The Role of Feature Parsing in Speech Processing and Phonology

David Gov, Mass General Hospital and Salem State College
Cheryl Zoll, Massachusetts Institute of Technology

This paper identifies the problem of feature parsing and explores the role of perceptual grouping processes in language processing and in a wide variety of phonological phenomena including assimilation, dissimilation and anti-gemination. We show that feature parsing is a fundamental component of speech perception, and that the threat of spurious feature association constrains the evolution of phonological systems.

1. Introduction

While it is universally recognized that segments are reducible to features, very little is known about how listeners integrate feature cues into segments. Recent work by Gov (2001, in press) suggests that this integration takes place at the level of abstract feature values, and that it is governed by perceptual grouping principles that reflect the salience and temporal distribution of feature cues in the speech signal. In this paper we explore ways in which this perceptual problem and the processes listeners use to resolve it shape a variety of phonological phenomena.

2. Feature Parsing

Feature parsing is the process of aligning feature cues recovered from the speech stream with abstract segmental representations. It poses an inescapable problem in speech perception because of the temporal distribution of feature cues. Feature cues associated with a given segment tend to be distributed over time due to a combination of articulatory and aerodynamic factors. The production of any segment involves the coordinated movement of multiple independent articulators. When the coordination of articulation is approximate, cues associated with different articulators are staggered in time. For example, the velopharyngeal gesture that gives [n] its nasality generally precedes the tongue tip gesture that gives it its place of articulation. Krakow (1993) demonstrates that these two gestures desynchronize to varying extents in different contexts. The distribution of feature cues is also influenced by aerodynamic factors. For example, the place of articulation of syllable-final stop consonants is marked by the movement of formants during voiced portions of the signal prior to closure. Place is also marked by the spectrum of the burst that follows the release of these segments. Evidence for trading relations between feature cues in segment
perception (c.f. Summerfield & Haggard, 1977) suggest that listeners integrate these multiple cues even though they arrive at different times.

3. Evidence from the processing of assimilated speech

The feature parsing problem becomes more complicated when one considers the complex mapping between features and segments. Goldsmith (1976) demonstrates that the traditional assumption of linear mapping between features and segments fails to account for a variety of phonological phenomena. The association of a single feature with multiple segmental positions in feature spreading is a case in point. In place assimilation a place feature spreads to a preceding segment that is itself already associated with its own place feature. In English, for example, a non-coronal place feature may spread to a preceding coronal. Hayes (1992) argues that in this case there is no delinking between the assimilated segment and its coronal place. This produces a segment with complex place. This analysis is consistent with articulatory data showing that coronal place assimilation generally produces a combination of partially reduced coronal closure and non-coronal closure (Barry, 1985; Nolan, 1992). Listeners therefore hear a segment with evidence for two places of articulation.

Psycholinguistic evidence supports Hayes’s analysis of coronal assimilation. Gow (in press; submitted) demonstrates that listeners who hear a token of a word like right in which the final segment has undergone labial assimilation show facilitated monitoring for a labial target (e.g. [b] in berries) that immediately follows it. Lahiri and Marslen-Wilson (1991) and Gow argue that this facilitation results from the association between the labiality of the assimilated segment and the labiality of the subsequent segment. Gow (in press, submitted) also demonstrates that listeners hearing this modified token of right show form priming for its underlying form (/ræɪt/), but not its apparent surface form (/ræɪp/). This appears to reflect the association between recovered evidence of coronality and the final segment of right. Combining these two results, one finds support for the claim that assimilated segments show complex place, and the hypothesis that listeners recover both of the underlying place features and successfully associate each feature with the correct segment.

