# Chapter 4

# Phonological exceptionality is localized to phonological elements: the argument from learnability and Yidiny word-final deletion

## Erich R. Round

University of Queensland

Anderson (2008) emphasizes that the space of possible grammars must be constrained by limits not only on what is cognitively representable, but on what is learnable. Focusing on word final deletion in Yidiny (Dixon 1977a), I show that the learning of exceptional phonological patterns is improved if we assume that Prince & Tesar's (2004) Biased Constraint Demotion (BCD) with Constraint Cloning (Pater 2009) is subject to a Morphological Coherence Principle (MCP), which operationalizes morphological analytic bias (Moreton 2008) during phonological learning. The existence of the MCP allows the initial state of Con to be simplified, and thus shifts explanatory weight away from the representation of the grammar per se, and towards the learning device.

I then argue that the theory of exceptionality must be phonological and diacritic. Specifically, I show that co-indexation between lexical forms and lexically indexed constraints must be via indices not on morphs but on individual phonological elements. Relative to indices on phonological elements, indices on morphs add computational cost for no benefit during constraint evaluation and learning; and a theory without indices on phonological elements is empirically insufficient. On the other hand, approaches which represent exceptionality by purely phonological means (e.g. Zoll 1996) are ill-suited to efficient learning. Concerns that a phonologically-indexed analysis would overgenerate (Gouskova 2012) are unfounded under realistic assumptions about the learner.

## 1 Exceptionality

What is the nature of representations which are passed from the morphology to the phonology? Anderson (1992) demonstrates that the processes that create those representations can be elaborate and complex. Operations that act upon morphological forms, to realize units of morphologically-relevant meaning, involve not only the concatenation

of formatives, but also selection among alternatives and non-concatenative modifications to intermediate representations (see also Anderson 2015; 2016; 2017). However, what of the final result, which comprises some number of morphs that must then be interpreted phonologically? A constant concern of generative phonology since its inception has been to account adequately for patterned phonological exceptionality, the phenomenon in which segments in a restricted class of morphs exhibit phonologically distinctive behavior as triggers, targets or blockers of alternations, or as participants in exceptional featural, phonotactic or prosodic surface structures. For example, in Yidiny (Dixon 1977a,b) vowels delete word-finally, if that deletion would prevent the word from surfacing with an unfooted syllable. This is seen in the root *gaɟara-* 'possum' in (1a) and the suffix *-ɲa* ACCUSATIVE in (1b), where feet are marked by parentheses. However, in a restricted set of morphs the final vowel behaves exceptionally, resisting deletion, as in the root *guɟara-* 'broom' (2a) and the suffix *-na* PURPOSIVE (2b).

(1)  a.  'possum.ABS'   /gaɟara/   →   (ga ɟa:r)
     b.  'father-ACC'   /bimbi-ɲa/   →   (bim bi:ɲ)

(2)  a.  'broom.ABS'   /guɟara/   →   (gu ɟa:) ra
     b.  'go-PURP'   /gali-na/   →   (ga li:) na

In order for the phonology to treat morph-specific, exceptional segments appropriately, it must receive from the morphology some kind of discriminating information which it can act upon. For much of the generative period it has been argued that this information is associated with morphs as a whole, and not with their individual phonological elements. Here I present an argument for the contrary view. The contribution, then, is to clarify the nature of one important aspect of the interaction between the morphological and phonological components of grammar. The principle line of evidence is learnability, namely the learnability of an optimality-theoretic grammar for phonological exceptionality. Anderson (2008) has emphasized that the space of possible human grammars must be constrained not only by limits on what is cognitively representable, but also on what is learnable. The crux of the argument here relies not on specifics, but ultimately on general properties of learnable grammars, and thus I would hope should remain valid even as specific theories undergo refinement as they move closer to answering Anderson's (2008) challenge.[1]

The chapter falls into two broad parts. In §2–§5 I discuss the processes and principles required to learn exceptionality. This leads to the positing of a Morphological Coherence

---

[1] A reviewer asks whether the machinery presented here is necessary if one assumes an exemplar-based model of phonology. I assume that learners do store rich, exemplar-like representations of linguistic experiences. However, natural language morphology in general has enough combinatorial complexity that reliance upon retrieved episodes will not be sufficient to reproduce the full range of creative behavior that humans display. Consequently some generative machinery is necessary, which performs not merely simple analogies and concatenations, but which can reproduce with precision the complex patterns generated by a realizational morphology such as Anderson's (1992), and by a formal phonological grammar such as entertained here.

Principle in §6, which operationalizes a morphological bias that ensures successful learning for certain cases. In §7–§9 I am concerned with the underlying theory of these processes and principles. I evaluate two broad approaches to phonological exceptionality: PHONOLOGICAL approaches, which represent exceptionality as a property of individual segments (Bloomfield 1939; Kiparsky 1982a; Inkelas 1994; Zoll 1996), and MORPHOLOGICAL approaches which represent it as a property of morphs (Chomsky 1964; Chomsky & Halle 1968; Zonneveld 1978; Pater 2000). The result is an argument in favor of a DIACRITIC PHONOLOGICAL approach. On this account, exceptionality is represented at the level of individual phonological elements, not morphs; however the means of marking it is by diacritics which are visible to the phonology but not manipulable by it, in contradistinction to the CONCRETE PHONOLOGICAL approach, where the crucial representations are themselves phonological elements. As I show, the function of these "Φ-indices" is essentially identical to "M-indices" which would mark morphs, only there is no assumption that all exponents of a morph *m* be indexed identically. As we shall see, freedom from that assumption is both coherent theoretically and desirable, computationally and empirically. The discussion is illustrated throughout by the facts of word final deletion in Yidiny, to which we turn now in §2.

# 2 Word-final deletion in Yidiny

## 2.1 The phenomenon

Yidiny (Dixon 1977a) belongs to the Yidinyic subgroup of the Pama-Nyungan language family. Traditionally it was spoken in the rainforest region southwest of Cairns, in Northeastern Australia. Most examples below are from Dixon's (1977) detailed descriptive grammar; examples marked † are from Dixon's (1991) dictionary and texts. An inventory of underlying segments is in Table 1.

Table 1: Yidiny underlying segments, after Dixon (1977a: 32).

|  | Labial | Apical | Laminal | Dorsal |
|---|---|---|---|---|
| **Stop** | b | d | ɟ | g |
| **Nasal** | m | n | ɲ | ŋ |
| **Lateral, trill** |  | l, r |  |  |
| **Approximant** | w | ɻ | y |  |
| **Vowels** | i, a, u, iː, aː, uː | | | |

Syllable shapes are tightly constrained. Onsets are obligatory and simple. Codas permit only sonorants other than /w/. Codas in word-final position are simple; word-internal codas also permit disegmental, continuant–nasal sequences. Morphologically, the language is almost entirely suffixing and largely agglutinative. Roots are minimally

disyllabic and suffixes are maximally disyllabic (Dixon 1977a: 35,90). An online appendix[2] discusses the morphological constituency of verbal inflection.

Of Yidiny's phonological alternations, those to receive the greatest attention have been stress placement, vowel length and to a lesser extent, word-final deletion (Dixon 1977a,b; Hayes 1982; 1985; Kager 1993; Crowhurst & Hewitt 1995; Halle & Idsardi 1995; Hall 2001; Pruitt 2010; Hyde 2012; Bowern, Round & Alpher in revision, *inter alia*). Yidiny's stress and length alternations in particular have featured in significant theoretical works on meter and prosody over the past four decades, and both are nontrivial topics in themselves. Word-final deletion, however, can be studied largely independently of them for reasons that follow.

Although stress placement in Yidiny has proven contentious (Pruitt 2010; Bowern, Round & Alpher in revision), word-final deletion is not sensitive to stress *per se*, but rather only to the position of foot boundaries. These have been uncontroversial since their analysis by Hayes (1982): feet in Yidiny are disyllabic and left-aligned within the phonological word.

Many words with word-final deletion also exhibit vowel lengthening; however the phenomena show little to no mutual interaction. In a rule-based theory permitting simultaneous application (Anderson 1974) lengthening and deletion would apply simultaneously; neither rule feeds or bleeds the other.[3] See Round (in progress) for an analysis of Yidiny lengthening.

Word-final deletion is sensitive to foot placement, and foot placement is sensitive to phonological word boundaries. In Yidiny, phonological words commence at the left edge of each root and each disyllabic suffix (Dixon 1977a: 88–98).[4] Phonological words therefore begin with either a polysyllabic root or a disyllabic suffix and are followed by zero or more monosyllabic or entirely consonantal suffixes. Word-final deletion targets unfooted syllables and therefore only affects prosodic words which, *modulo* deletion, would be at least trisyllabic. As a consequence, we are interested here in three kinds of phonological word: those comprised of bare roots of three or more syllables; those comprised of roots plus one or more monosyllabic suffixes; and those comprised of a disyllabic suffix plus one or more additional, monosyllabic suffixes. The third kind is rare,[5] and so discussion will focus on the first two.

Word-final deletion applies only if the word thereby avoids surfacing with an unfooted syllable. For example, the roots *gindanu-* 'moon' and *gubuma-* 'black pine' both contain three vowels, each of which is a potential syllabic nucleus at the surface. In (3a,4a) they have undergone deletion of their final vowel to prevent it from surfacing in an unfooted syllable; compare (3b,4b) where the roots are non-final in the word, and the final vowels surface.

---

[2] Available from `10.6084/m9.figshare.4579696`

[3] Deletion counter-bleeds lengthening, thus in a strictly serial analysis lengthening would precede word-final deletion (Dixon 1977a,b; Hayes 1985; Crowhurst & Hewitt 1995).

[4] Yidiny's only prefix, [ɟa:-] 'in a direction' occupies its own phonological word (Dixon 1977a: 98,162).

[5] For an illustration, see example (25).

(3) a. 'moon[ABS]'  /gindanu/  →  (gin da:n)
   *(gin da:) nu
  b. 'moon-ERG'  /gindanu-ŋgu/  →  (gin da) (nuŋ gu)

(4) a. 'black pine[ABS]'  /gubuma/  →  (gu bu:m)
   *(gu bu:) ma
  b. 'black pine-PURP'  /gubuma-gu/  →  (gu bu) (ma gu)

Final vowel deletion may also affect suffixes. In (5a,c,6a), the vowels of the nominal comitative suffix -*yi* and verbal comitative suffix -*ŋa* have undergone deletion, thereby preventing the surfacing of an unfooted syllable. In (5b,6b) the suffixes are non-final in the word, and the vowel surfaces.

(5) a. 'woman-COM'  /buɲa-yi/  →  (bu ɲa:y)
   *(bu ɲa:) yi
  b. 'woman-COM-ERG'  /buɲa-yi-ŋgu/  →  (bu ɲa) (yiŋ gu)
  c. 'black bream-COM'  /gulugulu-yi/  →  (gu lu) (gu lu:y)
   *(gu lu) (gu lu:) yi

(6) a. 'come-COM[IMP]'  /gada-ŋa/  →  (ga da:ŋ)
   *(ga da:) ŋa
  b. 'come-COM-PST'  /gada-ŋa-lɲu/  →  (ga da:) (ŋal ɲu)

Word-final deletion interacts with restrictions on word-final consonants, and the interaction plays out differently in roots versus suffixes. In roots, deletion will fail to apply if the result would be an illicit word-final coda, containing either a stop or /w/ (7) or a cluster (8). One conceivable alternative, to also delete the consonant, is not attested in roots (7–8).[6]

(7) a. 'man[ABS]'  /waguɟa/  →  (wa gu:) ɟa
   *(wa guɟ)
   *(wa gu:)
  b. 'dog[ABS]'  /gudaga/  →  (gu da:) ga
   *(gu da:g)
   *(gu da:)
  c. 'sugar ant[ABS]'  /balawa/  →  (ba la:) wa
   *(ba la:w)
   *(ba la:)
  d. 'place name[ABS]'  /ŋalumba/  →  (ŋa lu:m) ba
   *(ŋa lu:mb)
   *(ŋa lu:m)

---

[6] Neither Dixon's grammar (1977) nor dictionary (Dixon 1991, which cites underlying forms) records a surface form for the roots in (7c) and (7d), or for roots illustrating the same pre-final consonant or comparable consonant clusters. However, Dixon (1977a: 57–58) specifically reports that the roots *balawa-* and *gindalba-* do not undergo deletion; the surface forms provided here are what we would expect if this is so.

