

## Exceptions in Optimality Theory: Typology and Learnability

Joe Pater, UMass Amherst

*Goals of this talk:*

1. To provide typological support for an OT analysis of exceptions in terms of lexically indexed markedness and faithfulness constraints
2. To propose an account of the genesis of lexically specific constraints in learning
3. To discuss the problems raised by gradient phonotactic well-formedness, using novel data on consonant co-occurrence restrictions in Muna, an Austronesian language

### 1. Exceptionality as constraint indexation\*

#### 1.1 Introduction

- Since *SPE*, exceptions have been dealt with in two ways in generative phonology:
  - (1) a. *Diacritic approaches*: Exceptional lexical items are marked as subject to a rule or constraint that does not apply to regular items (or vice versa)
    - b. *Structural approaches*: Exceptional lexical items are specified with phonological structure that is not present on regular items
  - Within Optimality Theory, diacritic and structural approaches have continued to be taken:
    - (2) a. Diacritic approaches:
      - i. *Lexically specific ranking* - lexical items are specified for the ranking of two or more constraints:
        - A. Lexical items are specified for a ranking of constraints that differs from the 'core' ranking of the language (McCarthy and Prince 1993, Nouveau 1994; Pater 1994; Itô and Mester 1995a,b; Inkelas, Orgun and Zoll 1997; cf. Paradis and Lebel 1994)
        - B. Lexical items are specified to choose a particular ranking of unranked constraints (Antilla 2002)
      - ii. *Lexically specific constraints* - a lexically specific version of a constraint, ranked differently from the general one:
        - A. Only faithfulness constraints (Fukuzawa 1997; Kraska-Szlenk 1997, 1999; Itô and Mester 1999)
        - B. Markedness and faithfulness constraints (Pater 2000)
        - C. Only markedness constraints (?)

---

\* Acknowledgements for Part 1: Thanks to Arto Anttila, John McCarthy, Cheryl Zoll, and the participants in Ling 730, Fall 2002 UMass for useful discussion. Thanks also to Iwona Kraska-Szlenk and Wim Zonneveld for sending me copies of their work on exceptionality.

b. Structural approaches:

i. *Prespecification*: Specified structure that is protected by a faithfulness constraint (many: see esp. Inkelas, Orgun and Zoll 1997, Piñero 2004)

ii. *Allomorphy*: Similar to prespecification, except that entire morpheme alternants are stored (Kager 1999)

- Standardly, it is assumed that an adequate theory of exceptions should be able to:

- (3) a. Express the distinction between regular and exceptional forms  
 b. Express the distinction between exceptional and ungrammatical forms

- For example, in English (Hayes 1982; cf. Pater 1994):

- (4) a. If the penultimate syllable is light, the antepenult is usually stressed (e.g. *Cánada*), but in exceptional words, the penult receives stress (e.g. *banána*)  
 b. If the penultimate syllable is heavy, antepenultimate stress is impossible (\*‘póedektal’)

- To facilitate a comparison of different approaches, I will use a small set of constraints to construct hypothetical instances of lexical exceptionality.

- (5) NOCODA      Syllables end in vowels (\*C]<sub>σ</sub>)  
 MINWD        Words are minimally bimoraic (CVCV or CVC)  
 \*COMPLEX     No consonant clusters  
 MAX            Input segments have output correspondents (‘No deletion’)  
 DEP            Output segments have input correspondents (‘No epenthesis’)

- I will relate each of these cases to general predicted types of exceptionality, and provide some examples of how the typology is instantiated cross-linguistically

### 1.2 Exceptional Blocking by Faithfulness

- This is the most commonly discussed type of exception: an alternation is blocked in exceptional items
- Illustration: a language in which underlying codas are generally eliminated through deletion, and in which exceptional items retain their final consonants

- (6) /pak/ → [pa]            /pidot/ → [pido]  
 /lot/ → [lo]              /talak/ → [tala]  
 /tak/ → [tak]             /likot/ → [likot]

- Analysis of coda deletion:

- (7) NOCODA >> MAX

| Input | Output | NOCODA | MAX |
|-------|--------|--------|-----|
| pak   | pak    | * !    |     |
|       | pa     |        | *   |

- Under a lexically specific constraint analysis, the exceptional items are targeted by a ‘cloned’ MAX constraint, MAX-L, that ranks above NOCODA

(8) Grammar: MAX-L >> NOCODA >> MAX  
 Lexicon: /pak/ /lot/ /tak<sub>L</sub>/ /pidot/ /talak/ /likot<sub>L</sub>/

- Illustrative tableaux:

(9)

| Input            | Output | MAX-L | NOCODA | MAX |
|------------------|--------|-------|--------|-----|
| pak              | pak    |       | * !    |     |
|                  | pa     |       |        | *   |
| tak <sub>L</sub> | tak    |       | *      |     |
|                  | ta     | * !   |        | *   |

- Let us further assume that the language also bans consonant clusters, but unlike the restriction against codas, this one is absolute
- This is accounted for by \*COMPLEX >> MAX-L; if under Richness of the Base an underlying form with a cluster is given a lexical diacritic, the cluster is reduced:

(10)

| Input            | Output | *COMPLEX | MAX-L | NOCODA | MAX |
|------------------|--------|----------|-------|--------|-----|
| CCV              | CCV    | * !      |       |        |     |
|                  | CV     |          |       |        | *   |
| CCV <sub>L</sub> | CCV    | * !      |       |        |     |
|                  | CV     |          | *     |        | *   |

- Our desiderata for a theory of exceptions are fulfilled as follows:

(11) a. Exceptional forms, but not regular forms, are marked in the lexicon as subject to a lexically specific constraint  
 b. Exceptional forms are distinguished from ungrammatical forms by the ranking of the lexically specific constraint

- An alternative diacritic approach is to specify rankings in the lexicon
- In our example, exceptional lexical items would be specified for MAX >> NOCODA, which reverses their ranking in the tableau:

(12) Grammar: \*COMPLEX, NOCODA >> MAX  
 Lexicon: /pak/ /lot/ /tak<sub>Max>>NoCoda</sub>/ /pidot/ /talak/ /likot<sub>Max>>NoCoda</sub>/

| Input | Output | *COMPLEX | NOCODA | MAX |
|-------|--------|----------|--------|-----|
| pak   | pak    |          | * !    |     |
|       | pa     |          |        | *   |
|       | pakə   |          |        |     |

| Input                              | Output | *COMPLEX | MAX | NoCODA |
|------------------------------------|--------|----------|-----|--------|
| tak <sub>Max &gt;&gt; NoCoda</sub> | ☞ tak  |          |     | *      |
|                                    | ta     |          | * ! |        |
|                                    | takə   |          |     |        |

- The problem is that nothing rules out lexically specifying any other ranking
- An absolute restriction against clusters cannot be expressed, since lexical items can be specified for MAX >> \*COMPLEX