This raises the question of how listeners successfully associate the right features with the right segments. Gaskell and Marslen-Wilson (1996; 1998) argue that listeners use knowledge of phonology to infer the underlying form of assimilated segments. Their broad claim is that listeners invert the assimilation rule. For example, a listener hearing [lɪm#b] might infer that [m] takes its place from the [b]. Since English only allows the assimilation of coronal segments, the listener might infer that the segment was underlyingly coronal. This type of inference is problematic. For example, it would lead a listener to reanalyze trim beard as [trɪm] beard and to reanalyze ripe berries as [ræɪt] berries. The results of several priming experiments suggest that listeners do not make this inference (Gaskell & Marslen-Wilson, 2001; Gow, in press, submitted).\(^1\)

---

1 Gaskell and Marslen-Wilson could refine their position and claim that listeners invert the assimilation rule only when confronted with multiple feature cues on a single segment. However, Gow (submitted) shows that this is insufficient.
Lahiri and Marslen-Wilson (1991) demonstrate that the mapping problem may be constrained by representational factors. They find that English-speaking listeners expect a nasal to follow a nasal vowel, but Bengali-speaking listeners do not. They attribute this difference to the claim that vowel nasality/orality is specified in Bengali, but not in English. Thus, English listeners who detect evidence of nasality in these vowels are unable to associate it with the vowel and are forced to associate it with a neighboring segment. A similar availability constraint operating at a lexical level may help explain why listeners hearing an item like [grim] (green with the final segment assimilating labial place) anticipate that a labial segment will follow. Here, associating the relatively weak word-final acoustic evidence for labiality with [m] would yield the non-word *green*. However, associating evidence for coronality with the same segment and associating evidence of labiality with another segment would lead to the access of a familiar word.

Gow (in press; submitted) provides evidence that listeners solve the feature parsing problem posed by English coronal place assimilation through other means when availability constraints do not point to unique pattern of association. In a series of form priming experiments he uses cross-splicing techniques to manipulate the phonemic context following tokens of words like *right* in which the final coronal segment has undergone labial assimilation to produce segments with formant frequency values intermediate between those of labials and coronals. The results of these studies are summarized in (1):

(1) Priming by assimilated tokens with context manipulated by cross-splicing (Gow, in press, submitted)

<table>
<thead>
<tr>
<th>Prime</th>
<th>RIGHT</th>
<th>RIPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. …raʃʃ# labial…</td>
<td>accessed</td>
<td>not accessed</td>
</tr>
<tr>
<td>b. …raʃʃ# coronal…</td>
<td>not accessed</td>
<td>accessed</td>
</tr>
<tr>
<td>c. …raʃʃ# Ø</td>
<td>accessed</td>
<td>accessed</td>
</tr>
</tbody>
</table>

These results demonstrate that the context surrounding assimilated segments influences their interpretation. Surprisingly, they suggest that listeners do not rely on phonological knowledge to solve the feature parsing problem. If listeners used the knowledge that coronality and labiality are only combined in English segments when a coronal segment is followed by a labial they could infer the correct feature parsing solution without access to context. Furthermore, the results cannot be attributed to listeners’ ability to discriminate between contextually appropriate and inappropriate modification because 1b and 1c describe inappropriate contexts that lead to different patterns of lexical access, suggesting that listeners are not merely inverting a phonological rule.

Gow (in press, submitted) proposes a mechanism grounded in perceptual grouping that provides a straightforward explanation for the results summarized in 1a-c. Roughly, perceptually robust feature cues from context segments disambiguate complex place by drawing away weaker cues for assimilated features from the target of assimilation. In this set of experiments, when the coronal /t/ in *right* assimilates labial place it encodes a combination of coronal and labial place. In 1a subsequent labial context pulls away the labiality of the assimilated segment leaving only evidence of coronality to map onto the
final segment. This leads subjects to access RIGHT, but not RIPE. Similarly, in 1b the coronal context pulls away the coronality of the segment, leaving only evidence of labiality to associate with that segment. This causes subjects to access RIPE, but not RIGHT. Finally, in 1c the post-assimilation context is removed entirely. This means that there is nothing to pull away the coronal or labiality of the assimilated segment. As a result, listeners map both features onto the same segment. Mapping labiality onto the segment leads to the activation of RIPE, and mapping coronality onto the segment leads to the activation of RIGHT.