(8)  'warn[IMP]'  /binarŋa/  →  (bi na:r) ŋa
          *(bi na:rŋ)
          *(bi na:r)

In contrast, deletion in suffixes applies not only to the final vowel, but also to a single consonant that precedes it, if that consonant would be illicit word-finally, as in (9). This form of CV deletion respects phonotactic constraints while also avoiding unfooted syllables.[7]

(9)  a.  'grey possum-ERG'  /margu-ŋgu/      →  (mar gu:ŋ)
     b.  'see-PST'          /wawa-lɲu/        →  (wa wa:l)
     c.  'warn-DAT.SUB'     /binarŋa-lɲu-nda/ →  (bi nar) (ŋal ɲu:n)

However, word-final deletion never deletes the initial segment of a suffix (and consequently, it will never delete an entire suffix), as illustrated in (10).

(10)  a.  'woman-SET'     /buɲa-ba/   →  (bu ɲa:) ba
           *(buɲ ba)
      b.  'bandicoot-GEN' /guygal-ni/ →  (guy ga:l) ni
           *(guy ga:ln)
           *(guy ga:l)

Deletions do not occur word internally (11a,b), nor do word-final, licit codas delete (11b). All Yidiny roots and suffixes that are consonant-final end underlyingly with licit coda consonants, so no morph undergoes spontaneous deletion of an underlyingly-final consonant (11c).

(11)  a.   'woman-SET'  /buɲa-ba/  →  (bu ɲa:b)
            *(bu ɲa:)
      b.[†] 'name[ABS]'  /bagiram/  →  (ba gi:) ram
            *(ba gi:rm)
            *(ba gi:r)
      c.                */bagirag/  →  *(ba gi:r)

To summarize, word-final deletion applies only so as to avoid the surfacing of unfooted syllables. It may delete the final vowel from a root and the final (C)V sequence from a suffix, but will not delete a suffix-initial segment. Deletion is blocked (in roots) or expanded (in suffixes, from V deletion to CV deletion) in order to obey phonotactic restrictions on word-final codas. These are the regular conditions under which word-final deletion occurs.

   In addition to its regular application, Yidiny contains roots and suffixes which are exceptional non-undergoers of word-final deletion. In (13), the non-undergoer roots

---

[7] The "dative subordinate" is marked by what Round (2013: 26) has called "compound suffixation", comprising two monosyllabic suffixal morphs, /-lɲu; -nda/. That these are not a single, disyllabic suffix is evident in the fact that they fail to be parsed into a their own phonological word, separate from the root.

*mulari-, guɟara-, ɟudulu-, baŋgamu-* all resist word-final deletion despite their pre-final consonant being permissible as a coda, and despite the fact that the consequence is an unfooted, word-final syllable.

(12)  a.  'initiated man[ABS]'  /mulari/  →  (mu la:) ri
                                              *(mu la:r)
      b.  'broom[ABS]'  /guɟara/  →  (gu ɟa:) ra
                                       *(gu ɟa:r)
      c.  'brown pigeon[ABS]'  /ɟudulu/  →  (ɟu du:) lu
                                             *(ɟu du:l)
      d.  'potato[ABS]'  /baŋgamu/  →  (baŋ ga:) mu
                                        *(baŋ ga:m)

Dixon (1977a: 59) reports 115 trisyllabic roots whose phonotactic shape would, under regular conditions, expose them to word-final deletion. Of these, 34, or around 30%, are exceptional non-undergoers. The distinction is idiosyncratic; neither Dixon (1977a: 58) nor subsequent researchers have found any phonological, semantic or grammatical factor that categorically determines whether a root will be a non-undergoer.[8]

Suffixes also may be exceptional non-undergoers. In (17) the non-undergoer suffixes *-nda*, *-lɟi* and *-na* resist word-final deletion and allow an unfooted syllable to surface. Avoidance of regular, word-final CV deletion is seen in (13a,b) and V deletion in (13c).

(13)  a.  'grey possum-DAT'  margu-nda  →  (mar gu:n) da
                                            *(mar gu:n)
      b.  'see-LEST[ABS]'  wawa-lɟi  →  (wa wa:l) ɟi
                                          *(wa wa:l)
      c.  'go-PURP'  gali-na  →  (ga li:) na
                                  *(ga li:n)

Tables 2 and 3 list all suffixal allomorphs in Yidiny which, on phonotactic grounds, could plausibly delete.[9] Regular undergoers are in Table 2 and non-undergoers in Table 3.

---

[8] Historically speaking, borrowed forms may account for many of these items (Barry Alpher p.c.); synchronically, however, their motivation is opaque.

[9] Such suffixes must be vowel-final and monosyllabic. If just the final vowel is to delete, then it must leave behind a single, licit-coda consonant in word final position. This will require the suffix to be -CV, and be preceded by a vowel, not a consonant. Alternatively, if the final CV is to delete, then the suffix must be -CCV, since suffix-initial segments do not delete, and it too must attach to a vowel-final stem. Data here is from a comprehensive search of Dixon (1977a), in which relevant information can be found on pp.50–54, 151. "Emphatic" *-ɲa* (Dixon 1977a: 151) is excluded. It behaves as a phonological clitic that occupies a distinct phonological word, and does not undergo final deletion.

Table 2: Monosyllabic suffixes which undergo word-final deletion.

| Function | | -CV | -CCV |
|---|---|---|---|
| Case | ERGATIVE | | -ŋgu |
| | LOCATIVE | -la | |
| | ACCUSATIVE | -ɲa | |
| | COMITATIVE | -yi | |
| | GENITIVE | -ni, -nu | |
| Verbal | PAST tense inflection | -ɲu | -lɲu, -ɻɲu |
| | COMITATIVE derivation | -ŋa | |
| | DATIVE SUBORDINATE inflection | | -nda[10] |

Table 3: Monosyllabic suffixes which escape word-final deletion.

| Function | | -CV | -CCV |
|---|---|---|---|
| Case | DATIVE | | -nda |
| Verbal | PURPOSIVE inflection | -na | -lna, -ɻna |
| | LEST nominalizing derivation | | -nɟi, -lɟi, -ɻɟi |

Exceptional non-undergoers, both roots and suffixes, only block the deletion of their own segments; the exceptionality does not spread to neighboring morphs. Accordingly in (14), the exceptional non-undergoer LEST does not block deletion in the following, regular undergoer, ERGATIVE suffix.

(14)  wiwi-ɟi-nɟi-ŋgu  'give-ANTIP-LEST-ERG'  →  (wi wi:) (ɟin ɟi:ŋ)
                                                  *(wi wi:) (ɟin ɟi:ŋ) gu

Likewise, the presence of a regular undergoer will not undo the blocking effect of an exceptional non-undergoer. In (15) the regular undergoer COMITATIVE does not undermine the blocking of deletion in the exceptional non-undergoer PURPOSIVE, which follows it.

(15)  maɟinda-ŋa-lna  'walk up-COM-PURP'  →  (ma ɟin) (da ŋa:l) na
                                               *(ma ɟin) (da ŋa:l)

It will be recalled that roots in Yidiny can undergo word-final deletion of vowels, but not of the consonants that precede them. More specifically, roots that end in CCV do not delete final CV, whereas some suffixes do, and nor does final C′V delete from roots that end in VC′V, where C′ is an impermissible coda. Two conceivable accounts for this

---

[10] The dative subordinate is marked by a string of two monosyllabic suffixes -*lɲu-nda*, cf. fn.7.

may be distinguished. On one account, the grammar of Yidiny expressly prohibits root-final CV deletion. On the other, it happens just by chance that all CCV-final and VC′V-final roots are exceptional non-undergoers. On the latter account, the grammar WOULD enforce CV deletion from roots, if only the lexicon provided the right inputs; on the former account it would not. The level of empirical support for these hypotheses can be assessed statistically. Table 4 compares counts of CCV- and VC′V-final roots and CCV-final suffixes which either do or do not delete. The distribution is strongly unbalanced, and we can reject with confidence the null hypothesis that it is due to chance ($\chi^2_{df=1}$ = 47.9 $p < 10^{-10}$). Table 5 compares counts of roots that are CCV- and VC′V-final with those that are VCV-final, i.e., where C is a permissible coda. Again, the counts are highly unbalanced and we reject the hypothesis that the absence of deletion in CCV- and VC′V-finals is by chance ($\chi^2_{df=1}$ = 125.8. $p < 10^{-10}$). The only empirically-supported conclusion is that the lack of consonant deletion in Yidiny roots is systematic, not due to chance. A satisfactory formal analysis should reflect this.[11]

Table 4: Deletion of coda-ilicit pre-final C in roots versus suffixes.

|  | **CCV- and VC′V-final roots** | **CCV-final suffixes** |
|---|---|---|
| No deletion | 116 | 6 |
| Deletion | 0 | 4 |

Table 5: Deletion in roots with pre-final coda-illicit C versus prefixal coda-licit C.

|  | **CCV- and VC′V-final roots** | **VCV-final suffixes** |
|---|---|---|
| No deletion | 116 | 34 |
| Deletion | 0 | 81 |

## 2.2 Constraint rankings

A briefly sketch now follows of how the facts above would be analysed in OT. Foot placement in Yidiny is due to FootBinarity ≫ ParseSyllable ≫ Align(Ft,L,PrWd,L) (Prince & Smolensky 2004[1993], McCarthy & Prince 1993a, McCarthy & Prince 1995). Of these, only ParseSyllable (Prs) will be of interest for our purposes; I assume that other prosodic constraints are satisfied optimally. Absolute re-

---

[11] As a reviewer observes, there is an interesting historical background to be clarified here, and an account of it is planned. Naturally, the object of a synchronic analysis differs ontologically from that of a historical one. The two are complementary, but neither account would substitute for or serve as a counter-analysis to the other.

strictions against obstruents and /w/ in codas are due to Sonorant/Coda (e.g. Lombardi 2002) and *w/Coda; I assume these are unviolated.

Regular word-final deletion in Yidiny can be analysed straightforwardly by ranking Prs ≫ Maximality (Max, McCarthy & Prince 1995). This causes deletion of final vowels in preference to the surfacing of unfooted syllables, but not if an illicit coda results.

Segments may delete from the right edge of the word only, not the left or word-internally. High-ranking Anchor-Left(morph) penalizes deletion from the left edge of any morph and Contig-IO(PrWd) penalizes deletion internally (McCarthy & Prince 1995).

Yidiny permits complex codas word-internally, but not word-finally. Ranking Cntg ≫ *ComplexCoda (Bernhardt & Stemberger 1998) accounts for this; ranking both above Prs accounts for the absence of deletion after pre-final clusters in roots and the defeat of candidates which delete only a final vowel from word-final CCV suffixes.

Word-final deletion applies differently to roots and suffixes. Roots will not undergo consonant deletion, even if the consequence is an unfooted syllable. The ranking of undominated Max-C/root (McCarthy & Prince 1995) above Prs accounts for this. Suffixes do not violate Max-C/rt and consequently are free to undergo consonant deletion, however highly-ranked Anc penalizes the deletion of morph-initial segments. This accounts for the fact that a consonant may delete from a -CCV suffix but not from -CV.

At this point, regular word-final deletion occurs whenever satisfaction of the markedness constraint Prs requires the violation of the lower-ranked faithfulness constraint Max. Deletion is blocked unexceptionally whenever Prs itself is violated in order to satisfy higher-ranking constraints, which are of two kinds: those which penalize marked codas, son/Coda, *w/Coda, *Cplx; and those which penalize deletion in specific morphological contexts, namely at left edges of morphs, Anc, and consonants in roots, Max-C/rt. We see that the driver of word-final deletion in Yidiny is the ranking of Prs ≫ Max. Deletion occurs when Prs is satisfied but Max is not. Regular blocking results when Prs must be violated, in which case Max can be satisfied.

Exceptional non-undergoers avoid deletion. For them, Max is always satisfied, even at the expense of Prs. Consequently, while regular undergoers are subject to a ranking of Prs ≫ Max, exceptional non-undergoers must be subject to Max ≫ Prs. In §4 I consider two approaches that will ensure this is the case, one morphological and one phonological. First though, a remark about constraint violations.

## 3 Relativized constraint violation

I introduce here a simple expression for relating the violations of certain pairs of constraints, which will aid discussion in later sections.

For any constraint C and candidate cand, there will be zero or more violations of C. Given the definition of C, those violations will be due to certain parts, or loci, in cand, either in the output of cand or in the correspondences between input and output elements (McCarthy & Prince 1995). We can define the set of loci of violation, V(C, cand), as the loci in cand which cause violations of C (McCarthy 2003, Łubowicz 2005). Now,

some pairs of constraints $C_1$, $C_2$ are related such that for any CAND, the loci of violation of $C_2$ are a subset of the loci of violation of $C_1$. In many cases, the latter are precisely those members of the former which also contain some particular kind of phonological element. For example V(MAX-C, CAND) are those members of V(MAX, CAND) which also contain input consonants. In that case, we can express V($C_2$, CAND) terms of the INTERSECTION of the set V($C_1$, CAND) and some appropriately defined second set, that picks out loci containing the criterial elements. Let us define the set of "$\varphi$-loci", $L_\varphi$(D($\varphi$), CAND), as the set of loci in CAND that contain a phonological element $\varphi$ of the kind denoted by predicate D($\varphi$). For example, V(MAX-C, CAND) can be defined in relative terms, as in (16), where the predicate INPUT_CONSONANT($\varphi$) denotes input consonants. (For brevity I omit the "CAND" from the expression for each set.)