(13)

| Input                                | Output | *COMPLEX | NoCODA | MAX      |
|--------------------------------------|--------|----------|--------|----------|
| CCV                                  | CCV    |          |        | *        |
|                                      | ☞ CV   | *!       |        |          |
| Input                                | Output | MAX      | NoCODA | *COMPLEX |
| CCV <sub>Max &gt;&gt; *Complex</sub> | ☞ CCV  |          |        | *        |
|                                      | CV     | * !      |        |          |

- One way of restricting lexical rankings is to stipulate that only constraints that are unranked in the grammar can have lexically specified rankings (Anttila 2002):

(14) Grammar: \*COMPLEX >> NoCODA, MAX  
 Lexicon: /pak<sub>NoCoda >> Max</sub>/ /lot<sub>NoCoda >> Max</sub>/ /tak<sub>Max >> NoCoda</sub>/ /pidot<sub>NoCoda >> Max</sub>/ /talak<sub>NoCoda >> Max</sub>/  
 /likot<sub>Max >> NoCoda</sub>/

- Under this account, we lose the distinction between regulars and exceptions, since both of their rankings must be lexically specified, so as to block variation
- We must also posit a principle that stops grammatically specified rankings from being altered by lexical rankings
- Alternatively, we could allow words to choose between a specified set of rankings, or co-grammars; exceptions are subject to a special grammar, here marked the L-Grammar:

(15) Grammar: \*COMPLEX-ONSET >> NoCODA >> MAX  
 L-Grammar: \*COMPLEX-ONSET >> MAX >> NoCODA  
 Lexicon: /pak/ /lot/ /tak<sub>L</sub>/ /pidot/ /talak/ /likot<sub>L</sub>/

- This seems empirically indistinguishable from having lexically specific constraints, and has the notational disadvantage of replicating the entire hierarchy for each exceptional ranking
- Although either of these approaches could generate the same set of patterns as lexically specific constraints, lexically specific constraints have some conceptual and notational advantages (cf. Anttila 2002)

- A structural approach would posit an underlying distinction between deleting and non-deleting codas - let us say in terms of whether they bear a syllabic timing slot (x):

(16) Grammar: \*COMPLEX >> MAX-X >> NOCODA >> MAX

Lexicon: /pak/ /lot/ /tak<sup>x</sup>/ /pidot/ /talak/ /likot<sup>x</sup>/

- The tableaux for the previous analysis would be adapted by replacing the diacritic ‘L’ with the structure ‘x’, and our desiderata would be met in this minimally different fashion:

(17) a. Exceptional forms, but not regular forms, are marked in the lexicon with extra structure  
 b. Exceptional forms are distinguished from ungrammatical forms by the ranking of the constraint referring to that structure

- Note that the success of this structural analysis depends on constraint ranking: why should lexically specified structure override one constraint or rule, but not another? (Pater 2000)

- Examples of exceptional blocking by faithfulness:

(18) a. Turkish words regularly undergo coda devoicing, but it is blocked in some loanwords (Inkelas *et al.* 1997)

b. In Japanese, voiced geminates are regularly devoiced, but unassimilated foreign words retain their voicing (Itô and Mester 1999)

c. English sonorant-final medial syllables are regularly destressed (e.g. *information*), but sometimes retain the stress from their base (e.g. *èhortation*) (Pater 2000)

### 1.3 Exceptional triggering

- In this type of exceptionality, exceptional items undergo an alternation that does not apply to regular items; this is traditionally referred to as a minor rule
- Illustration - codas are generally tolerated, except in exceptional forms, in which they delete:

(19) /pak/ → [pak]            /pidot/ → [pidot]  
 /lot/ → [lot]            /talak/ → [talak]  
 /tak/ → [ta]            /likot/ → [liko]

- Here the exceptional items are indexed to a markedness constraint - in this case NOCODA:

(20) Grammar: NO-CODA-L >> MAX >> NOCODA

Lexicon: /pak/ /lot/ /tak<sub>L</sub>/ /pidot/ /talak/ /likot<sub>L</sub>/

- Illustrative tableaux:

(21)

| Input            | Output | NoCODA-L | MAX | NoCODA |
|------------------|--------|----------|-----|--------|
| pak              | ☞ pak  |          |     | *      |
|                  | pa     |          | * ! |        |
| tak <sub>L</sub> | tak    | * !      |     | *      |
|                  | ☞ ta   |          | *   |        |

- If lexical indexation is limited to faithfulness constraints, then one must analyze the regular forms as being lexically marked, and the exceptional pattern as regular:

(22) Grammar: MAX-L >> NoCODA >> MAX  
 Lexicon: /pak<sub>L</sub>/ /lot<sub>L</sub>/ /tak/ /pidot<sub>L</sub>/ /talak<sub>L</sub>/ /likot/

- A structural analysis faces a similar difficulty, since the non-alternating consonant must be lexically marked:

(23) Grammar: MAX-X >> NoCODA >> MAX  
 Lexicon: /pak<sup>X</sup>/ /lot<sup>X</sup>/ /tak/ /pidot<sup>X</sup>/ /talak<sup>X</sup>/ /likot/

- Since these approaches to exceptionality are forced to analyze exceptional blocking and triggering identically, one cannot account for the following in terms of the grammar:

(24) a. Native speaker's awareness of regulars vs. exceptions  
 b. Regularization of exceptional forms

- On the difficulty that regularization can pose for a structural account, see Albright (2002)
- In contrast, with lexically specific markedness and faithfulness constraints (or rule diacritics), one can maintain the standard view that regularization is loss of lexical marking

- Examples of exceptional triggering:

(25) a. In French, word-final velars are not normally deleted after oral vowels (e.g. *bac*, *avec*, *roc*, *mec*); however, there is a small group of words in which they are (*estomac*, *croc*, *escroc*; cf. *estomaquer*, *croquer*, *escroquer*); similarly for [l] (*calcul* vs. *cul*) (Schane 1970)

b. In Palauan, stressless high front vowels are not normally deleted, but they are in a small set of words (Flora, 1974 cited in Zonneveld 1978)

c. In Somali, codas are usually tolerated, but the autobenefactive suffix /-ad/ loses its consonant in final position (Saaed 1985, cited in Kraska-Szlenk 1999)

### 1.4 Exceptional blocking by markedness

- An alternation is sometimes blocked only when another markedness constraint is at stake
- An example would be if coda deletion were exceptionally blocked only in monosyllabic words, due to word minimality:

(26) /pak/ → [pa]                /pidot/ → [pido]  
 /lot/ → [lo]                /talak/ → [tala]  
 /tak/ → [tak]                /likot/ → [liko]

(27) Grammar: MINWD-L >> NoCODA >> MINWD, MAX  
 Lexicon: /pak/ /lot/ /tak<sub>L</sub>/ /pidot/ /talak/ /likot/

