Lexical Exceptions in Optimality Theory
Joe Pater
April 2003

1. Introduction

Several mechanisms have been proposed to handle lexical exceptions in Optimality Theory. Within generative phonology, the following two criteria are standardly held to characterize an adequate theory of exceptions:

(1) a. It should be able to express the distinction between regular and exceptional forms
   b. It should be able to express the distinction between exceptional and ungrammatical forms

In this paper, I argue that only one of the proposed mechanisms is adequate on both of these counts: lexically specific constraints. Furthermore, I argue that these criteria necessitate the lexical indexation of both markedness and faithfulness constraints.

The first of these claims contradicts the position of Anttila (2002), who argues that co-phonologies can better meet criterion (1b) than lexically specific constraints. Here I make the case that co-phonologies and lexically specific constraints are notational variants in terms of the kinds of exceptionality they can capture, so long as both markedness and faithfulness constraints can be lexically indexed (Pater 2000; cf. Fukazawa, Itô and Mester). However, it appears that no extant version of the co-phonology model simultaneously satisfies both (1a) and (1b), whereas lexically specific constraints do so straightforwardly.

To facilitate the comparison of the different models, I will use a small set of constraints to construct hypothetical instances of lexical exceptionality. I will relate each of these cases to general predicted types of exceptionality, and discuss how the typology is instantiated cross-linguistically. The constraints I adopt are the following:

(2)  \text{NOCODA} \quad \text{Syllables end in vowels} (*C|a)
    \text{MINWD} \quad \text{Words are minimally bimoraic} (CVCV or CVC)
    \text{MAX} \quad \text{Input segments have output correspondents}
    \text{DEP} \quad \text{Output segments have input correspondents}

The markedness constraints NOCODA and MINWD are standard prosodic constraints (though see McCarthy and Prince XX on the decomposition of MINWD into more basic constraints). The faithfulness constraints MAX and DEP ban deletion and epenthesis respectively (McCarthy and Prince 1999).

2. Exceptional blocking by faithfulness

This first type of exceptionality is the most often discussed, and perhaps the most common cross-linguistically. To illustrate it, let us consider a language in which underlying codas are generally eliminated through deletion. A hierarchy generating this pattern is as in (3):
(3) \textit{NoCoda, Dep} \gg \textit{Max, MinWd}

Schematic input-output mappings generated by this ranking are provided in (4). Adjacent to each Input-Output are failed output candidates (losers), and the rankings that are required to rule each of them out in favor of the optimal form:

\begin{tabular}{|c|c|c|c|c|c|}
\hline
Input & Output & Loser 1 & Rankings & Loser 2 & Rankings \\
\hline
CVC & CV & CVC & NoCoda >> Max & CV & CVCV \text{ Dep} >> Max \\
& & & NoCoda >> MinWd & & \text{ Dep} >> MinWd \\
CVCV & CVCV & CVCVC & NoCoda >> Max & CVCVCV & \text{ Dep} >> Max \\
\hline
\end{tabular}

Because CVC maps to the sub-minimal CV, not only must NoCoda and Dep outrank Max, but they must also dominate MinWd. Exceptional forms in this language retain their codas, in contravention of the NoCoda \gg Max ranking. One account of this is to posit a lexically specific version of the Max constraint, Max-L, which targets the exceptional lexical items:

(5) Grammar: \textit{Max-L, Dep} \gg \textit{NoCoda} \gg \textit{Max, MinWd}  
Lexicon: CVC, CVC, CVC, CVCVC, CVCVC, CVCVC

Under this ranking, the exceptional lexical items that are indexed to Max-L (notated with a subscript L) will surface with codas. Dep ranks above NoCoda so that the exceptional items do not trigger epenthesis. Other words in the language continue to undergo deletion.

In a co-phonology account, the exceptional ranking is specified in lexical entries. Lexical items are subject to the usual hierarchy of the language, except when a lexical ranking overrides it. One issue that arises for the co-phonology approach is how to specify the range of possible lexical re-rankings; Anttila’s (2002) solution is discussed in section XX.

(6) Grammar: \textit{Dep} \gg \textit{NoCoda} \gg \textit{Max, MinWord}  
Lexicon: CVC, CVC, CVC_{\text{Max}>>\text{NoCoda}}, CVCVC, CVCVC, CVCVC_{\text{Max}>>\text{NoCoda}}

Both of these accounts conform to the standard generative practice of marking exceptional items in the lexicon. The regular pattern is the one that occurs in the absence of lexical marking, and regularization is formalized as the loss of lexical marking.