(16)    V(MAX-C) $=_{\text{def}}$ V(MAX) $\cap$ $L_\varphi$(INPUT_CONSONANT($\varphi$))

This relativized method will be used below to define new constraints $C_N$ in terms of a reference constraint, $C_R$, and a set of phonological elements which restrict the violations of $C_N$ relative to those of $C_R$.

# 4 Preliminary analysis of word-final deletion

## 4.1 A morphological approach

We now consider an OT implementation of the morphological approach to Yidiny exceptionality, using lexically indexed constraints (Pater 2000; 2006; 2009). A lexically indexed constraint $C_M$ behaves precisely like its unindexed counterpart, C, except that it can be violated only by structures which contain exponents of a specific set $M$ of morphs, each of which has been assigned a diacritic mark which I will term a LEXICAL M-INDEX, that co-indexes it to $C_M$. The definition can be expressed relatively as in (17), following a similar formulation by Finley (2010).

(17)    V($C_M$) $=_{\text{def}}$ V(C) $\cap$ $L_\varphi$($m{\in}M$ & EXP($\varphi$, $m$)), *where*:
        $M$ is the set of morphs co-indexed to $C_M$.
        EXP($\varphi$, $m$) states that element $\varphi$ is an exponent of morph $m$

If we now define two sets of Yidiny morphs, $U$ the set of regular undergoers of word-final deletion, and $N$ the set of exceptional non-undergoers, then either of the rankings in (18) will ensure that the correct sets of morphs is subject to the desired partial ranking of PRS and MAX.

(18)    a.  PRS$_U$ $\gg$ MAX $\gg$ PRS
        b.  MAX$_N$ $\gg$ PRS $\gg$ MAX

In (18a), all phonological exponents of undergoer morphs will be subject to PRS$_U$ $\gg$ MAX, and non-undergoers to MAX $\gg$ PRS. In (18b), all phonological exponents of non-undergoer morphs will be subject to MAX$_N$ $\gg$ PRS, and undergoers to PRS $\gg$ MAX. For

now I will use ranking (18a); the reason for this will become clear in §5.[12,13] Examples in (19a–19b) illustrate word-final deletion of regular undergoers which are indexed *U*, the root *malanu-$_U$* and suffix ERGATIVE *-ŋgu$_U$*, while (19c–19d) show the absence of deletion for exceptional non-undergoers *mulari-* 'initiated man' and DATIVE *nda*.

(19)

|  |  |  | Prs$_U$ | Max | Prs |
|---|---|---|---|---|---|
| a. | /malanu$_U$/ 'right hand[ABS]' (ma la:n) > (ma la:) nu | | W | L | W |
| b. | /margu-ŋgu$_U$/ 'grey possum-ERG' (mar gu:ŋ) > (mar gu:ŋ) gu | | W | L | W |
| c. | /mulari/ 'initiated man[ABS]' (mu la:) ri > (mu la:r) | | | W | L |
| d. | /margu-nda/ 'grey possum-DAT' (mar gu:n) da > (mar gu:n) | | | W | L |
| e. | /maɟinda-ŋau$_U$-lna/ 'walk up-COM-PURP' (ma ɟin) (da ŋa:l) na > (ma ɟin) (da ŋa:l) | | | W | L |

Example (19e) illustrates the fact that violations of Prs$_U$ require not merely the presence of a *U*-indexed morph in the word, but a locus of violation which contains a phonological exponent of a *U*-indexed morph (17). Namely, the final syllable of (19e), *na*, is unfooted. However since that syllable contains no phonological exponent of a *U*-indexed morph, no violation of Prs$_U$ results. This is true despite the presence of a *U*-indexed morph elsewhere in the word.

## 4.2 A phonological approach

The phonological approach correlates the (un)exceptionality of a segment with representational properties of the segment itself. Implementations differ as to which property is used. Zoll (1996) analyses segments which resist deletion as having root nodes in their input, whereas segments that delete more readily lack root nodes, and are termed SUBSEGMENTS. Under these assumptions, a ranking MAX(SEG) ≫ PRS ≫ MAX(SUBSEG) ensures that segments with input root nodes are subjected to MAX(SEG) ≫ PRS, while those without are subjected to PRS ≫ MAX(SUBSEG).[14] Examples are in (20), where segments without root nodes are underlined.

---

[12] Briefly, procedures for learning OT grammars improve in performance if they opt to rank markedness higher than faithfulness when given a choice. Consequently the ranking in (18a) will be learned in preference to (18b); see §5.

[13] An early proposal that only faithfulness constraints be indexable (Benua 1997; Itô & Mester 1999; Fukazawa 1999) has proven untenable (Pater 2000; 2006; Flack 2007b; Flack 2007a; Inkelas & Zoll 2007; Gouskova 2007; Mahanta 2008; Jurgec 2010).

[14] Assuming undominated *FLOAT (Myers 1997), which prohibits surface subsegments, and low-ranked DEP(ROOT) (Zoll 2001).

(20)

|  |  | Max(Seg) | Prs | Max(Subseg) |
|---|---|---|---|---|
| a. | /malan<u>u</u>/<br>(ma la:n) ≻ (mu la:) nu |  | W | L |
| b. | /margu<u>ŋgu</u>/<br>(mar gu:ŋ) ≻ (mar guŋ) gu |  | W | L |
| c. | /mulari/<br>(mu la:) ri ≻ (mu la:r) | W | L | W |
| d. | /margu-nda/<br>(mar gu:n) da ≻ (mar gu:n) | W | L | W |

I wish to draw a distinction now between two conceivable kinds of phonological analysis. A CONCRETE phonological analysis represents exceptionality using regular phonological material, such as features, root nodes and prosodic units, or perhaps their absence. An ABSTRACT phonological analysis uses diacritic lexical indices, which I will term LEXICAL Φ-INDICES, on segments, much like the morphological analysis uses lexical M-indices on morphs. Some objections which have been raised to phonological analyses are specific to the concrete approach. These include doubts over whether sufficiently many concrete phonological contrasts would be available in languages with very many exceptional patterns (Gouskova 2012), and concerns over whether learners can choose between multiple, alternative concrete representations (Kiparsky 1973, Pater 2009). I will set these concrete-specific concerns aside for now, and instead assume an abstract phonological approach. I return to the concrete approach in §9, where I argue on independent grounds that it is poorly adapted to efficient learning.

Accordingly, I will use lexical Φ-indices $u$ and $n$ to index undergoer and non-undergoer segments respectively, and define Φ-indexed constraints, $C_\Phi$, in relative terms as in (21).

(21)   V($C_\Phi$) $=_{\text{def}}$ V(C) ∩ L$_\varphi$($\varphi \in \Phi$), *where:*
        Φ is the set of phonological elements co-indexed to $C_\Phi$.

Returning to the phonological account of Yidiny exceptionality, a constraint ranking Max-$n$ ≫ Prs ≫ Max, or Prs-$u$ ≫ Max ≫ Prs, will be sufficient for our purposes. Tableau (22) shows examples using the latter ranking; $u$-indexed segments are underlined.

(22)

|  |  | Prs-$u$ | Max | Prs |
|---|---|---|---|---|
| a. | (ma la:n) ≻ (mu la:) n<u>u</u> | W | L | W |
| b. | (mar gu:ŋ) ≻ (mar guŋ) g<u>u</u> | W | L |  |
| c. | (mu la:) ri ≻ (mu la:r) |  | W | L |
| d. | (mar gu:n) da ≻ (mar gu:n) |  | W | L |
| e. | (ma ɟin) (da ŋ<u>a:</u>l) na ≻ (ma ɟin) (da ŋa:l) |  | W | L |

A recent criticism of the phonological approach to exceptionality in OT is that it over-generates (Gouskova 2012). Adapting Gouskova's arguments to the facts of Yidiny: if we adopt the ranking Prs-*u* ≫ Max ≫ Prs, then it is no longer necessary to assign a high ranking to the morphologically-sensitive constraints Anc and Max-C/Rt, which penalize the deletion of morph-initial segments and root consonants. Rather, so long as all morph-initial segments and all root consonants lack a lexical *u*-index, then by virtue of the partial ranking Max ≫ Prs, they will resist deletion irrespective of the ranking of Anc and Max-C/Rt. By the same token however, if Anc and Max-C/Rt do receive a low ranking, then the analysis will fare poorly in the context of Richness of the Base (Prince & Smolensky 2004-1993]), since without high-ranked Anc and Max-C/Rt ensuring that morph-initial and root-consonant deletion is impossible, there is nothing to prevent seg-ments from deleting in those positions if they are *u*-indexed in the lexicon. For example, a root such as \**binarŋa* could undergo CV deletion; a suffix \*-*ni* could delete entirely; and \**mulari* could delete from the left, thereby failing to capture the generalization that the absence of such forms is not an accident of the lexicon, but a systematic property of the grammar. This is perhaps the most significant apparent flaw of the phonological approach: it fails to rule out unattested patterns. This is in contrast to the morphological approach, which does rule them out. Or at least, so it would seem. In §5 I show that the true situation can be otherwise, once learning is taken into account.

## 4.3 Alternatives

Before proceeding to learning, I mention two OT alternatives to the analysis of excep-tionality in Yidiny word-final deletion.

Co-phonological approaches handle exceptionality as a type of cyclicity effect (Orgun 1996, Kiparsky 2000, Inkelas & Zoll 2007, Bermúdez-Otero 2016). On each morphological cycle the result of a morphological operation is submitted to an appropriate phonological subgrammar, of which the language may possess many. Problematic for any cyclicity-based approach to exceptionality in Yidiny word-final deletion is that the Yidiny case is non-cyclic. Instead, undergoers are subject to deletion only if word-final. For example, in building both words in (23a,b) the first step would be to introduce the undergoer root *bigunu-* 'shield'. However at that point, the "deleting" subgrammar should only be applied if the root will end up word final, as in (23a) but not in (23b).

(23)  a.  'shield[ABS]'          /bigunu/           →   (bi gu:ŋ)
      b.  'shield-comit-ERG'     /bigunu-yi-ŋgu/    →   (bi gu) (nu yi:ŋ)
                                                        \*(bi gun) (yiŋ gu)

Selecting the correct subgrammar in (23) thus requires information about the next step in the derivation. Crucially though, it requires forewarning, not only of whether or not there is more morphology to come, but also of what the PHONOLOGICAL ramifications will be. This is because the relevant domain for word final-deletion in Yidiny is not the morphological word but the prosodic word. For example, in (24) the roots *gaɟula-* 'dirty' and *gumaɻi-* 'red' are followed by suffixes. Since the suffixes are monosyllabic, just one

prosodic word results and the roots are non-final in their prosodic word. In (25) however, the roots are followed by the disyllabic INCHOATIVE suffix *daga*, which commences a second prosodic word. As a consequence, the roots are final in their prosodic word and deletion is possible: the undergoer *gaɟula-* deletes while the non-undergoer *gumaɻi-* does not.

(24) a. 'dirty-CAUS-PST'
/gaɟula-ŋa-lɲu/    →    [(ga ɟu) (la ŋa:l)ₚwd]

b. 'red-CAUS-PST'
/gumaɻi-ŋa-lɲu/    →    [(gu ma) (ɻi ŋa:l)ₚwd]

(25) a. 'dirty-INCHO-PST'
/gaɟula-daga-ɲu/    →    [(ga ɟu:l)ₚwd] [(da ga:ɲ)ₚwd]

b. 'red-INCHO-PST'
/gumaɻi-daga-ɲu/    →    [(gu ma:) ɻi ₚwd] [(da ga:ɲ)ₚwd]

Any cyclic, look-ahead mechanism in Yidiny would therefore need to know how the word would be prosodically parsed on the NEXT cycle, before it can decide whether or not to apply the "deleting" subgrammar on the current cycle. The look-ahead mechanism would therefore require the power of a subgrammar itself, yet if the theory were augmented in this manner, then other core mechanisms such as scope, or "bracket erasure", effects (Inkelas & Zoll 2007) would be undermined. I conclude that co-phonology theory as it stands cannot analyse exceptionality in Yidiny word-final deletion.