(28)

| Input            | Output | MINWD-L | NoCODA | MAX | MINWD |
|------------------|--------|---------|--------|-----|-------|
| pak              | pak    |         | * !    |     |       |
|                  | ☞ pa   |         |        | *   | *     |
| tak <sub>L</sub> | ☞ tak  |         | *      |     |       |
|                  | ta     | * !     |        | *   | *     |
| pidot            | pidot  |         | * !    |     |       |
|                  | ☞ pido |         |        | *   |       |

- The following “Richness of the Base” tableau shows that if a CVCVC form is supplied with a lexical diacritic, it still undergoes deletion, since word minimality is satisfied:

(29)

| Input              | Output | MINWD-L | NoCODA | MAX | MINWD |
|--------------------|--------|---------|--------|-----|-------|
| CVCVC <sub>L</sub> | CVCVC  |         | * !    |     |       |
|                    | ☞ CVCV |         |        | *   |       |

- With only lexically specific faithfulness constraints, MAX must be indexed:

(30) Grammar: MAX-L >> NoCODA >> MINWD, MAX  
 Lexicon: /pak/ /lot/ /tak<sub>L</sub>/ /pidot/ /talak/ /likot/

- This grammar fails the Richness of the Base test, showing it is insufficiently restrictive:

(31)

| Input              | Output | MAX-L | NoCODA | MAX | MINWD |
|--------------------|--------|-------|--------|-----|-------|
| CVCVC <sub>L</sub> | CVCVC  |       | *      |     |       |
|                    | ☞ CVCV | * !   |        | *   |       |

- The structural analysis is similar, since it must use a faithfulness constraint to preserve the lexical material:<sup>1</sup>

<sup>1</sup> John McCarthy (p.c.) points out that an analysis is possible with conjoined constraints (Smolensky 1995): MINWD&Max-x.

(32) Grammar: MAX-X >> NoCODA >> MAX, MINWD

Lexicon: /pak/ /lot/ /tak<sup>x</sup>/ /pidot/ /talak/ /likot/

- This pattern of exceptionality is problematic for a purely rule-based theory in the same way as regular blocking, and also may pose additional formal difficulties: the exception feature must be in some way limited to CVC words
- Examples of exceptional blocking by markedness:

(33) a. In Russian, deletion of [i] from the suffix /-isk/ is sometimes blocked, but only when the preceding stem ends with a velar or palatal consonant (Chomsky and Halle 1968: 379)

b. In Finnish, stem-final /a/ usually rounds to [o] before suffixal [-i], however, when the preceding vowel is round, deletion usually occurs instead (Anttila 2002)<sup>2</sup>

c. In English, initial consonant-final pretonic syllables are usually stressed (e.g. *bàndána*), however in an exceptional group of words, mostly formed with Latinate prefixes (*admíre*), these syllables surface as stressless (Pater 2000)

### 1.5 Exceptional Repair

- A lexically driven conspiracy: exceptions use a different repair than regular words to satisfy the same constraint
- For example, codas generally delete, but epenthesis applies in exceptional cases:

(34) /pak/ → [pa]            /pidot/ → [pido]  
 /lot/ → [lo]            /talak/ → [tala]  
 /tak/ → [takə]        /likot/ → [likotə]

(35) Grammar: MAX-L, NoCODA >> DEP >> MAX  
 Lexicon: /pak/ /lot/ /tak<sub>L</sub>/ /pidot/ /talak/ /likot<sub>L</sub>/

(36)

| Input            | Output | MAX-L | NoCODA | DEP | MAX |
|------------------|--------|-------|--------|-----|-----|
| pak              | pak    |       | * !    |     |     |
|                  | ↗ pa   |       |        |     | *   |
|                  | pakə   |       |        | * ! |     |
| tak <sub>L</sub> | tak    |       | * !    |     |     |
|                  | ta     | * !   |        |     | *   |
|                  | ↗ takə |       |        | *   |     |

<sup>2</sup> This is only part of the story. As Anttila shows in detail, the choice between deletion and mutation is subject to a number of conditioning factors, both phonological and morphological. He also points out that lexically specific faithfulness constraints fail to capture these generalizations.

- This pattern shows that lexically specific faithfulness is required as well as markedness, for reasons beyond distinguishing exceptional blocking from triggering (i.e. §1.2 vs. §1.3)
- Examples of exceptional repair:
  - (37) a. In OshiKwanyama, nasal-voiceless obstruent clusters undergo nasal substitution with native stems (e.g. oN+pote → omote ‘good-for-nothing’). With non-native stems, post-nasal voicing occurs instead (e.g. oN+papila → ombapila ‘paper’) (Steinbergs 1985)
  - b. In Swahili loanwords, epenthesis applies to break up clusters and avoid codas (e.g. [gilasi] ‘glass’, [sampuli] ‘sample’). A set of suffixes, however, lose their initial consonants after consonant-final stems (e.g. the applicative /-li/ /pik+li+a/ → [pikia] ‘cook (APPL.)’) (Kraska-Szlenk 1999)

### 1.6 Conclusions

- A theory with lexically specific markedness and faithfulness constraints is more powerful than either a theory with no lexically specific constraints (i.e. a purely structural approach), or one with only lexically specific faithfulness constraints
- However, it seems that this extra power is warranted, since the full set of predicted constraint interactions are attested:
  - (38) i. Exceptional blocking by faithfulness: FAITH-LEX >> MARK >> FAITH
  - ii. Exceptional triggering: MARK-LEX >> FAITH >> MARK
  - iii. Exceptional blocking by markedness: MARK-B-LEX >> MARK-A >> FAITH, MARK-B
  - iv. Exceptional repair: FAITH-B-LEX, MARK >> FAITH-A >> FAITH-B
- Further research:
  - (39) i. Testing of native speaker awareness of the distinctions
    - Do people recognize that a class of exceptions is phonologically delimited (e.g. that only monosyllables are exceptions to coda deletion)?
  - ii. The scope of indexed constraints
    - Can all constraints be indexed?
    - Is the indexation of a root inherited in derived words? (cf. Mascaró 2003)
    - Is the indexation of an affix be inherited by the entire word? (cf. Coats 1970, Kisseberth 1970)
    - Can segments be indexed? (cf. Chomsky and Halle 1968, Inkelas *et al.* 1997)

## 2. Inconsistency Resolution and Constraint Cloning\*

### 2.1 Introduction

- By definition, exceptions (to alternations) give rise to inconsistent mark-data pairs (Tesar 1995, 2000) or vectors (Prince 2002):

(40)

| Input | W ~ L    | NoCODA | MAX |
|-------|----------|--------|-----|
| pak   | pa ~ pak | W      | L   |
| lok   | lo ~ lok | W      | L   |
| tok   | tok ~ to | L      | W   |

- The constraint demotion algorithm cannot rank these constraints, since neither prefers only winners, so it stalls (and “announces inconsistency” Prince 2002: 26)
- Proposal: this launches a routine for inconsistency resolution that involves cloning and lexically indexing a constraint (cf. Tesar 2000, Tesar and Prince to appear and Tesar *et al.* 2003 for other uses of inconsistency detection)
- How does this routine choose the correct constraint to be cloned?
- The typology we saw in section 1 presents two main challenges:

- (41)
- a. Distinguishing exceptional blocking by faithfulness from exceptional triggering
  - b. Distinguishing exceptional blocking by markedness from blocking by faithfulness

### 2.2 Fewest W's

- The mark-data pairs in (40) were for an instance of exceptional blocking by faithfulness markedness; the following exemplifies exceptional triggering:

(42)

| Input | W ~ L    | NoCODA | MAX |
|-------|----------|--------|-----|
| pak   | pak ~ pa | L      | W   |
| lok   | lok ~ lo | L      | W   |
| tok   | to ~ tok | W      | L   |

- Here we need a way of choosing the constraint that will wind being indexed to the smaller, exceptional set of items
- In our idealized example, this can be done straightforwardly by counting (cf. Prince and Tesar 2004):

- (43) *Fewest W's*  
Select the constraint(s) with the fewest W marks

---

\* Acknowledgement for Part 2: Many of these ideas were developed in discussion with John McCarthy and Nicholas Winslow during the preparation of Winslow (2003). They share the credit, not the blame. Thanks also to the Rutgers Learnability Group and the UMass Phonology Group for very useful feedback.

- We also need explicit cloning and indexation mechanisms:

(44) *Cloning*  
Create a copy of the constraint

*Indexation*

Identify the lexical items for which the cloned constraint prefers winners  
Provide them with a lexical diacritic, and mark the cloned constraint with the same diacritic, so that it applies only to those items

2.3 *Markedness bias*

- In exceptional blocking by markedness, we have three constraints to choose from:

(45)

| Input | W ~ L        | NOCODA | MAX | MINWD |
|-------|--------------|--------|-----|-------|
| pak   | pa ~ pak     | W      | L   | L     |
| lok   | lo ~ lok     | W      | L   | L     |
| tok   | tok ~ to     | L      | W   | W     |
| kopak | kopa ~ kopak | W      | L   |       |
| lotak | lota ~ lotak | W      | L   |       |
| ratok | rato ~ ratok | W      | L   |       |

- Fewest W's will correctly choose MINWD over NOCODA, but MAX and MINWD are tied
- Thus, to maintain restrictiveness, we must introduce a markedness bias analogous to that of Hayes (2004) and Prince and Tesar (2004):

(46) *Markedness bias*  
i. Select markedness constraints  
ii. If there are no markedness constraints, select a faithfulness constraint

- The markedness bias must be ordered after Fewest W's, so that a faithfulness constraint can be selected in the 'exceptional blocking by faithfulness' scenario
- In addition, we should have a random selection step in case there remains more than one constraint after fewest W's and the markedness bias

(47) i. Fewest W's  
ii. Markedness bias  
iii. Random selection  
iv. Cloning  
v. Indexation

2.4 *Inconsistency location*

- At this point, we are confronted by the familiar tension between restrictiveness and generality
- Inspection of the mark-data pairs for exceptional blocking by faithfulness shows that MINWD will also be selected in this case, since it has 'fewest W's':

(48)

| Input | W ~ L        | NoCODA | MAX | MINWD |
|-------|--------------|--------|-----|-------|
| pak   | pa ~ pak     | W      | L   | L     |
| lok   | lo ~ lok     | W      | L   | L     |
| tok   | tok ~ to     | L      | W   | W     |
| kopak | kopa ~ kopak | W      | L   |       |
| lotak | lota ~ lotak | W      | L   |       |
| ratok | ratok ~ rato | L      | W   |       |

- The result will ultimately be that both MINWD and MAX are cloned
- Empirically, this seems harmless, since the following rankings produce identical results:

- (49)
- a. MAX-L >> NoCODA >> MAX
  - b. MAX-L, MINWD-L >> NoCODA >> MAX

- However, to ensure generality (or elegance or parsimony), we need to provide a means of selecting the right constraint
- That is, we must determine the locus of inconsistency, when there is more than one pair of constraints remaining when RCD shuts down
- Pursuing a suggestion of John McCarthy's, we can identify the mark-data pairs that have only a single pair of constraints that are the source of inconsistency:

(50) *Inconsistency location*

- i. Select all rows with one and only one W and one and only one L
- ii. If any of these are inconsistent (fuse to  $L^+$  in Prince 2002), select the constraints dominating the non-empty columns
- iii. If no inconsistency is detected, select all constraints

- With Inconsistency Location ordered before Fewest W's, we avoid creating redundant lexically specific constraints
- Inconsistency Location could also prove key in resolving a well-known problem for the CDA: learning variation
- Variation is inconsistency between the mark-data pairs for a single lexical item; if that is detected, then the constraints should be left unranked
- To illustrate the workings of inconsistency resolution, let us go through the three test cases
- First, faithfulness blocking (for the mark-data pairs in (48))

- (51) Inconsistency Location: NoCODA, MAX  
 Fewest W's: MAX  
 Markedness bias: MAX

- Cloning and indexation then yields the revised set of mark-data pairs, which is rankable:

(52)

| Input              | W ~ L        | NoCODA | MAX | MINWD | MAX-L |
|--------------------|--------------|--------|-----|-------|-------|
| pak                | pa ~ pak     | W      | L   | L     |       |
| lok                | lo ~ lok     | W      | L   | L     |       |
| tok <sub>L</sub>   | tok ~ to     | L      | W   | W     | W     |
| kopak              | kopa ~ kopak | W      | L   |       |       |
| lotak              | lota ~ lotak | W      | L   |       |       |
| ratok <sub>L</sub> | ratok ~ rato | L      | W   |       | W     |

- Next, exceptional triggering:

(53)

| Input | W ~ L        | NoCODA | MAX | MINWD |
|-------|--------------|--------|-----|-------|
| pak   | pak ~ pa     | L      | W   | W     |
| lok   | lok ~ lo     | L      | W   | W     |
| tok   | to ~ tok     | W      | L   | L     |
| kopak | kopak ~ kopa | L      | W   |       |
| lotak | lotak ~ lota | L      | W   |       |
| ratok | rato ~ ratok | W      | L   |       |

- (54) Inconsistency Location: NoCODA, MAX  
 Fewest W's: NoCODA  
 Markedness bias: NoCODA

- Cloning and indexation then again produces a rankable set of mark-data pairs:

(55)

| Input              | W ~ L        | NoCODA | MAX | MINWD | NoCODA-L |
|--------------------|--------------|--------|-----|-------|----------|
| pak                | pak ~ pa     | L      | W   | W     |          |
| lok                | lok ~ lo     | L      | W   | W     |          |
| tok <sub>L</sub>   | to ~ tok     | W      | L   | L     | W        |
| kopak              | kopak ~ kopa | L      | W   |       |          |
| lotak              | lotak ~ lota | L      | W   |       |          |
| ratok <sub>L</sub> | rato ~ ratok | W      | L   |       | W        |