One objection that can be raised against both accounts is that diacritics are used in the lexicon, rather than phonological features. To avoid this problem, the grammar could be modified so as to permit codas, and the lexicon could be restructured accordingly:

(7) Grammar: \textit{Dep, Max} \gg \textit{NoCoda, MinWord}  
Lexicon: CV(C), CV(C), CVC, CVCV(C), CVCV(C), CVCVC

The parenthesized consonants in the lexical entries indicate consonants that surface only when in non-coda position, for instance, when a vowel-initial suffix is added. This could be accomplished by having two allomorphs in the lexical entry, one in which the parenthesized consonant is present, and one in which it is absent (Kager et al.). Alternatively, the consonant might be distinguished from a fixed one by leaving some structure unspecified (i.e. a ghost segment...
Inkelas, Orgun and Zoll and others). This creates a sort of ternary specification between consonants that are present, absent, and unspecified for presence. In both of these accounts, whether the consonant surfaces is determined by markedness constraints: NoCoda in the absence of the suffix, and Onset in its presence (see Kager, Inkelas et al. for details).

There is a cost to eliminating lexical diacritics: the distinction between regular and exceptional forms is not expressed. If our hypothetical language had only one lexical item that surfaced with a coda, then Max would have to be ranked above NoCoda for the language as a whole, and the grammar would not differ from a language that had a fully productive contrast in this position. This consideration argues in favor of grammatical accounts of exceptionality, such as lexically specific constraints or co-phonologies, over these representational approaches (see also Itô and Mester, cf. Inkelas, Kager).

There are a number of attested cases of this type: lexical stress placement, blocking of post-nasal voicing in non-Yamato Japanese, etc. Other cases...

3. Exceptional blocking by markedness

A hypothetical example of this type of exceptionality can be constructed on the basis of the same core grammar as in the previous section. In general, codas are deleted according to the hierarchy in (3), repeated as (8):

(8)  NoCoda, Dep >> Max, MinWD

The difference between the two languages is that in the present one the exceptional forms are limited to monosyllables. An account of this limitation is that codas are exceptionally retained to satisfy word minimality. That is, a lexically specific MinWD constraint outranks NoCoda, as in (9). Dep is ranked over MinWD-L to block epenthesis for CV inputs, which I assume does not occur in this language.

(9)  Grammar: Dep >> MinWD-L >> NoCoda >> Max, MinWD
Lexicon: CVC, CVC, CVC, CVCVC, CVCVC, CVCVC

In the lexical entries, I have marked both a CVC word, and a CVCVC word, as being subject to the lexically specific constraint. Under richness of the base (Prince and Smolensky 1993), lexical diacritics should be available to words of any phonological shape. Note that this does not entail that any particular lexical item could be so designated, any more than richness of the base entails that a particular lexical item can freely vary between other contrastive specifications.

The tableaux in (10) illustrate the result of this ranking for lexical items of CVC or CVCVC shape, indexed or not for MinWD-L: codas only surface in CVC forms indexed for MinWD-L.

---

1 Inkelas et al. (XX) argue that ternary specification works better than a co-phonology approach for exceptions to coda devoicing in Turkish. They show that individual segments in a single lexical item can differ as to whether they undergo devoicing. With lexically specific constraints, the most direct account of facts of this type would be to index individual segments to the specific constraint (cf. Chomsky and Halle 1968). With this modification, ternary specification and lexically specific constraints would be notational variants in terms of their ability to handle these cases.
This lexically specific constraint analysis can be straightforwardly translated into the co-phonology model by allowing lexical items to be specified for a ranking of MinWd-L above NoCODA. This type of pattern is out of the reach of a number of models of exceptionality in Optimality Theory, and also seems difficult to express in frameworks without ranked constraints. Lexically indexed faithfulness constraints cannot capture the generalization that exceptional violations of one markedness constraint are limited to cases when another markedness constraint is at stake. Similarly, the representational accounts of exceptionality discussed in section 1 (allomorphy and ternary specification) cannot explain why a representational device is limited to just the right phonological environment. In the present case, why should non-alternating final consonants be limited to monosyllables? A purely rule-based account is in general unable to account for blocking effects of constraints. In a theory with rules and constraints, the challenge would be to formally specify that a constraint blocks the application of only one rule, for only some lexical items.

Cases of this type are not often discussed, but it could be that they are under-reported, rather than rare: not only are they difficult to express in most theories, but the generalizations are fairly subtle. One instance of lexical stress that seems to fall into this category is discussed in Pater (2000). Lone pretonic syllables that have sufficient lexical material to form a heavy syllable (a long vowel or post-vocalic consonant) regularly surface with a full vowel (11a). However, there is a set of lexical items, mostly formed with Latinate suffixes, that show vowel reduction in this environment (11b). Both patterns give rise to alternations (11c) and (11d).