Another approach would be to lexically list two allomorphs for all undergoer morphs in the language, and have the grammar select them either optimally (Mester 1994, Kager 1996, Mascaró 1996, and Tranel 1996a,b) or with some degree of stipulation (Bonet, Lloret & Mascaró 2007; Round 2013; Wolf 2015). On this approach, "deletion" is apparent only, due in reality to the selection between two input allomorphs, one of which contains only a subset of the segments in the other (for a proposal not unlike this for Yidiny, see Hayes 1997). An example is shown in (26), where the grammar optimally selects between two input allomorphs of the undergoer root *bigunu-* 'shield'.

(26)

| {/bigunu/, /bigun/} 'shield[ABS]' | ANC | MAX-C/RT | PRS | MAX |
|---|---|---|---|---|
| a. ☞ /bigun/ :: (bi gu:n) | | | | |
| b. /bigun/ :: (bi gu:) nu | | | *W | |
| c. /bigunu/ :: (bi gu:n) | | | | *W |
| d. /bigunu/ :: (bi gu:) nu | | | *W | |

Two objections can be raised. First, because the approach simply lists alternant pairs, it misrepresents their resemblances as accidents, rather than relating them systematically. Relatedly, in the context of Richness of the Base, the analysis would allow the apparent deletion of morph-initial and -medial segments as well as root consonants, by leaving them out of an underlying allomorph, in a pair such as {/bigunu/, /gunu/}. Ranking ANC and MAX-C/ROOT highly would not ameliorate the problem, as shown in (27).

(27)

| {/bigunu/, /gunu/} | ANC | MAX-C/ROOT | PRS | MAX |
|---|---|---|---|---|
| a. ☞ /gunu/ :: (gu nu) | | | | |
| b. /bigunu/ :: (bi gu:) nu | | | *W | |

Second, it is unclear how the analysis would prevent apparent deletion in word medial positions in the event that it is optimsing, as in (28), where the true output *buɟala-ŋa:-lna* violates PRS while the more optimal false winner \**buɟal-ŋa-lna* does not. The constraint CNTG will not prevent this occurring.

(28)

| †{/buɟala, buɟal/}-ŋa-lna/ 'finely ground-CAUSE-PURP' | CNTG | *CPLX | PRS | MAX |
|---|---|---|---|---|
| a. ☞ /buɟala-ŋa-lna/ :: (bu ɟa) (la ŋa:l) na | | | *L | |
| b. \* /buɟal-ŋa-lna/ :: (bu ɟal) (ŋal na) | | | | |

I conclude that neither the co-phonological approach nor the allomorph-selection approach offers a viable alternative for Yidiny word-final deletion.

# 5 Learning exceptionality

## 5.1 Biased Constraint Demotion

I turn now to consider how exceptionality is, or isn't, learned. After introducing Prince and Tesar's (2004) Biased Constraint Demotion (BCD) algorithm and adaptations of it for the learning of indexed constraints, I show that the learning of Yidiny word-final deletion does not proceed as one might expect from the discussion in §4. A solution is then offered in §6.

Prince and Tesar's BCD is a computationally efficient algorithm for the learning of OT grammars. It builds upon Tesar's earlier Recursive Constraint Demotion (RCD) algorithm (Tesar 1995, Tesar & Smolensky 2000), deterministically learning a grammar, conditional on the data, by ranking constraints in a series of steps, or recursions. At the first step, one or more constraints is assigned to the highest-ranked CONSTRAINT STRATUM in the grammar. A stratum is a set of constraints whose relative ranking against one another is indeterminate given the data, but whose ranking relative to constraints in other strata is significant. The act of assigning constraints to a stratum is termed INSTALLATION. At each subsequent step, one or more additional constraints are installed in the next-highest stratum, and so on, until all constraints are ranked. The determination of which constraint(s) are installed next is based on evidence from winner–loser pairs (WLPs). For each WLP, any constraint yet to be installed will favor the winner in the pair, the loser, or neither. The full table of WLPs and constraints yet to be installed is termed the SUPPORT. A fragment of a support is shown in (29). The relative order of constraints and WLPs in a support is inconsequential, though for ease of inspection I set

out markedness constraints to the left of a vertical double line, and faithfulness to the right.

(29)

| | | FtBin | Prs-*u* | Prs | *Cplx | Max | Cntg | Anc |
|---|---|---|---|---|---|---|---|---|
| a. | /margu-n<u>i</u>/ <br> (mar gu:n) > (mar gu:) ni | | W | W | | L | | |
| b. | /guygal-n<u>i</u>/ <br> (guy ga:l) ni > (guy ga:l) | | L | L | | W | | W |
| c. | /guygal-n<u>i</u>/ <br> (guy ga:l) ni > (guy ga:ln) | | L | L | W | W | | |
| d. | /guygal-n<u>i</u>/ <br> (guy ga:l) ni > (guy ga:l ni) | W | L | L | | | | |
| e. | /bulmba/ <br> (bulm ba) > (bul ba) | | | L | L | W | W | |

In the original RCD algorithm, the sole criterion for installing a constraint was that it favor no losers. This is true of the constraints FtBin, Cntg and Anc in (29). When a constraint, C, is installed, all of the WLPs for which C favors the winner are removed from the support, since the constraint ranking has now accounted for them. In the RCD, all constraints meeting this criterion at any recursion are installed, and the result at the end of all recursions is a correct grammar for the data. Nevertheless, the grammars inferred by the RCD are not optimal (Prince & Tesar 2004). The suboptimality relates to the subset problem (Baker 1979; Angluin 1980), a general problem in algorithmic learning from positive evidence, namely that the system which results from learning will correctly assess as grammatical all attested items, but will fail to rule out certain systematically unattested items. This in turn relates to the notion of restrictiveness: a learning algorithm ought ideally to learn the most restrictive grammar consistent with the data. The RCD does not do this. In practice, meeting this desideratum is challenging for an efficient algorithm. However Prince & Tesar (2004) demonstrate that good headway can be made by enhancing the RCD with a small set of biases, hence the name Biased Constraint Demotion, or BCD. The BCD differs from the RCD in two main respects. The first is the principle of faithfulness delay. According to this, at every recursion faithfulness constraints are not installed, even when they favor no losers, unless there are no other installable constraints. In (29) for example, the BCD would install the markedness constraint FtBin but not the faithfulness constraints Cntg and Anc. If we do this, and remove from (29) all the WLPs for which FtBin favors the winner, namely (29d), and remove FtBin, we have (30), in which only faithfulness constraints, Cntg and Anc favor no losers; under these conditions, faithfulness delay would permit the installation of Cntg and Anc.

(30)

| | | Prs-*u* | Prs | *Cplx | Max | Cntg | Anc |
|---|---|---|---|---|---|---|---|
| a. | /margu-n<u>i</u>/ (mar gu:n) > (mar gu:) ni | W | W | | L | | |
| b. | /guygal-n<u>i</u>/ (guy ga:l) ni > (guy ga:l) | L | L | | W | | W |
| c. | /guygal-n<u>i</u>/ (guy ga:l) ni > (guy ga:ln) | L | L | W | W | | |
| e. | /bulmba/ (bulm ba) > (bul ba) | | L | L | W | W | |

However, there is a second principle to consider also. A principle of "freeing up markedness" states that when there is a choice between installing several faithfulness constraints, the algorithm should install the smallest subset possible, whose installment would cause a markedness constraint to become installable in the next recursion. For example, in (30), installing Cntg would remove WLP (30e), thereby freeing up the markedness constraint *Cplx at the next recursion; no comparable gain would flow from installing Anc. On those grounds, from (30) the BCD would install Cntg.

## 5.2 A support for learning Yidiny exceptionality

I now consider several learning scenarios for Yidiny exceptionality. Each begins directly after the installation of undominated constraints. Table 6 contains a set of WLPs that is representive of all combinations of roots and suffixes which are relevant to the grammar of word-final deletion: it is not the complete support, but it represents the complete support well. Segments which can delete are underlined. To economize on space below, WLPs will be referred to by the letters in the first column of Table 6.

## 5.3 Learning the phonological account (preliminary version)

We begin with the learning of the phonological account of Yidiny exceptionality described previously in §4.2. For the moment, I assume that input segments are already lexically Φ-indexed as *u* or *n*. We begin after undominated constraints have been installed, with a support as in (31).

Table 6: Support for learning Yidiny exceptionality.

| | | | |
|---|---|---|---|
| a. | /margu-ni̲/ | (mar gu:n) | >(mar gu:) ni |
| b. | /guygal-ni̲/ | (guy ga:l) ni | >(guy ga:l) |
| c. | /guygal-ni̲/ | (guy ga:l) ni | >(guy ga:ln) |
| d. | /margu-ŋgu̲/ | (mar gu:ŋ) | >(mar gu:ŋ) gu |
| e. | /margu-ŋgu̲/ | (mar gu:ŋ) | >(mar gu:ŋg) |
| f. | /bigunu̲-yi-ŋgu̲/ | (bi gu) (nu yi:ŋ) | >(bi gun) (yiŋ gu) |
| g. | /wawa-lṉu/ | (wa wa:l) | >(wa wa:l) ɲu |
| h. | /gali-ŋa̲/ | (ga li:ŋ) | >(ga li:) ŋa |
| i. | /gaɟara̲/ | (ga ɟa:r) | >(ga ɟa:) ra |
| k. | /margu-nda/ | (mar gu:n) da | >(mar gu:n) |
| l. | /wawa-lna/ | (wa wa:l) na | >(wa wa:l) |
| m. | /gali-na/ | (ga li:) na | >(ga li:n) |
| n. | /guɟara/ | (gu ɟa:) ra | >(gu ɟa:r) |
| o. | /maɟinda-ŋa̲-lna/ | (ma ɟin) (da ŋa:l) na | >(ma ɟin) (da ŋa:l) |
| p. | /bulmba/ | (bulm ba) | >(bul ba) |

(31)

| | Prs-*u* | Prs | *Cplx | Max | Max-C | Cntg | Anc |
|---|---|---|---|---|---|---|---|
| a, h, i. | W | W | | L | | | |
| b. | L | L | | W | W | | W |
| c. | L | L | W | W | | | |
| d, g. | W | | | L | L | | |
| e. | | | W | L | L | | |
| f. | | | | L | L | W | |
| j, k, l, o. | | L | | W | W | | |
| m, n. | | L | | W | | | |
| p. | | | L | W | | W | |

Support (31) does not contain any markedness constraints that favor no losers. Two faithfulness constraints favor no losers: Cntg, which would free up *Cplx if installed, and Anc, which would not free up any markedness constraints. Consequently, Cntg is installed next, removing WLPs (f) and (p) from the support. After that, the newly freed-up *Cplx is installed, removing WLPs (c) and (e), and leaving (32).

(32)

|  | Prs-*u* | Prs | Max | Max-C | Anc |
|---|---|---|---|---|---|
| a, h, i. | W | W | L | | |
| b. | L | L | W | W | W |
| d, g. | W | W | L | L | |
| j, k, l, o. | | L | W | W | |
| m, n. | | L | W | | |

In 32 only Anc favors no losers, and so is installed. This removes (b), freeing up Prs-*u*, which is installed next, removing (a,h,i) and (d.,g), leaving (33). From (33), Max will be installed since it frees up Prs. This leaves Prs and Max-C, which according to faithfulness delay, will be ranked last as Prs ≫ Max-C, as in (34).

(33)

|  | Prs | Max | Max-C |
|---|---|---|---|
| j, k, l, o. | L | W | W |
| m, n. | L | W | |

(34)   Cntg ≫ *Cplx ≫ Anc ≫ Prs-*u* ≫ Max ≫ Prs ≫ Max-C

Some comments are in order. First, the BCD algorithm has learned the key constraint ranking Prs-*u* ≫ Max ≫ Prs responsible for the core of Yidiny exceptionality. Secondly however, it has also ranked Anc ≫ Prs-*u*, in which case the learned grammar expressly prohibits morph-initial deletion. Indeed, had Max-C/rt been included in (31), it would also have been ranked highly since it only ever favors winners, meaning the grammar would also expressly prohibit CV deletion in roots (the reasons for my excluding Max-C/rt are clarified in §6). This means that the algorithm is learning precisely the rankings required to prevent the phonological solution from overgenerating, thereby voiding the major criticism of the phonological approach which was introduced in §4.2. This is perhaps surprising, so why is the ranking learned? It is learned because the BCD algorithm attempts to construct a restrictive grammar. The typical assumption, that grammars implementing a phonological approach would not assign redundant, high rankings to constraints like Anc, is predicated on an implicit assumption that the learner would be seeking a permissive grammar; doing so leads to overgeneration. However no successful learner would adopt that assumption, because successful learning in general requires a restrictive approach. For the theory of exceptionality, this is significant. It means the result obtained here, in which a phonological approach to exceptionality has been learned without overgeneration, is not dependent on some minor detail of the BCD, or the constraints used, or even OT. Rather, it follows from a general principle of learning. Consequently, the adoption of realistic assumptions about learning narrows the performance gap between the phonological and morphological approaches. I will examine the phonological approach further in §7.3.