- Finally, markedness blocking:

(56)

| Input | W ~ L        | NoCODA | MAX | MINWD |
|-------|--------------|--------|-----|-------|
| pak   | pa ~ pak     | W      | L   | L     |
| lok   | lo ~ lok     | W      | L   | L     |
| tok   | tok ~ to     | L      | W   | W     |
| kopak | kopa ~ kopak | W      | L   |       |
| lotak | lota ~ lotak | W      | L   |       |
| ratok | rato ~ ratok | W      | L   |       |

- (57) Inconsistency Location: NoCODA, MAX, MINWD  
 Fewest W's: MAX, MINWD  
 Markedness bias: MINWD

(58)

| Input            | W ~ L        | NoCODA | MAX | MINWD | MINWD-L |
|------------------|--------------|--------|-----|-------|---------|
| pak              | pa ~ pak     | W      | L   | L     |         |
| lok              | lo ~ lok     | W      | L   | L     |         |
| tok <sub>L</sub> | tok ~ to     | L      | W   | W     | W       |
| kopak            | kopa ~ kopak | W      | L   |       |         |
| lotak            | lota ~ lotak | W      | L   |       |         |
| ratok            | rato ~ ratok | W      | L   |       |         |

#### 2.4 Ranking lexically specific constraints

- Since lexically specific constraints prefer only winners, one might think that they could simply be installed in the topmost stratum
- However, this can result in overgeneration of exceptions
- Recall our example of a language that generally lacks both clusters and codas, and has exceptional codas but no exceptional clusters, which requires the following ranking:

(59) \*COMPLEX >> MAX-L >> NoCODA >> MAX

- To get \*COMPLEX-ONSET >> MAX-L, we could rely on the markedness preference of Biased Constraint Demotion (Prince and Tesar 2004)
- However, pairs of markedness constraints can give rise to the same problem
- Take a language that has a minimal word restriction, in which CVC is acceptable, but CV is not; this indicates that MINWD >> NoCODA:

(60)

| Input | Output | MINWD | NoCODA |
|-------|--------|-------|--------|
| CVC   | CV     | *!    |        |
|       | ☞ CVC  |       | *      |

- Let us further assume that codas are generally permitted (MAX >> NoCODA), but exceptional items exhibit deletion (NoCODA-L >> MAX) - that is, a case of exceptional triggering
- However, in contrast with the earlier example, CVC words absolutely resist deletion, which requires a ranking of MINWD >> NoCODA-L:

#### (61) Restricted exceptional triggering

| Input              | Output  | MINWD | NoCODA-L | MAX | NoCODA |
|--------------------|---------|-------|----------|-----|--------|
| CVC                | CV      | *     |          | *   |        |
|                    | ☞ CVC   |       |          |     | *      |
| CVC <sub>L</sub>   | ☞ CV    | *!    |          | *   |        |
|                    | CVC     |       | *        |     | *      |
| CVCVC              | CVCV    |       |          | *   |        |
|                    | ☞ CVCVC |       |          |     | *      |
| CVCVC <sub>L</sub> | ☞ CVCV  |       |          | *   |        |
|                    | CVCVC   |       | *        |     | *      |

- Both MINWD and NoCODA-L are markedness constraints, and both prefer only winners
- Therefore, a bias against installing lexically specific constraints must be incorporated into RCD in order to ensure restrictiveness

## 2.5 Error-driven learning and exceptions

- All of the above rests on a set of mark-data pairs that is complete in various respects
- With an incomplete set of data, a learner might mistakenly clone too many constraints, or more problematically, clone a constraint with a too broad scope
- There is a simple, if somewhat brutal, solution:

(62) When a new piece of data is encountered, eliminate all lexically specific constraints

- Could there be problems in not generating the informative losers? Probably not, since exceptions themselves need to be dealt with.
- The biases that real-life learners exhibit against lexically specific constraints suggest that a more subtle approach might ultimately be necessary
- As well as regularizing exceptional forms in the input data, learners also seem to regularize exceptions to their self-imposed rules:

(63) Daniel's nasal harmony (Menn 1971 cited in Menn and Matthei 1992)

|       | <i>Stage 1</i> | <i>Stage 2</i> | <i>Stage 3</i> | <i>Stage 4</i> |
|-------|----------------|----------------|----------------|----------------|
| down  | [dæwn]         | [dæwn]         | [dæwn]/[næwn]  | [næwn]         |
| stone | [don]          | [don]          | [don]/[non]    | [non]          |
| beans | –              | [minz]         | [minz]         | [minz]         |
| dance | –              | [næns]         | [næns]         | [næns]         |

- The lexically specific constraint needed for stage 2 is eliminated by stage 4
- Since this is counter to the evidence from the language, it suggests a system internal pressure against lexically specific constraints

## 3. Gradient phonotactic well-formedness: problems and prospects\*

### 3.1 Introduction

- Two problems:

(64) a. Lexically specific constraints capture three degrees of well-formedness – regular, exceptional, and ill-formed, but speakers seem to be sensitive to finer degrees of acceptability, which often reflect frequency distributions in the language (e.g. Zuraw 2000, Frisch and Zawadyeh 2001)

b. Exceptions to phonotactic generalizations do not involve inconsistency, so the learning algorithm proposed above will not work for them

---

\* Acknowledgements for Section 3: The work on Muna is being done collaboratively with René van den Berg and Andries Coetzee.

- To illustrate these problems, I will present data on consonantal co-occurrence restrictions in Muna, a Western Austronesian language spoken on an island off Southern Sulawesi, described by van den Berg (1989)
- These restrictions have interesting resemblances to, and differences from, the better-known Semitic restrictions

### 3.2 The Muna data

- Inspired by Uhlenbeck's (1949) study of Javanese, van den Berg (1989: 30-31) examined the consonantal phonotactic constraints in CVCV roots (based on 1,100 roots)
- As in the closely related Timugon Murut (Prentice 1971); as well as Gurindji (McConvell 1988 and Evans 1995), van den Berg found the following gap in Muna:

(65) Prenasalized plosives do not co-occur

- And as in Sundanese (Cohn 1992):

(66) Unlike liquids do not co-occur

- As far as place-related restrictions, van den Berg notes the following:

(67) "Obstruents and prenasalized consonants do not co-occur with homorganic nasals"  
 "Initial plosives do not co-occur with contra-voiced homorganic plosives"

- The restrictions in (67) suggest a broader generalization: a ban on non-identical homorganic segments, as in Semitic (Greenberg 1950) and Javanese (Uhlenbeck 1949)
- The following table, based on van den Berg's (1989) data, supports this generalization (exceptions are bolded):