(2) Feature parsing schema

\[
\begin{array}{ccc}
\text{speech signal} & \text{feature attraction} & \text{phonemic percept} \\
r_{p}^{1} \text{berries} & r_{p}^{1} \text{berries} & \text{rait berries} \\
[\text{coronal}] & [\text{coronal}] & [\text{coronal}] \\
[\text{labial}] & [\text{labial}] & [\text{labial}] \\
\end{array}
\]

Gow’s account of these results (in press, submitted) is grounded in basic perceptual grouping mechanisms. Research in both visual and auditory perception indicates that perceivers group perceptual elements based on a variety of principles including the ability to participate in higher order organizations, proximity, and similarity along a number of dimensions (see Bregman, 1990 for a review). The availability constraint reflects the first of these factors. In auditory processing, proximity refers to the tendency to group elements that appear in temporal proximity to one another. In the case of assimilation or coarticulation, normal proximity constraints may lead to incorrect grouping as articulatory gestures and the cues they produce are pushed together in time. Finally, within the domain of speech perception similarity may act at several levels of abstraction. Vowels appear to be grouped together on the basis of common or continuous fundamental frequency. For example, Bregman notes that when the word *sissy* is repeated continuously listeners report hearing two independent streams of sound – a series of frication sounds and a series of alternating vowel sounds. Results such as these suggest that under some circumstances vowels and consonants are grouped independently.

While fundamental frequency may bind vowels together into a common stream, speech perception typically does not appear to rely on grouping based on physical similarity. Remez et al. (1994) note that such a strategy would replace phonemic perception with the perception of simultaneous streams of sequences of noise and tones with separate streams for different formants, or instances of noise with different frequency characteristics. We would like to propose that Gestalt grouping principles would allow for segmental perception if they were applied to the level of features or feature cues rather than the raw signal. Julesz (1965) argues that Gestalt grouping by similarity takes place at the level of spontaneous pre-attentive features. Low-level feature detectors (e.g. VOT or burst spectrum detectors) meet this requirement. Marr (1976) suggests that Gestalt principles in visual perception are rooted in the sensible assumptions that can be made about the structure of the physical world. Similarly, one might argue that Gestalt principles in speech perception may be rooted in sensible...
assumptions that may be made about the relationship between articulation and the structure of the speech signal. This would lead listeners to group nearby cues based on the fact that they mark the same feature value, even though they may be physically dissimilar. For example, listeners might group evidence for labial place based on F2 values prior to a consonantal closure with subsequent evidence for labial place based on the spectrum of the burst, because it may be reasonably assumed that they are derived from the same articulation.

Ideally these principles converge to support a single feature parsing solution. Adjacent segments would have maximally contrasting features, feature cues associated with each segment would occur at the same time, temporally separated from cues associated with adjacent segments, and groupings would be consistent with higher order organizations. However, that is seldom the case in speech. In speech as in the visual domain different organizational mechanisms may promote mutually exclusive groupings of elements. Just as Gow found that different contexts lead to different interpretations of assimilated items in 1a-c, manipulations of context cause different organizations to dominate in grouping problems. Consider the grouping of the two X’s that appear in examples 3a-3d. Example 3a can be viewed either as two independent elements, or as one group that includes two elements. While viewers are able to see the example either way, they cannot entertain both organizations simultaneously. While the two X’s bear the same relationship with one another in terms of similarity and proximity in all four examples, context manipulations lead them to be grouped together in 3c, and separately in 3b, and 3d. For example, the context placed after the two X’s in 3c pulls the X’s together into a group that is set apart from the three O’s that follow, just as the consonant clusters in 3b push them apart. In 3d higher order organization into the word SEX and the familiar sequence XYZ likewise pulls the X’s apart.

(3a) X X
(3b) YZX XRW
(3c) X X X O O O
(3d) SEX XYZ

Bregman (1990) notes that when different groupings compete, the strength of a particular grouping is a function of the amount of evidence it encompasses. As a grouping gains more support, it has an increased ability to draw in new elements. In the feature parsing domain, phonetically strong positions may constitute powerful attractors for cues marking similar features. For example, the place of word-initial stops is reliably marked by several acoustic cues that occur in relatively close temporal proximity to one another. This explains why onsets attract feature cues from assimilated segments as in examples (1a-b) even when the onset plays no role in assimilation (1b). When two robust cues to labiality are grouped, they become a strong attractor for other neighboring cues for labiality, including those that do not necessarily belong to the same segment.
To summarize, the temporal distribution of feature cues in the speech stream poses a perceptual grouping problem for listeners. Feature parsing may be constrained in some instances by the availability of segments or lexical representations that can be associated with recovered features. However, in many instances, feature parsing is driven by perceptual grouping principles acting at the level of features or feature cues. Evidence from English coronal place assimilation suggests that particularly salient feature cues may attract weaker feature cues from neighboring segments.