## 5.4 Learning indexed constraints and the morphological analysis

We consider next the learning of the morphological approach. The support begins, after installation of undominated constraints, as (35). These are the same constraints and WLPs as in the previous section, but without Prs-*u*. The support begins with no lexically indexed constraints; how they are learned is considered shortly. I also do not include Max-C/rt in the support. Max-C/rt is essentially a variant of Max-C, indexed to all root morphs. This is the kind of constraint we might reasonably expect the morphological approach to learn.

(35)

|            | Prs | *Cplx | Max | Max-C | Cntg | Anc |
|------------|-----|-------|-----|-------|------|-----|
| a, h, i.   | W   |       | L   |       |      |     |
| b.         | L   |       | W   | W     |      | W   |
| c.         | L   | W     | W   |       |      |     |
| d, g.      | W   |       | L   | L     |      |     |
| e.         |     | W     | L   | L     |      |     |
| f.         |     |       | L   | L     | W    |     |
| j, k, l, o.| L   |       | W   | W     |      |     |
| m, n.      | L   |       | W   |       |      |     |
| p.         |     | L     | W   | W     | W    |     |

Turning now to the BCD algorithm, neither of the markedness constraints in support (35) favors no losers. Cntg does, and would free up *Cplx. Anc also does, but would not free up any markedness constraints. Accordingly, Cntg is installed next, removing WLPs (f) and (p) are from the support, and *Cplx after that, removing (c) and (e), leaving (36).

(36)

|            | Prs | Max | Max-C | Ans |
|------------|-----|-----|-------|-----|
| a, h, i.   | W   | L   |       |     |
| b.         | L   | W   | W     | W   |
| d, g.      | W   | L   | L     |     |
| j, k, l, o.| L   | W   | W     |     |
| m, n.      | L   | W   |       |     |

Anc is installed next, removing WLP (b), which leaves (37), a support in which there is no constraint which favors no losers.

(37)

|           | Prs | Max | Max-C |
|-----------|-----|-----|-------|
| a, h, i.  | W   | L   |       |
| d, g.     | W   | L   | L     |
| j, k, l, o. | L | W   | W     |
| m, n.     | L   | W   |       |

Supports in this state are said to have reached INCONSISTENCY. An inconsistency, however, is not a failure.

Inconsistencies indicate that the combination of data and assumptions currently under consideration have not led to a working grammar. Accordingly (assuming the data is correct), a revision of the assumptions is warranted. Suppose, in this case, that a revision could be made which leaves intact all previously installed constraints and their rankings, and the validity of all previously accounted-for WLPs, that is, a revision that would change only what is in the support. Suppose also that as a result of this revision the support came to contain a constraint that favors no losers. Such a revision would resolve the inconsistency. The BCD could restart and, one hopes, lead to a working grammar. Revisions that meet these criteria can be considered a type of learning. One such revision is to add a new, lexically M-indexed constraint to Con.

Pater (2009) describes a method for learning M-indexed constraints and assigning co-indices to morphs, which takes a BCD inconsistency as its starting point. Coetzee (2009) extends this to Output-Output constraints, which I will not consider here. Becker's modifications (Becker 2009; Becker, Ketrez & Nevins 2011) are addressed in §8.

Central to Pater's method is the operation of CONSTRAINT CLONING, a process I describe informally here and return to in detail in §8. Within the stalled support, a constraint C is sought which, if it were indexed to some set *M* of morphs, would (i) favor at least one winner[15] and (ii) favor no losers. Assuming such a constraint C can be identified, it is then cloned, which is to say, a lexically M-indexed version of it, $C_M$, is added to the support. Because $C_M$ favors no losers, it is installed next. For example, support (38) is the same as (37) but now displays information about which morphs are involved. I have annotated relevant undergoers as *U* and non-undergoers as *N*.

According to the criteria for cloning, all three of Prs, Max and MaxC are candidates for cloning (indexed to sets *U*, *N* and *N* respectively). I assume that owing to faithfulness delay, markedness constraints are cloned in preference to faithfulness when both are available, in which case Prs will be cloned. In (39) the cloned, lexically M-indexed constraint $Prs_U$ is added to the support. Installing it removes WLPs (a,d,g,h,i) which frees up Max, whose installation is followed by Prs and Max-C. The resulting ranking is (40), which requires comment.

---

[15] The new constraint needs to favor at least one winner to have any chance of freeing up another constraint once it is installed.

(38)

|  |  | Prs | Max | Max-C |
|---|---|---|---|---|
| a. | /margu-ni$_U$/ | W | L |  |
| d. | /margu-ni$_U$/ | W | L | L |
| g. | /margu-ŋgu$_U$/ | W | L | L |
| h. | /gali-ŋa$_U$/ | W | L |  |
| i. | /gaɟara$_U$/ | W | L |  |
| j. | /binarŋa/ | L | W | W |
| k. | /margu-nda$_N$/ | L | W | W |
| l. | /wawa-lna$_N$/ | L | W | W |
| m. | /gali-na/ | L | W |  |
| n. | /guɟara$_N$/ | L | W |  |
| o. | /maɟinda-ŋa-lna$_N$/ | L | W | W |

(39)

|  | Prs-$_U$ | Prs | Max | Max-C |
|---|---|---|---|---|
| a, h, i. | W | W | L |  |
| d, g. | W | W | L | L |
| j, k, l, o. |  | L | W | W |
| m, n. |  | L | W |  |

(40)  Cntg ≫ *Cplx ≫ Anc ≫ Prs$_U$ ≫ Max ≫ Prs ≫ Max-C

The algorithm has successfully learned the key constraint ranking Prs$_U$ ≫ Max ≫ Prs. However, it did not create an indexed version of Max-C for roots, and thus has not learned to expressly prohibit CV deletion in roots. To be sure, no individual roots ending in CCV or VC′V (where C′ would be an illicit coda) will have been co-indexed to Prs$_U$ during the cloning operation (see §8 for details) and so none of those roots will be subject to CV deletion, however the ranking in (40) predicts that if the lexicon did contain a root such as *binarŋa$_U$, then that root and any like it would undergo CV deletion. This is overgeneration of the same kind which was believed to beset phonological accounts. Thus, while §5.3 showed that grammars learned for the phonological account may suffer less than expected from overgeneration once learning is taken into consideration, §5.4 shows that grammars for the morphological account may suffer from overgeneration more than expected. In the next section, I propose a solution.

# 6 Morphological analytic bias: the Morphological Coherence Principle

In §5.4 the grammar which was learned for a morphological analysis of Yidiny exceptionality suffers from a manifestation of the subset problem. Although the algorithm

correctly handled all attested data, it did not learn the more restrictive generalization which applies also to unattested data, that roots in Yidiny do not undergo consonant deletion. The problem arises because the cloning procedure assesses morphology on a morph-by-morph basis only, whereas the true generalization in Yidiny applies to a class of morphs, in this instance, to roots. The remedy to be pursued here has two parts. It adds a new kind of constraint cloning, which indexes a constraint not to an idiosyncratic lexical list of morphs, but to a general class. It then biases constraint cloning so that class-indexed (or K-INDEXED) cloning is preferred over lexically indexed cloning. Effectively, this introduces an analytic bias (Moreton 2008) from morphology to phonological learning at BCD inconsistencies.

Now, supposing that the algorithm is seeking a constraint that it will clone and K-index to some non-idiosyncratic class of morphs, which classes should be available for the learner to consider? Important here is the fact that human phonological learning will need to proceed in parallel with, and interleaved with, morphological learning (Tesar 2007, Merchant 2008: 6). Accordingly, I assume the learner has access both to universally-defined classes such as "root", and those classes which have been morphologically learned, such as ERGATIVE CASE. The biasing principle, which I term the Morphological Coherence Principle is stated in (41), where criterion 2 provides an additional bias towards maximal restrictiveness.

(41) The Morphological Coherence Principle:
    1. At a BCD inconsistency, attempt to create a K-indexed constraint, co-indexed to some universal or learned morphological class *K*, before attempting to create a lexically-indexed constraint.
    2. If multiple constraints are eligible for K-indexation, select the one whose co-indexed class is most general.

The MCP has some desirable theoretical properties. If the universal state of CON at the commencement of learning is CON$_{init}$, then the MCP obviates the need for CON$_{init}$ to contain any constraints that are relativized to universal or learned morphological classes, since such constraints will be learned on demand, if and only if needed. In effect, this reduces the size of CON$_{init}$ without any change in the explanatory capacity of the theory. And, since it allows the grammar to BUILD constraints for language-specific morphological classes it makes those constraints available to the learner without problematically assuming them universal (Russell 1995, Hammond 2000, see also Smith 2004, Flack 2007b). The MCP operationalizes, in a specific manner, the kind of insight into linguistic theory that Anderson (2008) argues ought to follow from an improved understanding of the learning device.

Let us now return to Yidiny exceptionality, equipped with the MCP. Learning begins and proceeds as in §5.4 until the inconsistency in (38), at which point a constraint is sought for cloning. The MCP states that if possible, a constraint should be cloned and K-indexed. In (38) MAX-C would favor no losers if it were K-indexed to the entire class of roots, so it is cloned and accordingly K-indexed. This is the functional equivalent of adding MAX-C/RT to CON, and the reason why in §5 I did not include MAX-C/RT in the

support at the outset. Adding Max-C/rt to the support results in (42). From (42), Max-C/rt is installed and WLP (j) is removed, whereupon we return to inconsistency, in (51). As in §5.4, the process from that point results in the cloning of Prs and the installation of Prs$_U$, then Max, Prs and Max-C, yielding the desired constraint ranking (43).

(42)

|  | Prs | Max | Max-C/rt | Max-C |
|---|---|---|---|---|
| a, h, i. | W | L |  |  |
| d, g. | W | L |  | L |
| j. | L | W | W | W |
| k, l, o. | L | W |  | W |
| m, n. | L | W |  |  |

(43)

|  | Prs | Max | Max-C |
|---|---|---|---|
| a, h, i. | W | L |  |
| d, g. | W | L | L |
| k, l, o. | L | W | W |
| m, n. | L | W |  |

(44)  Cntg ≫ *Cplx ≫ Anc ≫ Max-C/rt ≫ Prs$_U$ ≫ Max ≫ Prs ≫ Max-C

To summarize, results from §5.3 suggested that, provided a learner is seeking a restrictive grammar, the phonological approach to exceptionality may not suffer from overgeneration. This contradicts recent arguments, which on examination appear to adopt the implausible assumption that a learner would be seeking a permissive grammar. That being said, I have not yet clarified how the learner would arrive at the requisite Φ-indices required by the phonological approach. That will be discussed in §7.3. Meanwhile, §5.4 revealed that without further refinement, the BCD is prone to learning grammars that overgenerate even in a morphological approach to exceptionality, due to an overly atomistic method of morphological generalization. This was remedied in §6 by the Morphological Coherence Principle (41), which solves the learning problem and simplifies Con$_{init}$.

# 7 The theoretical status of lexical indices

In §7 I set Yidiny to one side and consider some matters of theory.

## 7.1 Lexical M-indices

Lexical M-indices are representations which are visible to the phonology, but they are not phonological elements *per se*. In OT, Gen cannot alter M-indices. It cannot add or remove them, or displace them from one morph to another. There is therefore no need for mechanisms such as M-index "faithfulness", rather it is simply assumed that

the lexical affiliation of a morph *m* with an M-index *M* is identical in the input and output. This set of properties is shared with other kinds of lexical affiliation, such as the affiliation of a phonological element with its morph, and is termed Consistency of Exponence (McCarthy & Prince 1993b, Van Oostendorp 2007).

Taking a historical view, M-indices closely resemble the RULE FEATURES and ALPHABET FEATURES of early generative phonology (GP) (Chomsky & Halle 1968, Lakoff 1970, Coats 1970, Zonneveld 1978, *inter alia*). Both sets of formalisms fulfill the function of determining for cases of exceptionality whether a morph *m* participates in certain phonological patterns or not, by ensuring that *m* is visible or not visible as required, to OT's constraints or GP's phonological rules. Diacritic features were investigated extensively in GP. It was argued that the theory should not allow the phonology to manipulate diacritic features (Kiparsky 1973, Zonneveld 1978). The same applies to M-indices in OT. It was argued that not all idiosyncrasies in the phonology can be analysed satisfactorily in terms of rule exception features, and that there is an additional role for cyclicity (Chomsky & Halle 1968, Kiparsky 1982b) and the same has been recognized for M-indices (Pater 2009). In GP, it was also assumed that the diacritic features of morph *m* were distributed across, and directly characterized, each of the phonological elements (namely, segments) in *m*. We might ask whether this is also true of M-indices in OT. Suppose that it is, so that the *M*-indices of a morph *m* directly characterize each phonological element *φ* that is lexically affiliated with *m* (that is all *φ* which are EXPONENTS of *m*). In that case, the relative definition of an M-indexed constraint (25), repeated here as (45), can be revised and simplified as (46).