(68) Muna homorganic consonant co-occurrence patterns

| C1 | Occurring (n≥2) | Non-occurring (n=0)     | Marginal (n=1) | Total <sup>3</sup> |
|----|-----------------|-------------------------|----------------|--------------------|
| p- | p-p             | p-b, p-ḃ, p-f, p-w, p-m |                | 84                 |
| b- | b-b             | b-p, b-ḃ, b-f, b-w, b-m |                | 50                 |
| ḃ- | ḃ-ḃ, <b>ḃ-w</b> | ḃ-p, ḃ-b, ḃ-f, ḃ-m      |                | 73                 |
| f- | f-f, <b>f-p</b> | f-b, f-ḃ, f-w, f-m      |                | 17                 |
| w- | w-w             | w-p, w-b, w-f, w-m      | w-ḃ            | 59                 |
| m- | m-m             | m-p, m-b, m-ḃ           | m-f, m-w       | 44                 |
| k- |                 | k-k, k-g, k-ɣ, k-ŋ      |                | 86                 |
| g- | g-g             | g-k, g-ɣ, g-ŋ           |                | 46                 |
| ɣ- | ɣ-ɣ             | ɣ-k, ɣ-g, ɣ-ŋ           |                | 62                 |
| ŋ- |                 | ŋ-ŋ, ŋ-k, ŋ-ɣ, ŋ-g      |                | 5                  |

<sup>3</sup> Van den Berg provides a percentage representing the relative occurrence in initial position of each consonant. The totals were arrived at by multiplying that percentage by 1100.

|    |                                     |          |               |     |
|----|-------------------------------------|----------|---------------|-----|
| t- | <b>t-t, t-d, t-s, t-l, t-r, t-n</b> |          |               | 127 |
| d- | <b>d-d, d-t, d-l, d-n</b>           |          | d-s, d-r      | 50  |
| s- | <b>s-s, s-d, s-l, s-r</b>           | s-t, s-n |               | 89  |
| l- | <b>l-l, l-t, l-d, l-s, l-n</b>      |          | l-r           | 97  |
| r- | <b>r-r, r-t, r-d, r-s, r-n</b>      | r-l      |               | 72  |
| n- | <b>n-t, n-d, n-l</b>                |          | n-s, n-n, n-r | 20  |

- Based on these data, one might conclude that coronals are exempt from the restriction
- In collaboration with van den Berg and Andries Coetzee, I have been analyzing a larger set of data from van den Berg's dictionary of Muna
- We selected all of the CVCV and CVCVCV roots minus those with prenasalized stops, which yielded a total of 4901 roots
- Co-occurrence examined between contiguous consonants
- Observed/Expected, for each place of articulation:

(69) **Coronal**

|          |          |          |          |          |          |          |
|----------|----------|----------|----------|----------|----------|----------|
|          | <b>t</b> | <b>d</b> | <b>s</b> | <b>l</b> | <b>r</b> | <b>n</b> |
| <b>t</b> |          |          |          |          |          |          |
| <b>d</b> | 0.74     |          |          |          |          |          |
| <b>s</b> | 0.43     | 0.66     |          |          |          |          |
| <b>l</b> | 1.00     | 1.04     | 1.40     |          |          |          |
| <b>r</b> | 1.07     | 1.05     | 1.27     | 0.25     |          |          |
| <b>n</b> | 0.83     | 0.30     | 1.35     | 0.40     | 0.66     |          |

(70) **Labial**

|          |          |          |          |          |          |          |
|----------|----------|----------|----------|----------|----------|----------|
|          | <b>p</b> | <b>b</b> | <b>ɸ</b> | <b>f</b> | <b>w</b> | <b>m</b> |
| <b>p</b> |          |          |          |          |          |          |
| <b>b</b> | 0.10     |          |          |          |          |          |
| <b>ɸ</b> | 0.12     | 0.30     |          |          |          |          |
| <b>f</b> | 0.20     | 0.53     | 0.20     |          |          |          |
| <b>w</b> | 0.17     | 0.00     | 0.43     | 0.58     |          |          |
| <b>m</b> | 0.41     | 0.07     | 0.00     | 0.94     | 0.09     |          |

(71) **Dorsal**

|          |          |          |          |          |          |
|----------|----------|----------|----------|----------|----------|
|          | <b>k</b> | <b>g</b> | <b>ɠ</b> | <b>h</b> | <b>ŋ</b> |
| <b>k</b> |          |          |          |          |          |
| <b>g</b> | 0.08     |          |          |          |          |
| <b>ɠ</b> | 0.44     | 0.00     |          |          |          |
| <b>h</b> | 0.77     | 1.01     | 0.47     |          |          |
| <b>ŋ</b> | 0.11     | 0.00     | 0.00     | 0.70     |          |

- Observations:

(72) **Coronal effect**

Pairs of non-identical coronals *are* underrepresented in Muna, but they occur more frequently than dorsals or labials

e.g. t-d > k-g, p-b; t-n > p-m, k-ŋ; d-n > b-m, g-ŋ

(73) **Similarity effects**

*Voicing 1a*: Nasals co-occur less frequently with voiced stops than with voiceless ones  
t-n > d-n; p-m > b-m; k-ŋ > g-ŋ

*Voicing 1b*: Voiced stops co-occur less frequently with nasals than with voiceless stops  
t-d > d-n; p-b > b-m; k-g > g-ŋ

*Voicing 2*: Fricatives occur less frequently with voiceless stops than with voiced ones  
d-s > t-s; b-f > p-f; g-h > k-h

*Sonorancy*: The nasal [n] co-occurs less frequently with [l] and [r] than do [d] and [t].  
Furthermore, [l] and [r] co-occur less frequently with each other than with [n]

*Voicing plus manner*: Nasals and voiceless fricatives co-occur freely

- These data provide general support for Frisch, Broe and Pierrehumbert's (2004) proposal that OCP effects increase with similarity, and especially for the claim that redundant features count in this assessment (much better evidence for redundant voice in Muna than Arabic)
- They also demonstrate one way that languages can weight features differently in computing similarity: sonority is more important than voicing in Arabic, the reverse of Muna
- Muna is also of interest in comparison with Arabic because the coronals co-occur more freely in Muna, yet the set of coronals is only marginally bigger than labials or dorsals (cf. Frisch *et al.*'s 2004 discussion of Arabic)

### 3.3 Towards an account of gradient phonotactics

- This is an extremely sketchy analysis, which I am presenting in this form so as to have some concrete constraints to illustrate the issues that I'd like to discuss
- Along the lines of Alderete (1997), we could posit a set of OCP constraints divided by place, with dorsal and labial in a fixed ranking above coronal
- In Muna, dorsals seem to occur less freely than labials, so we have the following ranking:

(74) \*K-K >> \*P-P >> \*T-T

- Each of these constraints is subdivided into more specific constraints that apply when segments have the same feature value (or set of feature values)

- Here are some of the coronal constraints (ultimately need to deal with liquids too):