(11)
The analysis in terms of lexically specific markedness constraints in Pater (2000) is as follows. The regular pattern is produced by Parse-σ outranking a constraint against adjacent stresses. The exceptional forms are targeted by lexically specific *CLASH constraint.

It is possible to reanalyze this case without lexically specific markedness if one departs from the standard assumption that the forms in (11a) are the regular ones. Full vowels or stress could then be specified in the initial syllables of those words, and left unspecified in the words in (11b). One problem for this approach is that this pattern is evidenced in alternations (11c), though this might be given an interpretation in terms of allomorphy or ternary specification.

Anttila (2002) discusses a case that is a closer parallel to our hypothetical one, and is more resistant to reanalysis. Other cases...

4. Exceptional triggering

Exceptional triggering is the mirror image of exceptional blocking by faithfulness: the forms that undergo a process are the exceptions. A hypothetical case would be one in which codas are usually permitted, but are deleted in a specific set of forms. This could be captured by having a general ranking of MAX >> NoCODA, with a lexically specific NoCODA-L ranked above MAX:

(12) **Grammar:** DEP, NoCODA-L >> MAX >> NoCODA, MINWD  
    **Lexicon:** CVC, CVC, CVC_L, CVCVC, CVCVC, CVCVC_L

A lexical ranking of NoCODA >> MAX would produce the same effect:

(13) **Grammar:** DEP, MAX >> NoCODA, MINWD  
    **Lexicon:** CVC, CVC, CVC-NoCODA-L >> MAX*, CVCVC, CVCVC, CVCVC-NoCODA, MINWD

This sort of pattern cannot be expressed by lexical specific faithfulness constraints, except by treating the regular pattern as exceptional. Itô and Mester XX raise the same problem for a theory without lexically specific faithfulness: that a case of exceptional blocking by faithfulness can only be analyzed with a lexically specific markedness constraint by treating the blocking pattern as regular. Because the allomorphic and ternary specification approaches make no distinction between regular and exceptional patterns, they also cannot distinguish between exceptional triggering and blocking by faithfulness.

Exceptional triggering is traditionally described as a minor rule. Kenstowicz and Kisseberth (1979: 396) discuss the case of the Russian comparative suffix, which surfaces as -eje after most roots, but as -je after “a relatively small number of roots”. Such roots would be targeted by a lexically specific markedness constraint that requires deletion of the vowel. The English stress case discussed in section 4 might also be analyzed in these terms: vowel-faithfulness generally outranks the constraint causing reduction, but a lexically specific version of the reduction constraint outranks vowel-faithfulness.

5. Exceptional repair

The final type of exceptionality is one that requires lexically specific faithfulness. Two processes occur to satisfy a single output restriction, with one of them being the regular process and the other exceptional. That is, it is a lexically conditioned conspiracy. For instance, a language might
regularly use deletion to satisfy NoCoda, but in some forms use epenthesis instead. This can be captured with a general ranking of Dep >> Max, with a lexically specific Max-L dominating Dep:

(14) Grammar: Max-L, NoCoda >> Dep >> Max >> MinWd
Lexicon: CVC, CVC, CVC_L, CVCVC, CVCVC, CVCVC_L

Alternatively, one could lexically specify the exceptional ranking:

(15) Grammar: NoCoda >> Dep >> Max >> MinWd
Lexicon: CVC, CVC, CVCMax>>Dep, CVCVC, CVCVC, CVCVCMax>>Dep

Lexically specific markedness cannot capture this pattern. The representational models might be able to distinguish morphemes so that they surface with one repair or the other (for example, by listing the correct allomorphs), but this would again fail to differentiate the regular from the exceptional pattern.

Examples: Japanese clusters, OshiKwanyama nasal substitution and post-nasal voicing.

These four types of exceptionality exhaust the logical possibilities provided by lexically specific constraints or co-phonologies. A faithfulness constraint can be exceptionally ranked above a markedness constraint (exceptional blocking) or another faithfulness constraint (exceptional repair), and a markedness constraint can be exceptionally ranked above a faithfulness constraint (exceptional triggering) or another markedness constraint (exceptional blocking).