4. Evidence from phonology

The evidence from processing suggests that local context plays a critical role in feature parsing and thus in the perception of segments. If this is the case, perceptual grouping considerations should be reflected in phonology in at least two ways. First, grouping processes can create systematic distortion in speech perception by realigning the mapping between cues and segments. This may develop into sound change. Second, given the potential for distortion, language systems may be expected to evolve characteristics that simplify the feature parsing problem and minimize the chances of spurious association. This section considers these predictions.

Ohala 1981 describes an example of sound change in Shona that is consistent with systematic feature parsing distortion caused by perceptual grouping processes (4). After labial consonants, the labiovelar glide /w/ became velar /f/.

(4) Shona dissimilation
(Ohala 1981: 188)

<table>
<thead>
<tr>
<th>Proto-Bantu</th>
<th>Pre-Shona</th>
<th>Shona</th>
</tr>
</thead>
<tbody>
<tr>
<td>*-bua</td>
<td>*-bwa</td>
<td>-b/a</td>
</tr>
<tr>
<td>*ku-mu-a</td>
<td>ku-mw-a</td>
<td>ku-mf-a</td>
</tr>
</tbody>
</table>

‘dog’
‘to drink’

Ohala 1981 argues that dissimilation arises when a listener wrongly interprets the labial glide as resulting from the interaction of a velar glide with the preceding labial consonant. This scenario is illustrated here in (5).

(5) Dissimilation scenario (Ohala 1981:187)

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Listener</th>
<th>Listener-turned-Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>/bwa/</td>
<td>/b/a/</td>
<td></td>
</tr>
<tr>
<td>produced as</td>
<td>↑</td>
<td>produced as</td>
</tr>
<tr>
<td>↓</td>
<td>↓</td>
<td></td>
</tr>
<tr>
<td>[bwa]          —heard as→</td>
<td>[bwa]</td>
<td>[b/a]</td>
</tr>
</tbody>
</table>
Feature parsing theory explains the reconstruction process. It predicts that when equivalent cues are in proximity, the more salient cue may attract its weaker neighbors. In the Shona case, the shared labiality feature is drawn away by the more salient word-initial consonant. This resulted in a sound change that eliminated labiality from the glide.

Dissimilation is likewise a well-attested synchronic phenomenon. As the examples from Kikuria in (6) illustrate, when the prefix *oko* precedes a stem whose first consonant is a voiceless supralaryngeal such as {t,s,k} its /k/ is pronounced as [g] (6e-h). Example (6b) shows that this effect does not occur when there is any intervening consonant.

(6) Kikuria (Dahl's law) (Odden 1994: 304)

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>oko-ráára</td>
<td>'to sleep'</td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>oko-mE@nE@nEkánya</td>
<td>'to make each other shine'</td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>oko-báára</td>
<td>'to count'</td>
<td></td>
</tr>
<tr>
<td>d.</td>
<td>oko-hágaacha</td>
<td>'to build'</td>
<td></td>
</tr>
<tr>
<td>e.</td>
<td>ogo-téma</td>
<td>'to hit'</td>
<td></td>
</tr>
<tr>
<td>f.</td>
<td>ogo-sóóka</td>
<td>'to respect'</td>
<td></td>
</tr>
<tr>
<td>g.</td>
<td>ogo-kóNóónta</td>
<td>'to slap'</td>
<td></td>
</tr>
<tr>
<td>h.</td>
<td>ogo-ótyá</td>
<td>'to split (tr.)'</td>
<td></td>
</tr>
</tbody>
</table>

Like the Shona example above, this kind of dissimilation can be straightforwardly understood as the result of spurious feature parsing diachronically (7). The cues for voicelessness are attracted to the first stem consonant. As a result, the prefixal /k/ surfaces as the voiced [g]. The perceptual mechanisms underlying feature attraction may allow grouping across the intervening vowel. If consonants and vowels stream separately, or if the vowels have a short duration, then the cues for voicelessness in the prefix *oko* are perceptually contiguous to those of the first consonant in the root and may group by similarity.