(45)  $V(C_M)$ $=_{\mathrm{def}}$ $V(C)$ $\cap$ $L_\varphi(m \in M$ & $\mathrm{Exp}(\varphi, m))$, *where*:
   $M$ is the set of morphs co-indexed to $C_M$.
   $\mathrm{Exp}(\varphi, m)$ states that element *φ* is an exponent of morph *m*

(46)  $V(C_M)$ $=_{\mathrm{def}}$ $V(C)$ $\cap$ $L_\varphi(\varphi \in \Phi_M)$, *where*:
   $\Phi_M$ is the set of phonological elements co-indexed to $C_M$.

It will be recalled that the relative definition of a constraint $C_M$ is expressed as the set intersection between the loci of variation of the unindexed constraint C, written $V(C)$, and the set of loci, $L_\varphi(D(\varphi))$ which contain some criterial type of phonological element *φ*, described by predicate $D(\varphi)$. Importantly, this means that M-indexed constraints are defined DIRECTLY in terms of phonological elements, *φ*, and only indirectly in terms of morphs *m*. The indirectness shows up in the complexity of $D(\varphi)$ in (45), which links morphs to their exponent *φ* elements via the function $\mathrm{Exp}(\varphi, m)$. This is in contrast with (46), where the assumption is that all *φ* elements are directly characterized by the M-index borne by their affiliated morph. The constraint definition no longer refers to the morph itself, and so the predicate $D(\varphi)$ is simpler.

At risk of laboring the point, the phonology itself assesses violations of M-indexed constraints directly in terms of *φ* elements, not morphs. While it is possible to refer to the morphs in the definitions of M-indexed constraints as in (17/45), it is not necessary. Nor is it possible to refer only to the morphs and not to the *φ* elements, since the loci of violation of these constraints are defined inherently at a sub-morphological, phonological level.

## 7.2 Lexical Φ-indices

Let us now consider the nature of lexical Φ-indices of the type I invoked in §4.2 and §5.3. My proposal is that these are exactly like M-indices: they are non-phonological indices of lexical affiliation, visible to, but not manipulable by, the phonology and used for making particular phonological elements visible or not, as required, to OT's constraints in order to provide a coherent account of exceptionality. The only distinction between Φ-indices and M-indices lies in the supplementary assumption attached to M-indices, in (47).

(47)   The M-index assumption:
       A lexical index which characterizes phonological element $\varphi_i$ will also
       characterize all other phonological elements $\varphi_j$ affiliated with the same morph *m*.

Φ-indices are not subject to this redundancy; they are affiliated with only those $\varphi$ elements for which the affiliation makes any difference to the analysis of language. As I will show in §8, that makes Φ-indices somewhat simpler to learn, since they correspond more directly to the evidence in the data.

The reader may also have noticed that the definition of a Φ-indexed constraint in (21) is almost exactly like the simplified definition of an M-indexed constraint in (46). This reflects the fact that for the operation of the phonology, it is $\varphi$ elements, and the indexation of specific $\varphi$ elements, that matter. Whether or not one chooses to adopt supplementary assumption (47) in fact has no material consequence for the evaluation of an individual indexed constraint. The question of whether there are other consequences, and whether they are desirable, is taken up in §9.

## 7.3 Learning lexical Φ-indices

Given the proposal above, the learning of Φ-indices is quite parallel to the learning of M-indices. I assume that the MCP still applies, so that class-based exceptionality and K-indexed constraints continue to be learned with priority over idiosyncratic exceptionality, even though the latter will now be accounted for by Φ-indexed constraints, not M-indexed. This is a coherent assumption to make. The MCP is concerned with the learning of class-based generalizations, whereas Φ- and M-indexed constraints are alternative devices for learning idiosyncrasies. Accordingly, in a stalled support once there are no K-indexed constraints available for cloning, the algorithm seeks a constraint C which, were it indexed to some set Φ of phonological elements, would (i) favor at least one winner and (ii) favor no losers. All else proceeds as for M-indexed constraints. In the learning of Yidiny word-final deletion, the process begins as in §6, leading to a first inconsistency resolved by the addition of Max-C/rt to Con, and proceeding from there to the second inconsistency (43), repeated here in part and in more detail as (48).

(48)

|  |  | Prs | Max | Max-C |
|---|---|---|---|---|
| a. | /margu-ni̱/<br>(mar gu:n) > (mar gu:) ni | W | L |  |
| d. | /margu-ŋgu̱/<br>(mar gu:ŋ) > (mar gu:ŋ) gu | W | L | L |
| h. | /gali-ŋa̱/<br>(ga li:ŋ) > (ga li:) ŋa | W | L |  |
| i. | /gaɟara̱/<br>(ga ɟa:r) > (ga ɟa:) ra | W | L |  |
| k. | /margu-nda/<br>(mar gu:n) da > (mar gu:n) | L | W | W |
| m. | /gali-na/<br>(ga li:) na > (ga li:n) | L | W |  |
| n. | /guɟara/<br>(gu ɟa:) ra > (gu ɟa:r) | L | W |  |
| o. | /maɟinda-ŋa̱-lna/<br>(ma ɟin) (da ŋa:l) na > (ma ɟin) (da ŋa:l) | L | W | W |

In (48), no K-indexed constraint is available for cloning.[16] Turning to potential Φ-indexed constraints, we see that the constraint Prs would, if it were co-indexed to all underlined phonological elements, favor at least one winner and favor no losers, and so it is cloned and co-indexed resulting in (49).

(49)

|  |  | Prs | Prs-*u* | Max | Max-C |
|---|---|---|---|---|---|
| a. | /margu-ni̱/ | W | W | L |  |
| d. | /margu-ŋgu̱/ | W | W | L | L |
| h. | /gali-ŋa̱/ | W | W | L |  |
| i. | /gaɟara̱/ | W | W | L |  |
| k. | /margu-nda/ | L |  | W | W |
| m. | /gali-na/ | L |  | W |  |
| n. | /guɟara/ | L |  | W |  |
| o. | /maɟinda-ŋa̱-lna/ | L |  | W | W |

(50)   Cntg ≫ *Cplx ≫ Anc ≫ Max-C/rt ≫ Prs-*u* ≫ Max ≫ Prs

---

[16] Actually this is not strictly true. All past suffixes for example are undergoers, in which case the MCP would generate and rank Prs$_{\text{PST}}$. Notwithstanding this, the essential argument remains, since other morphological classes exist, such as ergative and "root", that are not uniformly (non)undergoers, and still need to be handled by lexically-indexed, not K-indexed, constraints. This minor correction applies equally to the learning process in §6.

From there the algorithm proceeds in the now-familiar fashion, resulting in grammar (50). With its high-ranking Max-C/rt and Anc, (50) does not overgenerate. Moreover, given the argument in §7.2, that for Eval there is no detectable difference between M-indexed and Φ-indexed constraints, we can see that grammar (50) is in all material aspects identical to grammar (44) learned in §6.

# 8 Constraint cloning

## 8.1 Assessing eligibility for cloning

It is necessary now to examine more precisely the processes by which constraints are deemed eligible for cloning (§8.1), by which a viable set of co-indexed elements is identified (§8.2), and by which a selection is made between multiple eligible constraints (§8.3).

Earlier, I introduced criteria by virtue of which a constraint becomes eligible for cloning. These are restated in (51) in a generalized from, so that the set *S* is: a coherent class of morphs for K-indexing; an idiosyncratic set of morphs for M-indexing; or an idiosyncratic set of lexical phonological elements for Φ-indexing.

(51)     A constraint should be sought for cloning which, if it were indexed to set *S*, would (i) favor at least one winner, and (ii) favor no losers.

Criterion (51ii) ensures that once the cloned constraint is added to the support, it can be installed; (51i) ensures that its installation will remove at least one WLP from the support, and thereby have some hope of freeing up other constraints. The formulation in (51) improves upon Pater's (2009: 144) criterion, which is to seek a constraint that favors no losers "for all instances" of some morph.[17] To see why Pater's criterion fails, consider WLPs (h,l,o) from the stalled support (38), reproduced in part and in detail in (52). For the purposes of discussion, I assume we are attempting to learn an M-indexed constraint, though the argument generalizes to other kinds.

(52)

|  |  | Prs | Max | Max-C |
|---|---|---|---|---|
| h. | /gali-ŋa/ 'go-com[imp]' <br> (ga li:ŋ) > (ga li:) ŋa | W |  |  |
| l. | /wawa-lna/ 'see-purp' <br> (wa wa:l) na > (wa wa:l) | L | W | W |
| o. | /maɟinda-ŋa-lna/ 'walk up-com-purp' <br> (ma ɟin) (da ŋa:l) na > (ma ɟin) (da ŋa:l) | L | W | W |

In (52), WLPs (h) and (o) both contain the suffix -*ŋa*, a regular undergoer which our procedure ought to co-index to the M-indexed constraint Prs$_\text{U}$. In WLP (h) word-final *ŋa* is subject to deletion, and Prs favors the winner. In WLP (o) non-final *ŋa* is parsed

---

[17] Pater's phrase "favors only winners" is equivalent to my "favors no losers".

into a foot and escapes deletion. Nevertheless, for WLP (o) Prs favors the loser. This has nothing to do with *ŋa*, but is due to the non-deletion of the unparsed, word-final non-undergoer *-lna*. Pater's co-indexing criterion asks whether Prs favors no losers "for all instances" of *-ŋa* in the support. The answer is "no", because (o) contains an instance of *-ŋa* and Prs favors the loser for (o). This is the wrong result; the suffix *-ŋa* ought to get co-indexed to Prs$_U$. It comes about because Pater's criterion does not discriminate between morphs that contribute to violations and those which are present in the word, but do not contribute. The criteria in (51) avoid this problem because they refer directly to how the co-indexed constraint would perform, were it created. The next two sections detail how to operationalize them.

## 8.2 Specifying co-indexed sets

The question considered here is which set *S* ought to be co-indexed to a given constraint C if we wish to clone C? The answer varies depending on which kind of indexed constraint we are constructing. One possible answer is that no such set exists, and C cannot be cloned. Seen from that angle, the question here is also: is C eligible for cloning?

K-indexed constraints can be co-indexed only to the morphological classes $K_1, K_2 \ldots K_n$ in the language (§6). In (41) I suggested that the preferred class for co-indexation is the MOST GENERAL one. Thus, to efficiently assess if constraint C is eligible for cloning and K-indexing, the learner should proceed stepwise through the available classes, ordered by decreasing generality. The process is one of trial and error. At each step, the constraint $C_K$ is built and applied to all WLPs in the support. If $C_K$ meets criteria (51) then it is successful; the process halts and $C_K$ is used, otherwise the trial and error continues. If by the end, no successful constraint $C_{K1} \ldots C_{Kn}$ is found, then C is ineligible for cloning.

For M-indexed and Φ-indexed constraints, the desired set *S* can be identified by focusing attention on loci of violation. Suppose we are considering constraint C for cloning. For any WLP, *p*, its loci of violation of constraint C fall into three classes: the class w(*p*), responsible for violations of C that favor the winner (i.e., the locus occurs in the loser only), class L(*p*) which favor the loser (locus occurs in the winner only) and class N(*p*) which favor neither (occurs in both). Next define $\Phi_{w(p)}$ as the set of phonological elements *φ* contained in any of the loci in w(*p*), and $\Phi_{L(p)}$ as the set of *φ* elements contained in any of the loci in L(*p*). Finally, define $\Phi_W$ as the union of $\Phi_{w(p)}$ for all WLPs, $p_1, p_2 \ldots p_n$, in the support, and $\Phi_L$ as the union of all $\Phi_{L(p)}$ in the support. Now, consider the set ($\Phi_W - \Phi_L$), the set difference between $\Phi_W$ and $\Phi_L$. This is the set of all *φ* elements which both (i) appear in at least one locus that in at least one WLP causes C to favor a winner, and (ii) never appear in a locus that causes C to favor a loser. For a Φ-indexed constraint this is an optimal set *S*. If for a given constraint C, ($\Phi_W - \Phi_L$) is the null set, then we may conclude that C is ineligible for cloning.[18]

---

[18] To be precise, if ($\Phi_W - \Phi_L$) is the null set then it is possible that there still exists some additional, viable set *S* which contains fortuitous elements $\varphi_i$ which are elements of both $\Phi_W$ and $\Phi_L$ such that in EVERY WLP *p* in which $\varphi_i$ is contained in some number *n* of the loci w(*p*) there are at least *n* offsetting loci in L(*p*) which contain other elements $\varphi_j$ which are also in *S*. Identifying these fortuitous elements $\varphi_i$, or even determining if any exist, would very likely be prohibitively expensive computationally.