6. Limiting exceptionality

To distinguish exceptional from ungrammatical forms, limits must be placed on the ability of lexical items to select for particular rankings. With lexically specific constraints, these limits are expressed by the position of the constraints in the hierarchy. For example, a language might generally lack onset clusters and codas, and have exceptional codas but no exceptional clusters. As shown in section 2, exceptional codas require a ranking of Max-L >> NoCoda >> Max. To ban onset clusters entirely, *Complex-Onset ranks above Max-L:

(16) Grammar: *Complex-Onset >> Max-L >> NoCoda >> Max
Lexicon: CVC, CVC, CVC_L, CCVC, CCVC, CCVC_L

As shown in (17), regular items surface without codas or clusters, while codas are permitted in lexically marked forms. Onset clusters, however, are completely prohibited.
Limited exceptionality: *COMPLEX-ONSET >> MAX-L >> NoCODA

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
<th>*COMPLEX-ONSET</th>
<th>MAX-L</th>
<th>NoCODA</th>
<th>MAX</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCVC</td>
<td>CV</td>
<td>∅ CV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CVC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CCVC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCVC̄</td>
<td>CV</td>
<td>* CV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>∅ CV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The co-phonology model that we have been considering thus far does not incorporate a means expressing such limits. There are two existing proposals for doing so. The first is to take the term “co-phonology” literally, and have separate grammars for regular and exceptional forms. Lexical items would be limited to selecting from the available grammars. This approach is illustrated in (18). Lexically marked items are subject to the co-phonology, labeled “L-Grammar”:

(18) Grammar: *COMPLEX-ONSET >> NoCODA >> MAX
    L-Grammar: *COMPLEX-ONSET >> MAX >> NoCODA
    Lexicon: CVC, CVC, CVC̄, CCVC, CCVC, CCVC̄

This seems empirically indistinguishable from having lexically specific constraints, and it has the notational disadvantage of requiring the entire constraint hierarchy to be replicated for each exceptional ranking.

Anttila (2002) proposes instead to restrict lexical items to fixing rankings of constraints that are unordered in the grammar. For the hypothetical case we have been considering, NoCODA and MAX would be left unranked, and lexical items would select one ranking or the other.

(19) Grammar: *COMPLEX-ONSET >> NoCODA, MAX
    Lexicon: CVC_{MAX>>NoCODA}, CVC_{MAX>>NoCODA}, CVC, CCVC, CCVC̄, CCVC̄_{MAX>>NoCODA}

In this proposal, lexical items do not have the capacity of overriding grammatically specified rankings. As such, the ranking between *COMPLEX-ONSET and MAX cannot be reversed, and clusters are categorically ruled out.

While this proposal allows the co-phonology model to distinguish between exceptions and ungrammatical forms, it erases the distinction between regulars and exceptions. Both regulars and exceptions must be lexically specified for the ranking of the unordered constraints, since leaving a form unspecified results in random choice between the rankings, and free variation between the outcomes.

Anttila (2002) argues that it is a virtue of this model that it unifies the treatment of variation and morphological exceptions. However, whether lexical items are specified for rankings or indexed to lexically specific constraints is independent of whether variation is captured with unranked constraints. It is perhaps more straightforward to restrict lexical items to fixing unordered constraints in the co-phonology model. However, limits on exceptional ranking are already expressed for lexically specific constraints by their position in the hierarchy, so it is unnecessary to impose such a restriction under that approach.
7. Conclusions

Lexically specific constraints and co-phonologies appear to be equivalent in terms of the types of exceptionality they can capture, so long as both markedness constraints and faithfulness constraints can both be lexically indexed. A theory with only lexically specific faithfulness is more restrictive, but it fails to express the patterns of exceptional triggering (section 4) and exceptional blocking by markedness (section 3). One might also consider allowing only markedness constraints to be lexically indexed, but this would fail to deal with exceptional blocking by faithfulness (section 2) and exceptional repair (section 5). In this respect, these two models of exceptionality are notational variants of one another. However, the co-phonology model appears to face a dilemma that does not trouble lexically specific constraints. If it is formulated in such a way as to place restrictions on the set of exceptional rankings in a language (Anttila 2002), it fails to distinguish between exceptional and regular forms. Purely representational models of exceptionality such as allomorphy and ternary specification also fail to distinguish exceptions from regular forms, and in some cases face difficulties in distinguishing exceptionality from ungrammaticality (section 3). Making the three-way distinction between regular, exceptional and ungrammatical forms seems to be a fairly minimal criterion of adequacy for a theory of exceptions, especially in light of proposals to deal with even finer gradations in grammaticality (e.g. Frisch et al., Zuraw, Boersma and Hayes). Lexically specific constraints meet this minimal criterion using a relatively minimal elaboration of the basic theory of Prince and Smolensky (1993).