(7) Feature parsing as source of dissimilation

<table>
<thead>
<tr>
<th>underlying form</th>
<th>feature attraction</th>
<th>output</th>
</tr>
</thead>
<tbody>
<tr>
<td>/o k o - t é m a/</td>
<td>o k o - t é m a</td>
<td>o g o - t é m a</td>
</tr>
<tr>
<td>[-voice] [-voice]</td>
<td>[-voice]</td>
<td>[+voice] [-voice]</td>
</tr>
</tbody>
</table>

Feature parsing explains the way in which Dahl's Law could have developed Voicelessness is drawn away from the prefix by the voiceless stem-initial obstruent, and the otherwise more marked [+voice] feature becomes a context-specific default. An Optimality Theoretic analysis of the synchronous alternation is straightforward. The pattern is derived from the interaction of the
OCP (Leben 1973) with the universal voice markedness hierarchy (8). The OCP penalizes voiceless obstruents between which no consonant intervenes. The markedness constraints (8b-c) are violated by voiced and voiceless obstruents respectively. The hierarchy in (8d) reflects the relative markedness of the two types of obstruents. All else being equal, voiceless obstruents are considered less marked and are generally the default.

(8) Constraints:

a. OCP [voice] Voiceless obstruents are not C-
adjacent
b. *[+voice, -son] No voiced obstruents
c. *[-voice, -son] No voiceless obstruents
d. *[+voice, -son] » *[+voice, -son] fixed markedness hierarchy

The tableau in (9) illustrates the basic analysis. In Kikuria, it must be the case that the OCP outranks both [voice] markedness constraints. The form most faithful to the input, oko-têma, has less marked obstruents, since both are voiceless, but it fatally violates the OCP (9a). The optimal output candidate (9b) avoids the OCP violation by inserting [+voice]. Under normal circumstances this is the more marked obstruent, but in the environment of another voiceless obstruent it becomes optimal.

(9) Kikuria dissimilation

<table>
<thead>
<tr>
<th>/oko têma/</th>
<th>OCP [voice]</th>
<th>*+[voice]</th>
<th>*-[voice]</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="image" /></td>
<td><img src="image" alt="image" /></td>
<td><img src="image" alt="image" /></td>
<td><img src="image" alt="image" /></td>
</tr>
<tr>
<td><img src="image" alt="image" /></td>
<td><img src="image" alt="image" /></td>
<td><img src="image" alt="image" /></td>
<td><img src="image" alt="image" /></td>
</tr>
<tr>
<td><img src="image" alt="image" /></td>
<td><img src="image" alt="image" /></td>
<td><img src="image" alt="image" /></td>
<td><img src="image" alt="image" /></td>
</tr>
</tbody>
</table>

Although the synchronic analysis does not invoke feature parsing explicitly, the OCP is a natural consequence of the feature parsing framework. By maximizing the contrast between local segments, the OCP simplifies the feature parsing problem, in this case by voicing one of the obstruents.

Another process that can be understood as the attraction of feature cues from a weak to a strong host is degemination. In the transition from Latin to Spanish, for example, geminates were reduced to their singleton counterparts.

---

(10). In this case it appears that the onset attracts away all of the features from the coda slot, resulting in the perception of a single segment.

(10) Degemination (Hock 1991: 82)

<table>
<thead>
<tr>
<th>Latin cippus</th>
<th>Spanish cepo</th>
<th>‘pole; branch’</th>
</tr>
</thead>
<tbody>
<tr>
<td>rattus</td>
<td>rato</td>
<td>‘rat’</td>
</tr>
<tr>
<td>vacca</td>
<td>vaca</td>
<td>‘cow’</td>
</tr>
</tbody>
</table>

If the feature parsing idea is correct, then we also expect that phonological systems will be organized in part to avoid the possibility of such spurious feature cue groupings. The well-known process of anti-gemination can be understood in this way (McCarthy 1986). Afar provides a simple illustration (11-12). In Afar, vowels delete in double-sided open syllables (12a-b). However, when the consonants flanking the vowel are identical no deletion occurs (12c). The failure of vowel deletion serves to separate the two identical consonants and minimizes the possibility of degemination via spurious feature association.

