrast in [continuant]. Stops seem to be stable after liquids, which contrast in both [sonorant] and [continuant], or in [sonorant] and [approximant] depending on the feature system one adopts. All the reduced clusters also show no contrast in voicing. Darzi is less restrictive with respect to place of articulation, and allows the deletion of stops that are not coronal and not

[32]The role of coronality is not clear. Is it the case that non-coronal consonants may not drop in the variety described by Mahootian, or are non-coronal stops disregarded because they are much less frequent, as is the case in English (see chapter 1, section 1.2.3.3.)? Recall that Darzi does allow deletion of non-coronal stops.

homorganic with the preceding consonant. The conditions on manner of articulation, however, are identical as in Mahootian.