To find the equivalent for an M-indexed constraint, it is necessary to extrapolate from $\Phi_W$ and $\Phi_L$ to morphs: set $S$ will be the set $(M_W - M_L)$ where $M_W$ is the set of all morphs $m_w$, such that any of $m_w$'s phonological exponents is an element of $\Phi_W$; and $M_L$ is the set of all morphs $m_l$, such that any of $m_l$'s phonological exponents is an element of $\Phi_L$. Note that $M_W$ and $M_L$ can be calculated only after the calculation of $\Phi_W$ and $\Phi_L$ is performed.

In §7.1 I considered what is involved computationally in assessing violations of $\Phi$- and M-indexed constraints, and argued that the calculations for both are essentially concerned with $\varphi$ elements, not morphs. Here we see that the same is true when learning the co-indexed set. As in §7.1, one can bring morphs into the picture, to be sure, but in both cases doing so requires additional computational effort, for no effective difference in how the grammar will work. In §9 I will argue the theory to be preferred is one which admits lexically $\Phi$-indexed constraints, but not M-indexed.

## 8.3 Selecting among eligible constraints

Suppose there are multiple lexically-indexed constraints which are eligible for cloning; which do we choose? The principles of faithfulness delay and freeing-up of markedness constraints will eliminate some options (§5.1). Beyond that, I suggest the learner chooses the constraint which favors the most winners, and whose installment would therefore remove the greatest number of WLPs. A desirable consequence will be a bias toward restrictiveness. For example, suppose Max is eligible. If so, then so too is Max-C, Max-V, Max-p, etc. This "maximize-winners" criterion would select Max, and increase the restrictiveness of the grammar, relative to the other options.

Interestingly, Becker (2009) proposes a MINIMIZE-winners criterion, whose effect is to generate many, very specific cloned constraints, each indexed to highly specific subclasses in the lexicon. The aim is to account for a particular phenomenon, which I describe here. I argue that other accounts are possible, and that Becker's solution has undesirable consequences.

When language learners assign novel words to existing grammatical categories, they do so on the basis of statistical correlations that exist in the lexicon, for example between category membership and aspects of the members' phonological forms (Poplack, Pousada & Sankoff 1982; Albright 2002). One such task is to assign a word as exceptional or non-exceptional, given evidence which underdetermines that choice. The key question here is, what existing statistical knowledge do speakers use, and what do they ignore? In Turkish, speakers appear to ignore correlations between the (non)alternation of a stop's laryngeal features and the quality of its neighboring vowel. It is proposed (Becker 2009; Becker, Ketrez & Nevins 2011) that this is because speakers do not access lexical statistics *per se*, rather they attend to the statistics of constraint indexation. Importantly, Con lacks constraints such as *[+HIGH]tV which refer to a stop and the quality of its vocalic neighbor. Consequently, no such constraint can be indexed, making such correlations invisible and hence irrelevant to a speaker when she assigns a novel word to a (non)exceptional lexical category. Assuming this is the case, then in order for fine-grained knowledge to be available to speakers, an atomizing, "minimize-winners"

criterion for cloning is needed. However, this solution would seem neither necessary nor warranted.

Notwithstanding the facts of Turkish, speakers in other languages and performing other novel-word tasks do use lexical correlations which lack a corresponding constraint in CON (Moreton & Amano 1999, Albright 2002, Albright & Hayes 2002, Ernestus & Baayen 2003), indicating that speakers are capable of such computation. In that case, atomized indexed constraints alone are not enough to produce the Turkish results. An additional stipulation is required, that this ability is suppressed when assigning novel words to exceptionality classes; yet this leads to a curious view of phonology. Whereas the grammar is usually the store of generalizations, just in the case of exceptionality, it is a store of highly detailed idiosyncrasy, and just in that case speakers ignore their usual, lexical store of idiosyncrasy and turn to the grammar. More satisfying would be to find some other explanation of the Turkish data. While that would take us well beyond this scope of this paper, it can be noted that what is required is a mechanism that can filter the lexical information in some way. That mechanism needn't be part of the OT grammar. Indeed, if it is true that learners build certain constraints during learning (Flack 2007b, Hayes & Wilson 2008, Hayes 2014), then there must exist EXTRA-grammatical generalization devices, which may provide the lexicon-filtering power needed. For now I conclude that that Becker's proposal follows from just one possible solution to an interesting puzzle, however both the puzzle and solution are outliers relative to what else we know. In contrast, a "maximize-winners" criterion leads to the learning of restrictive grammars, and on those general grounds would appear correct.

# 9  Discussion

## 9.1  The case against concrete accounts

Throughout this paper, I have considered only the ABSTRACT phonological approach to analyzing exceptionality, gradually building the argument that its superiority to the morphological approach lies in the fact that it localizes exceptionality to specific $\varphi$ elements, which are the elements in terms of which the relevant computation must be carried out. CONCRETE phonological approaches also localize exceptionality at a sub-morphological level, but compared to the abstract approach they are ill-suited to learning, and to seriality, as follows.

Lexical indexation is an ideal response to BCD inconsistency, because it annotates the lexicon with indices which are invisible to all previously installed constraints. This guarantees, without needing to check, that all previously accounted-for WLPs remain accounted for. Even if some of them contain lexical $\varphi$ elements which acquire a new index, their violations of all previously ranked constraints remain unchanged, since no previously-ranked constraint is sensitive to the new index. In contrast, the alteration of phonological form — for example, removing a root node from certain segments — may very well alter the evaluation of WLPs by already-ranked constraints, thus it requires a re-evaluation of the entire ranking. It is not possible to simply repair an inconsistency

and resume the BCD process. An ABSTRACT phonological account is therefore easier to learn.

In serial theories, concrete phonological approaches face the problem that in non-initial strata, it is possible that a preceding stratum will have removed, altered, moved or introduced, those aspects of phonological form which should function as pseudo-indices, which lack Consistency of Exponence. This opens up the possibility of all manner of phonological manipulations of exceptionality, for which I am unaware of any evidence.

Taking a more historical view, Chomsky (1964) criticized concrete phonological accounts espoused by structuralists (e.g. Bloomfield 1939) for the proliferation of underlying segments that they entailed. To the extent that such concerns matter to modern phonological theories, Φ-indexation avoids such proliferation by augmenting representations with non-phonological indices (cf §7), rather than additional underlying phonological distinctions.

## 9.2 The case against M-indexing

In §7 and §8 I showed that for both constraint evaluation and constraint learning, exceptionality is calculated in terms of phonological elements, not morphs. Morphs can be brought into the picture, but at additional computational cost and to no effect. Perhaps, however, it is nevertheless empirically true that exceptionality is inherently morph-bound. If that were so, then phonological exceptionality in any morph $m$ would always be either (i) uniform throughout all phonological exponents of $m$ or (ii) entirely predictably located within $m$. Yet this is not the case. If we accept something along the lines of Anderson's (1982) analysis of French schwa as an exceptionally-deleting /ø/ vowel, then that exceptional property is neither uniform throughout morphs nor does it have a predictable location. Similarly, in Turkish, non-high round vowels are phonotactically exceptional outside the first syllable (Clements & Sezer 1982; Hulst & Weyer 1991), yet the location of the exception is not predictable, as seen in a comparison of *otoban* 'highway', *monoton* 'monotone', *fenomen* 'phenomenon' and *paradoks* 'paradox'. There is no doubt that in most known cases, exceptionality does happen to be either uniform or predictable within a morph, but this follows uninterestingly from the fact that most exceptional morphs are short, or that most phonological alternations are either local, in which case their location inside a morph is predictably restricted to an edge, or domain-spanning, in which case the morph acts uniformly. However, when such uninformative cases are set aside, the small, informative residue of evidence does not support the morph-based view.

A second argument in defense of M-indices might be that morphs, and not $\varphi$ elements, belong to lexical strata, and that a single morphological diacritic can therefore coherently index a whole set of phonological exceptionality patterns, patterns which impact different parts of the morph and which therefore would be only incoherently represented by individual diacritics on $\varphi$ elements. Yet the empirical falsity of this claim has long been recognized. SPE (Chomsky & Halle 1968) permitted both stratal diacritics, later labeled MORPHOLOGICAL FEATURES (Postal 1968) and more specific RULE FEATURES (Lakoff 1970),

in view of the fact that distinct phonological patterns associated with strata are not uniformly attested in all morphs. For more recent work, see for example Labrune (2012: 71,72,85ff) on Japanese.

A third argument in defense of M-indices might be that since some kinds of phonological exceptionality are cyclic (§7.1), and since cycles are inherently tied to morphology, not $\varphi$ elements, then something like M-indices are required anyhow, in which case $\Phi$-indices are redundant. I would suggest that this is a category mistake. While it is true that cycles are inherently tied to morphology, they are tied not to morphs, but to morphological operations. Some operations are non-concatenative and hence morph-free (Anderson 1992). Cyclicity effects, therefore, are about how phonological subgrammars correlate with OPERATIONS; in contrast, $\Phi$-indices are about correlation with FORMS. M-indices fall uncomfortably in between. Since they are inherently attached to morphs, they will be unavailable for the triggering of cyclicity effects associated with non-concatenative operations. And, as we have seen above, they are inefficient, and in all likelihood insufficient, devices for exceptionality of forms.

## 10  Conclusion

For most of the generative period, an implicit assumption has been that we must choose between a concrete phonological and a diacritic morphological approach to phonological exceptionality.[19] But the argument from learning is that the correct theory is phonological and diacritic, based on lexical phonological indices which are visible to the phonology but not manipulable by it. The concrete phonological approach, whose pseudo-indices are manipulable by the phonology, is ill-suited to efficient learning (§9.1). Diacritic approaches are well suited to learning; however the computation of exceptionality is simply not carried out in terms of morphs, rather its currency is lexical phonological elements. This is true for both constraint evaluation (§7.1) and the learning of co-indexation (§8.2). Concurrently, plausible assumptions about learning ensure that a diacritic phonological account does not suffer from overgeneration (§5.3), and reveal the need for a morphological analytic bias, operationalized here as the Morphological Coherence Principle (§6). Finally, a morph-based diacritic theory appears empirically insufficient in the inevitably small number of cases that are informative (§9.2). No doubt there is much more to be said on the topic of exceptionality, but I hope to have established that the nature of exceptionality is, in essence, phonological and diacritic.

## Abbreviations

Abbreviations conform with the Leipzig glossing rules; in addition are: LEST 'lest' and SET 'inclusion/one of a group' (Dixon 1977a).

---

[19] Except, trivially, in purely abstract theories (e.g. Lamb 1966, Fudge 1967).

# References

Albright, Adam. 2002. Islands of reliability for regular morphology: evidence from Italian. *Language* 78(4). 684–709.

Albright, Adam & Bruce Hayes. 2002. Modeling English past tense intuitions with minimal generalization. In *Proceedings of the ACL-02 workshop on Morphological and phonological learning-Volume 6*, 58–69. Association for Computational Linguistics.

Anderson, Stephen R. 1974. *The organization of phonology*. New York: Academic Press.

Anderson, Stephen R. 1982. Where's morphology? *Linguistic Inquiry* 13(4). 571–612.

Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Anderson, Stephen R. 2008. Phonologically conditioned allomorphy in the morphology of Surmiran (Rumantsch). *Word Structure* 1(2). 109–134.

Anderson, Stephen R. 2015. Morphological change. In Claire Bowern & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 264–285. New York: Routledge.

Anderson, Stephen R. 2016. The role of morphology in transformational grammar. In Andrew Hippisley & Gregory T. Stump (eds.), *The Cambridge handbook of morphology*, 587–608. Cambridge: Cambridge University Press.

Anderson, Stephen R. 2017. Words and paradigms: Peter H. Matthews and the development of morphological theory. *Transactions of the Philological Society* 115. 1–13.

Angluin, Dana. 1980. Inductive inference of formal languages from positive data. *Information and control* 45(2). 117–135.

Baker, Carl L. 1979. Syntactic theory and the projection problem. *Linguistic Inquiry* 10(4). 533–581.

Becker, Michael. 2009. *Phonological trends in the lexicon: The role of constraints*. Amherst, MA: University of Massachusetts Ph.D. dissertation.

Becker, Michael, Nihan Ketrez & Andrew Nevins. 2011. The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish laryngeal alternations. *Language* 87(1). 84–125.

Benua, Laura. 1997. *Transderivational identity: phonological relations between words*. University of Massachusetts, Amherst PhD thesis.

Bermúdez-Otero, Ricardo. 2016. Stratal Phonology. In S.J. Hannahs & Anna R. K. Bosch (eds.), *The Routledge handbook of phonological theory*. Abingdon: Routledge.