(11) Anti-gemination in Afar (from McCarthy 1986)

V deletion rule: \( V \rightarrow O/VC_i \_C_jV \quad (C_i \neq C_j) \)

(12) Afar (Bliese 1981)

a. wager-n-é 'we reconciled' wagr-é 'he reconciled'
b. alif-tee-ni 'you (pl) closed' al.f-é 'he closed'
c. adud-t-é 'she was wet' adud-é 'he was wet'

5. Direction of attraction

One remaining question is what determines the direction of attraction. The framework of perceptual grouping offers two potential mechanisms. First, grouping may be constrained by availability or other higher order organizational constraints. Just as vowel nasality is not associated with vowels by English speakers because nasality is not contrastive (Lahiri and Marslen-Wilson 1991), feature association may be limited by a range of markedness constraints. Following the Gestalt Law of Pragnanz which states that simple, stable organizations are favored in perceptual grouping, feature association in the assimilation case will favor a simple, stable organization over a more marked one. The simplest and most stable pattern of association between feature cues and segments is one in which each segment has a single feature value and so groupings that favor this arrangement will be favored over ones that associate multiple contrasting place features with a single segment. By extension, other markedness constraints may reflect perceptually stable groupings.

Dissimilation illustrates the role of a second mechanism sensitive to the relative strength of the participating segments. Following the psychophysical
argument that more encompassing groupings dominate weaker ones, Gow (submitted) argues that phonetically strong positions dominate weak ones as potential feature cue attractors. This is supported by substantial phonological evidence that strong elements sustain a broader range of features than weak ones do in many languages (Beckman 1998 and references therein). In many cases of dissimilation where there is a phonetic asymmetry between two neighboring segments, it is the weaker segment that loses its features.

A good illustration of this is the well-known Seri glottal dissimilation in (13) described by Marlett and Stemberger 1983 and Yip 1988. Here when glottal stops occur in both onset and coda position of the same syllable, the coda glottal disappears (13a). This does not reflect a prohibition against glottal codas per se, since when the onset is not identical to the coda, a coda glottal stop is allowed to remain (13b). Example (13c) likewise shows dissimilation between the second and third glottal stops, but notice that the two onset glottals remain, despite the fact that they are no farther apart than the dissimilating onset/coda pair.3

(13) Seri glottal dissimilation in same syllable (Yip 1988: 75)

<table>
<thead>
<tr>
<th>underlying form</th>
<th>surface form</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. /a-a:/-sanx</td>
<td>/a:-sanx</td>
</tr>
<tr>
<td>b. ko/pans&amp;X</td>
<td>‘run like him!’</td>
</tr>
<tr>
<td>c. /i-/-a: -kas&amp;ni</td>
<td>/i-/-a: -kas&amp;ni</td>
</tr>
</tbody>
</table>

The diagram in (14) shows how Seri glottal dissimilation could have developed. When a strong (onset) and a weak (coda) glottal are close enough, the strong element attracts the relevant feature (14a). When there is no asymmetry, as in (14c), which has only a single preconsonantal glottal, or in (14b), where both surviving glottal stops are onsets, no spurious feature parsing occurs.

(14) Feature parsing as source of dissimilation

<table>
<thead>
<tr>
<th>a. feature attraction</th>
<th>b. some feature attraction</th>
<th>c. no feature attraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>/a-a: -sanx</td>
<td>/i-/- a: -kas&amp;ni</td>
<td>ko/pans&amp;X</td>
</tr>
<tr>
<td>[c.g.] ←[c.g.]</td>
<td>[c.g.] [c.g.] ←[c.g.]</td>
<td>[c.g.]</td>
</tr>
</tbody>
</table>

The synchronic Seri pattern is a logical solution to the threat of feature absorption. The core of an Optimality Theoretic analysis is provided in (15-16). The OCP penalizes glottal stops that are not separated by a consonant, and thus

---

3 In other cases of dissimilation, such as Kikuria above, feature attraction does occur between onsets, but here there is another asymmetry: root vs. affix, Fukazawa, H. (1999). Theoretical Implications of OCP Effects on Features in Optimality Theory, University of Maryland. lists a number of such cases. This is either a phonetic effect (since in many cases roots are pronounced more robustly than affixes), or it reflects a processing distinction between closed and open class morphemes.
can be satisfied by the deletion of one of them. The ONSET constraint is violated by syllables that do not begin with a consonant (Prince and Smolensky 1993).