(75) \*T-T[+Vce] - Violated by [d-n], not by [t-d], [t-n], [t-s], [s-d], [s-n]  
 \*T-T[-Vce] - Violated by [t-s], not by [t-d], [d-n], [t-n], [s-d], [s-n]  
 \*T-T[-Cont] - Violated by [t-n], [t-d], [d-n], not by [t-s], [s-d], [s-n]  
 \*T-T[-Son] - Violated by [s-d], [t-s], [t-d] not [t-n], [d-n], [s-n]

- Rankings and observed lexical distributions:

(76) \*T-T[+Vce] >> \*T-T[-Vce] >> \*T-T[-Son], \*T-T[-Cont]

d-n (.30) < t-s (.43) < s-d (.66), t-d (.74), t-n (.83) < s-n (1.35)

- One problem - under this analysis, it's a coincidence that the similarity generalizations hold across places of articulation (i.e. why are both \*T-T[+Vce] and \*P-P[+Vce] dominant?)
- The bigger problem - all of these sequences are allowed in Muna, so all of these constraints are dominated by Faith. Standard OT does not distinguish between them.
- Maybe this isn't a problem - a theory of phonological grammar isn't necessarily obligated to account for statistical phonotactic generalizations (see e.g. Inkelas, Orgun and Zoll 1997)
- Pushed to its logical conclusion, though, this point of view would hold that the grammar for a language with a single word with a coda is identical to the grammar for a language that freely tolerates codas

(77) FAITH >> NOCODA

- With lexically indexed constraints, the language that has exceptional codas is instead:

(78) FAITH-L >> NOCODA >> FAITH

- To get finer grades of acceptability as in Muna, we could extend this approach (cf. Itô and Mester 1999):

(79) FAITH-L-A >> \*T-T[+VCE] >> FAITH-L-B >> \*T-T[-VCE] >> FAITH-L-C >>  
 \*T-T[-SON], \*TT[-CONT] >> FAITH-L-D

- The acceptability of a nonce form is then determined by asking what the odds are that it would surface faithfully, given all of the possible lexical indexations:

(80) **s-n:** 4/4 (1.35) **s-d /t-d/t-n:** 3/4 (.66, .74, .83) **t-s:** 2/4 (.43) **d-n:** 1/4 (.30)

- Three learnability problems:

(81) i. How do we get the markedness constraints in the right order?  
 ii. How do we create lexically specific constraints for exceptions to phonotactics?  
 iii. How do we get the faithfulness constraints interspersed correctly?

- The first of these problems might be solved by applying Boersma’s (1998) Gradual Learning Algorithm, which is sensitive to frequency - relatively large numbers of \*T-T[-SON] violators should push that constraint lower than \*T-T[+VCE] (see Zuraw 2000)
- However, the GLA does not detect inconsistency, which was crucial in section 2, and in the recent learnability results of Tesar and colleagues
- I’ll sketch one possible solution, using an abstract example parallel to those in §1 and 2
- In this language, there are generally no codas or clusters. However, there is a set of words with codas (here two words), and a smaller set of words with clusters (here one word):

(82) Words: [pa] [ti] [la] [mat] [net] [fle]

- The only novel assumption I will make is that the learner initially assumes that every word in the set of mark-data pairs is targeted by lexically specific faithfulness constraints, which later get merged
- Otherwise I simply follow Prince and Tesar’s (2004) Biased Constraint Demotion algorithm

(83) Faithfulness: MAX  
 Markedness: NO-CODA, \*COMPLEX

- There are no constraint conflicts for the unmarked CV words, so mark-data pairs will only be constructed for the CCV and CVC words:

(84)

| Input | W ~ L    | NOCODA | *COMPLEX | MAX-[mat] | MAX-[net] | MAX-[fle] |
|-------|----------|--------|----------|-----------|-----------|-----------|
| mat   | mat ~ ma | L      |          | W         |           |           |
| net   | net ~ ne | L      |          |           | W         |           |
| fle   | fle ~ fe |        | L        |           |           | W         |

- BCD iteratively places constraints in strata, starting at the top of the hierarchy

(85) i. Identify constraints that prefer no losers  
 ii. If any of these are markedness constraints, install them in the current stratum  
 iii. If there are no available markedness constraints, install faithfulness constraints

- In phonotactic acquisition, faithfulness constraints will never prefer a loser, since there is no evidence for divergent underlying representations
- Prince and Tesar (2004: 267) propose that the choice amongst faithfulness constraints is done by identifying ones that “free up” markedness constraints for ranking:

(86) “**Smallest effective F sets.** When placing faithfulness constraints into the hierarchy, place the *smallest set* of F constraints that *frees up some markedness constraint.*”

- According to (86), BCD will first install MAX-[fle]; this eliminates the mark-data pair for [fle] from further consideration:<sup>4</sup>

(87) Grammar:  
MAX-[fle] >>

(88)

| Input | W ~ L    | NoCODA | *COMPLEX | MAX-[mat] | MAX-[net] | MAX-[fle] |
|-------|----------|--------|----------|-----------|-----------|-----------|
| mat   | mat ~ ma | L      |          | W         |           |           |
| net   | net ~ ne | L      |          |           | W         |           |
| fle   | fle ~ fe |        | L        |           |           | W         |

- Next, \*COMPLEX will be installed due to the markedness bias, and then MAX-[mat] and MAX-[net] will be installed to free up NoCODA

(89) Grammar:  
MAX-[fle] >> \*COMPLEX >> MAX-[mat], MAX-[net]

(90)

| Input | W ~ L    | NoCODA | *COMPLEX | MAX-[mat] | MAX-[net] | MAX-[fle] |
|-------|----------|--------|----------|-----------|-----------|-----------|
| mat   | mat ~ ma | L      |          | W         |           |           |
| net   | net ~ ne | L      |          |           | W         |           |
| fle   | fle ~ fe |        | L        |           |           | W         |

- With no more data to account for, NoCODA will be installed next, and then the general MAX constraint:

(91) Grammar:  
MAX-[fle] >> \*COMPLEX >> MAX-[mat], MAX-[net] >> NoCODA >> MAX

- The lexically specific constraints could then be collapsed as follows:

(92) Merge instantiations of any constraint that occupy the same stratum and index the constraint to the respective lexical entries

(93) Grammar:  
MAX-L1 >> \*COMPLEX >> MAX-L2 >> NoCODA >> MAX

Lexicon:

/fle<sub>L1</sub>/ /mat<sub>L2</sub>/ /net<sub>L2</sub>/ /pa/ /ti/ /la/

- The success of this approach is encouraging, since it so minimally elaborates on Biased Constraint Demotion
- It also yields a grammar that accounts for “Impossible Nativizations” (Itô and Mester 1999)

---

<sup>4</sup> When we consider a realistic case with more than just one faithfulness constraint, some principle may be needed to choose between a set of instantiations of a single constraint, versus a set of different constraints.