Bernhardt, Barbara H. & Joseph P. Stemberger. 1998. *Handbook of phonological development from the perspective of constraint-based nonlinear phonology*. Academic press.

Bloomfield, Leonard. 1939. Menomini morphophonemics. *Travaux du Cercle Linguistique de Prague* 8. 105–115.

Bonet, Eulàlia, Maria-Rosa Lloret & Joan Mascaró. 2007. Allomorph selection and lexical preferences: Two case studies. *Lingua* 117(6). 903–927.

Bowern, Claire, Erich R. Round & Barry J. Alpher. in revision. The Phonetics and Phonology of Yidiny Stress.

Chomsky, Noam. 1964. *Current issues in linguistic theory*. Mouton: The Hague.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English.* New York: Harper & Row.

Clements, George N. & Engin Sezer. 1982. Vowel and consonant disharmony in Turkish. In Harry Van der Hulst & Norval Smith (eds.), *The structure of phonological representations*, vol. 2, 213–255.

Coats, Herbert S. 1970. Rule environment features in phonology. *Papers in Linguistics* 2(?). 110–140.

Coetzee, Andries W. 2009. Learning lexical indexation. *Phonology* 26(1). 109–145.

Crowhurst, Megan & Mark Hewitt. 1995. Prosodic overlay and headless feet in Yidiny. *Phonology* 12(1). 39–84.

Dixon, Robert M. W. 1977a. *A grammar of Yidiny.* Cambridge: Cambridge University Press.

Dixon, Robert M. W. 1977b. Some Phonological Rules in Yidiny. *Linguistic Inquiry* 8(1). 1–34.

Dixon, Robert M. W. 1991. *Words of our country: stories, place names and vocabulary in Yidiny, the Aboriginal language of the Cairns-Yarrabah region.* Brisbane: University of Queensland Press.

Ernestus, Mirjam & R. Harald Baayen. 2003. Predicting the unpredictable: interpreting neutralized segments in dutch. *Language* 79(1). 5–38.

Finley, Sara. 2010. Exceptions in vowel harmony are local. *Lingua* 120(6). 1549–1566.

Flack, Kathryn. 2007a. Templatic morphology and indexed markedness constraints. *Linguistic Inquiry* 38(4). 749–758.

Flack, Kathryn G. 2007b. *The sources of phonological markedness.* Amherst, MA: University of Massachusetts Ph.D. dissertation.

Fudge, Eric C. 1967. The nature of phonological primes. *Journal of Linguistics* 3(1). 1–36.

Fukazawa, Haruka. 1999. *Theoretical implications of OCP effects on features in Optimality Theory.* University of Massachusetts, Amherst Ph.D. Dissertation. https://rucore.libraries.rutgers.edu/rutgers-lib/38461/.

Gouskova, Maria. 2007. The reduplicative template in Tonkawa. *Phonology* 24(3). 367–396.

Gouskova, Maria. 2012. Unexceptional segments. *Natural Language & Linguistic Theory* 30(1). 79–133.

Hall, Nancy. 2001. Max-Position drives iterative footing. In Karine Megerdoomian & Leora A. Bar-el (eds.), *Proceedings of the 20th West Coast Conference on Formal Linguistics.* Cascadilla Press.

Halle, Morris & William J. Idsardi. 1995. General properties of stress and metrical structure. In John Goldsmith (ed.), *The handbook of phonological theory*, 403–443. Oxford: Blackwell Publishing.

Hammond, Michael. 2000. There is no lexicon. *Coyote Papers* 10. 55–77.

Hayes, Bruce. 1982. Metrical structure as the organizing principle of Yidiny phonology. In Harry Van der Hulst & Norval Smith (eds.), *The Structure of Phonological Representations, Part I*, 97–110. Dordrecht: Foris.

Hayes, Bruce. 1985. *A metrical theory of stress rules.* New York: Garland.

Hayes, Bruce. 1997. Anticorrespondence in Yidiɲ. Ms., University of California, Los Angeles.

Hayes, Bruce. 2014. Comparative phonotactics. In. Second International Workshop on Phonotactics. Pisa: Dept. of Linguistics Scuola Normale Superiore.

Hayes, Bruce & Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39(3). 379–440.

Hulst, Harry van der & Jeroen van de Weyer. 1991. Topics in Turkish Phonology. In R. Boeschoten & L. Verhoeven (eds.), *Turkish Linguistics Today*, 11–159. Leiden: Brill.

Hyde, Brett. 2012. The odd-parity input problem in metrical stress theory. *Phonology* 29(3). 383–431.

Inkelas, Sharon. 1994. The consequences of optimization for underspecification. In Jill Beckman (ed.), *Proceedings of NELS 25*, 287–302. Amherst: GLSA.

Inkelas, Sharon & Cheryl Zoll. 2007. Is grammar dependence real? A comparison between cophonological and indexed constraint approaches to morphologically conditioned phonology. *Linguistics* 45(1). 133–171.

Itô, Junko & Armin Mester. 1999. The phonological lexicon. In Natsuko Tsujimura (ed.), *The handbook of Japanese linguistics*, 62–100. Oxford: Blackwell Publishing.

Jurgec, Peter. 2010. Disjunctive lexical stratification. *Linguistic Inquiry* 41(1). 149–161.

Kager, René. 1993. Alternatives to the iambic-trochaic law. *Natural Language & Linguistic Theory* 11(3). 381–432.

Kager, René. 1996. On affix allomorphy and syllable counting. In Ursula Kleinhenz (ed.), *Interfaces in phonology* (Studia Grammatica 41), 155–171. Berlin: Akademie-Verlag.

Kiparsky, Paul. 1973. *Abstractness, opacity and global rules*. Bloomington: Indiana University Linguistics Club.

Kiparsky, Paul. 1982a. From cyclic phonology to lexical phonology. In Harry van der Hulst & Norval Smith (eds.), *The structure of phonological representations*, 131–175. Dordrecht: Foris.

Kiparsky, Paul. 1982b. From cyclic phonology to lexical phonology. In Harry Van der Hulst & Norval Smith (eds.), *The Structure of Phonological Representations, Part I*, 131–175. Dordrecht: Foris.

Kiparsky, Paul. 2000. Opacity and cyclicity. *The Linguistic Review* 17(2–4). 351–366.

Labrune, Laurence. 2012. *The Phonology of Japanese*. Oxford: Oxford University Press.

Lakoff, George. 1970. *Irregularity in syntax*. New York: Holt, Rinehart & Winston.

Lamb, Sydney M. 1966. Prolegomena to a theory of phonology. *Language* 42(2). 536–573.

Lombardi, Linda. 2002. Coronal epenthesis and markedness. *Phonology* 19(02). 219–251.

Łubowicz, Anna. 2005. Locality of conjunction. In John Alderete, Chung-hye Han & Alexei Kochetov (eds.), *Proceedings of the 24th West Coast Conference on Formal Linguistics*, 254–262.

Mahanta, Shakuntala. 2008. *Directionality and locality in vowel harmony: With special reference to vowel harmony in Assamese*. Netherlands Graduate School of Linguistics Ph.D. Dissertation.

Mascaró, Joan. 1996. External allomorphy and contractions in Romance. *Probus* 8(2). 181–206.

McCarthy, John & Alan Prince. 1993a. Generalized alignment. In Geert Booij & Jaap van Marle (eds.), *The yearbook of morphology 1993*, 79–153. Dordrecht: Kluwer Academic Press.

McCarthy, John & Alan Prince. 1993b. Prosodic morphology i: constraint interaction and satisfaction. Ms., University of Massachusetts, Amherst.

McCarthy, John & Alan Prince. 1995. Faithfulness and reduplicative identity. In *University of Massachusetts occasional papers 18: Papers in Optimality Theory*, 249–384. Amherst: Graduate Linguistic Student Association, UMass.

McCarthy, John J. 2003. Comparative markedness. *Theoretical linguistics* 29(1-2). 1–51.

Merchant, Nazarré Nathaniel. 2008. *Discovering underlying forms: Contrast pairs and ranking*. New Brunswick, NJ: Rutgers University PhD thesis.

Mester, Armin. 1994. The quantitative trochee in Latin. *Natural Language and Linguistic Theory* 12(1). 1–61.

Moreton, Elliott. 2008. Analytic bias and phonological typology. *Phonology* 25(1). 83–127.

Moreton, Elliott & Shigeaki Amano. 1999. Phonotactics in the perception of Japanese vowel length: evidence for long-distance dependencies. In *Proceedings of the 6th European Conference on Speech Communication and Technology*, 82. https://pdfs.semanticscholar.org/f76d/0ebc91f9414ecc23bc36420662cc33776267.pdf.

Myers, Scott. 1997. OCP effects in Optimality Theory. *Natural Language and Linguistic Theory* 15(4). 847–892.

Orgun, Cemil. 1996. *Sign-based morphology and phonology with special attention to Optimality Theory*. Berkeley, CA: University of California, Berkeley PhD thesis.

Pater, Joe. 2000. Non-uniformity in English secondary stress: the role of ranked and lexically specific constraints. *Phonology* 17(2). 237–274.

Pater, Joe. 2006. The locus of exceptionality: morpheme-specific phonology as constraint indexation. In Leah Bateman & Adam Werle (eds.), *University of Massachusetts Occasional Papers 32: Papers in Optimality Theory*, 1–36. Amherst, MA: GLSA.

Pater, Joe. 2009. Morpheme-specific phonology: Constraint indexation and inconsistency resolution. In Steve Parker (ed.), *Phonological Argumentation: Essays on Evidence and Motivation*, 123–154.

Poplack, Shana, Alicia Pousada & David Sankoff. 1982. Competing influences on gender assignment: Variable process, stable outcome. *Lingua* 57(1). 1–28. http://www.sciencedirect.com/science/article/pii/0024384182900687, accessed 2016-10-02.

Postal, Paul. 1968. *Aspects of phonological theory*. New York: Harper & Row.

Prince, Alan & Paul Smolensky. 2004. *Optimality Theory: Constraint interaction in generative grammar*. Malden, MA: Wiley-Blackwell.

Prince, Alan & Bruce Tesar. 2004. Learning phonotactic distributions. In René Kager, Joe Pater & Wim Zonneveld (eds.), *Constraints in phonological acquisition*, 245–291. Cambridge: Cambridge University Press.

Pruitt, Kathryn. 2010. Serialism and locality in constraint-based metrical parsing. *Phonology* 27(3). 481–526.

Round, Erich R. 2013. *Kayardild Morphology and Syntax*. English. Oxford: Oxford University Press.

Round, Erich R. in progress. Unsyllabified moras and length in Yidiny.

Russell, Kevin. 1995. Morphemes and candidates in Optimality Theory.

Smith, Jennifer L. 2004. Making constraints positional: toward a compositional model of CON. *Lingua* 114(12). 1433–1464.

Tesar, Bruce & Paul Smolensky. 2000. *Learnability in Optimality Theory*. Cambridge: MIT Press.

Tesar, Bruce B. 1995. *Computational optimality theory*. Boulder: University of Colorado Ph.D. dissertation.

Tesar, Bruce B. 2007. Learnability. In Paul de Lacy (ed.), *The Cambridge handbook of phonology*, 555–574. Cambridge: Cambridge University Press.

Tranel, Bernard. 1996a. Exceptionality in Optimality Theory and final consonants in French. In Karen Zagona (ed.), *Grammatical Theory and Romance Languages*, 275–293. Amsterdam: John Benjamins.

Tranel, Bernard. 1996b. French liaison and elision revisited: A unified account within Optimality Theory. In Claudia Parodi, Carlos Quicoli & Mario Saltarelli (eds.), *Aspects of Romance linguistics*, 433–455. Washington, DC: Georgetown University Press.

Van Oostendorp, Marc. 2007. Derived environment effects and consistency of exponence. In Sylvia Blaho, Patrick Bye & Martin Krämer (eds.), *Freedom of analysis?*, 123–148. Berlin/New York: Mouton de Gruyter.

Wolf, Matthew. 2015. Lexical insertion occurs in the phonological component. In Eulàlia Bonet, Maria-Rosa Lloret & Joan Mascaró (eds.), *Understanding Allomorphy: Perspectives from Optimality Theory*, 361–407. London: Equinox Publishing.

Zoll, Cheryl. 2001. Constraints and representation in subsegmental phonology. In Linda Lombardi (ed.), *Segmental phonology in Optimality Theory: Constraints and representations*, 46–78.

Zoll, Cheryl C. 1996. *Parsing below the segment in a constraint based framework*. Berkeley, CA: University of California, Berkeley Ph.D. dissertation.

Zonneveld, Wim. 1978. *A formal theory of exceptions in generative phonology*. Lisse: Peter de Ridder.