(15) Constraints:

a. OCP (constricted glottis) No C-adjacent [constricted glottis]
b. ONSET A syllable has an onset
c. Ranking: ONSET » OCP

The tableau in (16) provides a ranking argument for ONSET and the OCP. The form most faithful to the input (16a) fatally violates the OCP, since better candidates exist that violate the OCP fewer times. Reduction of OCP violations by deletion of an onset glottal (16b-c) fatally violates the ONSET constraint. The optimal candidate minimizes OCP violations by deleting only the coda glottal stop.

(16)

<table>
<thead>
<tr>
<th></th>
<th>/i-</th>
<th>/- a/-kas&amp;ni</th>
<th>ONSET</th>
<th>OCP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a.</td>
<td>/i-</td>
<td>/- a/-kas&amp;ni</td>
<td></td>
<td>**!</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>/i.</td>
<td>a/-kas&amp;ni</td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>i.</td>
<td>/- a/-kas&amp;ni</td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. ex</td>
<td>/i.</td>
<td>/- a/-kas&amp;ni</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

Like in the account developed for Kikuria above, the synchronic analysis does not itself invoke feature parsing, but the OCP-driven effect makes sense given the feature parsing framework. In this case, the resolution of the OCP reflects an extreme case of spurious feature parsing, since the output contains fewer glottal stops than the lexical form. Seri demonstrates further the importance of relative positional strength in feature grouping, since here feature attraction occurs only between a more encompassing grouping and a weaker one, not between two positions of equal strength.

In sum the perceptual account provides two mechanisms that account for the direction of attraction: relative markedness and strength. A third logical possibility is that each language or process arbitrarily chooses a direction of attraction. While formally possible, this has two problems. First, the stipulative nature of directionality predicts that languages will randomly choose one direction or another. This appears to be the incorrect prediction. In a recent survey of dissimilation by Fukazawa 1999, directionality of dissimilation does
not appear to be random or exclusive. Rather it is a function of syllabic or morphological constituency, considerations that may be tied to the strength argument. Furthermore, parameterizing directionality does not make sense in perceptual terms. Within a perceptual account one might explain a tendency for right-to-left grouping on the grounds that the first element initially constitutes the only, and thus, the strongest (at least temporarily), grouping of feature cues. However, exclusive left-to-right attraction cannot be explained in perceptual terms.

6. Conclusion:

This paper has introduced the problem of feature parsing and explored the role of perceptual grouping processes in language processing and in a wide variety of phonological phenomena including assimilation, dissimilation and anti-gemination. We have shown that feature parsing is a fundamental component of speech perception, and that the threat of spurious feature association constrains the evolution of phonological systems.

Many questions remain. With respect to the feature parsing mechanism itself, it is still necessary to develop a more precise definition of strength. In addition it is important to ascertain which factors define the range of interaction between feature cues in different contexts and languages. To the extent that grouping is determined by context, we hypothesize that languages will differ in their grouping possibilities depending on the available contexts, phonemic inventory and phonotactic constraints. This hypothesis can only be addressed with psychophysical data. More generally, although we presented a number of cases here as strictly the result of feature parsing, a fuller account must address the contributions of articulatory and formal constraints on these and similar phenomena.

References


Department of Linguistics and Philosophy  
E39-245 MIT  
77 Massachusetts Avenue  
Cambridge MA 02139  
USA  

Neuropsychology Laboratory  
VBK 827  
Massachusetts General Hospital  
55 Fruit Street  
Boston, MA 02114  

czoll@mit.edu  
gow@helix.mgh.harvard.edu  

* We are grateful to Zhiqiang Li and the editors of this volume for their assistance in preparing this manuscript.