- It does, however, have the strange property of creating a lexically specific constraint for a language with a full-fledged contrast
- That is, the grammars for the “one-coda” and “many-coda” languages will now both be:

(94) FAITH-L >> NOCODA >> FAITH

- It should also be possible to more tightly integrate the learnability accounts for exceptionality in phonotactics and exceptionality in alternations

#### **4. Conclusions**

- Exceptions provide many interesting challenges, both for phonological theory, and for learnability
- The aim here has been to show how many of these challenges can be met by assuming a fairly minimal elaboration of standard Optimality Theory: lexically indexed constraints

## References

- Albright, Adam. 2002. A restricted model of UR discovery: Evidence from Lakhota. Ms, University of California at Santa Cruz.
- Anttila, Arto. 2002. Morphologically Conditioned Phonological Alternations. *NLLT* 20, 1, 1-42.
- Berg, René van den. 1989. *A grammar of the Muna language*. Dordrecht: Foris.
- Chomsky, Noam and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.
- Coats, Herbert. 1970. Rule Environment Features in Phonology. *Papers in Linguistics* 2. 110-140.
- Flora, Marie Jo-Ann. 1974. *Palauan Phonology and Morphology*. Ph.D. dissertation, University of California at San Diego.
- Frisch, Stefan, and Bushra Zawadyeh. 2001. The Psychological Reality of OCP-Place in Arabic. *Language* 77. 91-106.
- Frisch, Stefan, Janet Pierrehumbert and Michael Broe. 2004. Similarity Avoidance and the OCP. *Natural Language and Linguistic Theory*. 179-228.
- Fukuzawa, Haruka. 1997. Multiple Input-Output Faithfulness Relations in Japanese. Ms., University of Maryland.
- Greenberg, J. 1950. The Patterning of Root Morphemes in Semitic. *Word* 6: 162-181.
- Hayes, Bruce. 1982. Extrametricality and English Stress. *Linguistic Inquiry* 13, 227-76.
- Hayes, Bruce. 2004. Phonological acquisition in Optimality Theory: the Early Stages. In R. Kager, J. Pater and W. Zonneveld (eds.), *Constraints in Phonological Acquisition*. Cambridge University Press.
- Inkelas, Sharon, Orhan Orgun and Cheryl Zoll. 1997. The implications of lexical exceptions for the nature of grammar. In *Derivations and Constraints in Phonology*, ed. I Roca, 393-418. New York: Oxford University Press.
- Itô, Junko and Armin Mester. 1995a. The Core-periphery Structure in the Lexicon and Constraints on Re-ranking. In Jill Beckman, Laura Walsh Dickey and Suzanne Urbanczyk, eds. *Papers in Optimality Theory*. GLSA, University of Massachusetts, Amherst. 181-210.
- Itô, Junko and Armin Mester. 1995b. Japanese phonology. In John Goldsmith, ed. *The Handbook of Phonological Theory*. Blackwell, Oxford. 817-38.
- Itô, Junko and Armin Mester. 1999. The Phonological Lexicon. In the *Handbook of Japanese Linguistics*, ed. N. Tsujimura. 62-100. Oxford: Blackwell.
- Kager, René. 1999. Ternary Alternations and Lexical Allomorphy. Paper presented at NELS 30.
- Kisseberth, Charles. 1970. The Treatment of Exceptions. *Papers in Linguistics* 2, 44-58.
- Kraska-Szlenk, Iwona. 1997. Exceptions in Phonological Theory. *Proceedings of the 16<sup>th</sup> International Congress of Linguists*. Pergamon, Oxford. Paper No. 0173.
- Kraska-Szlenk, Iwona. 1999. Syllable structure constraints in exceptions. In John Rennison and Klaus Kühnhammer, eds. *Phonologica 1996: Syllables !?*. The Hague: Thesus. 113-131.
- Mascaró, Joan. 2003. Comparative markedness and derived environments. *Theoretical Linguistics* 29. 113-122.
- McCarthy, John and Alan Prince. 1993. Generalized Alignment. *Yearbook of Morphology*. 79-153.
- McConvell, P. 1988. Nasal cluster dissimilation and constraints on phonological variables in Gurindji and related languages. In N. Evans and S. Johnson (eds.) *Aboriginal Linguistics* 1. Armidale: Department of Linguistics, University of New England.

- Nouveau, Dominique. 1994. *Language Acquisition, Metrical Theory, and Optimality: A Study of Dutch Word Stress*. PhD Dissertation, Utrecht University.
- Paradis, Carole and Caroline Lebel. 1994. Contrasts from Segmental Parameter Settings in Loanwords: Core and Periphery in Quebec French. In C. Dyck (ed.) *Toronto Working Papers in Linguistics* 13:1. 75-94.
- Pater, Joe. 1994. Against the Underlying Specification of an 'Exceptional' English Stress Pattern. In C. Dyck (ed.) *Toronto Working Papers in Linguistics* 13:1. 95-121.
- Pater, Joe. 2000. Nonuniformity in English stress: the role of ranked and lexically specific constraints. *Phonology* 17:2. 237-274.
- Piñeros, Carlos-Eduardo. 2004. Vowel Weightlessness and Stress Retraction in Spanish. Ms., University of Iowa.
- Prentice, D. J. *The Murut Languages of Sabah*. Pacific Linguistics Series C, no. 18. Canberra: Australian National University.
- Prince, Alan. 2002. Entailed Ranking Arguments. Ms., Rutgers University.
- Prince, Alan, and Bruce Tesar. 2004. Learning Phonotactic Distributions. In R. Kager, J. Pater and W. Zonneveld (eds.) *Constraints in Phonological Acquisition*. CUP.
- Rice, Keren. 1997. Japanese NC Clusters and the Redundancy of Postnasal Voicing. *Linguistic Inquiry* 28 (3): 541-51.
- Russell, Kevin. 1995. Morphemes and candidates in Optimality Theory. Ms., University of Manitoba.
- Schane, Sanford. 1973. The Formalization of Exceptions in Phonology. In Maurice Gross, Morris Halle, and Marcel-Paul Schützenberger, eds. *The Formal Analysis of Natural Languages*. Mouton: The Hague.
- Tesar, Bruce. 2000. Using Inconsistency Detection to Overcome Structural Ambiguity in Language Learning. Ms., Rutgers University.
- Tesar, Bruce, and Alan Prince. (to appear). Using Phonotactics to Learn Phonological Alternations. Revised version to appear in *The Proceedings of CLS 39, Vol. II: The Panels*. ROA-620.
- Tesar, Bruce, John Alderete, Graham Horwood, Nazarre Merchant, Koichi Nishitani, and Alan Prince. 2003. Surgery in language learning. In *The Proceedings of WCCFL 22*, pp. 477-490. ROA-619.
- Winslow, Nicholas. 2003. Incorporating Exceptions in Optimality Theoretic Learnability. Honor's Thesis, University of Massachusetts, Amherst.
- Zonneveld, Wim. 1978. *A Formal Theory of Exceptions in Generative Phonology*. Lisse: The Peter de Ridder Press.
- Zuraw, Kie 2000. *Patterned Exceptions in Phonology*. Ph.D. dissertation, UCLA.