Bjorkman, Bronwyn; Peter Jurgec

Regularizing effects of indexation to complex constituents

This abstract presents a modest extension of indexed constraints, one that allows us to capture a class of long-distance morphologically derived environment effects (MDEEs) that have been previously unexplained. Our central proposal is that indexed constraints can apply not only to individual morphemes, but also to potentially complex constituents such as the *stem*. This modification allows us to derive the regularization patterns typical of long-distance MDEEs, if complex constituents such as stems are treated as lexically exceptional only when every morpheme contained within them is independently exceptional.

Background: Phonology is frequently sensitive to properties of the morphemes to which an operation or constraint applies. In Optimality Theory (OT), one way this has been accounted for is by allowing constraints to be *indexed* to certain classes of words, e.g. to roots (McCarthy & Prince, 1993), loanwords (Itô & Mester, 1995, 2001), nouns (Smith, 2001, 2006), specific lexical items (Pater, 2000; Becker et al., 2011), or exceptional suffixes (Pater, 2007, 2009).

An important observation of work on constraint indexation has been that morphological sensitive constraint evaluation is *local*: the presence of an exceptional affix in a word does not cause all other affixes to behave as though they were also exceptional. To account for this, Pater (2007, 2009) explicitly limits the reach of indexed constraints, so that the locus of violation of an indexed constraint must be part of the morpheme with that index.

Puzzle: The locality of indexed constraint evaluation is challenged by the existence of clearly non-local MDEEs, described in recent work by Jurgec (2014) and Gouskova & Linzen (to appear). They describe cases in which an exceptional property of a root is *suppressed* in certain morphological contexts. Jurgec (2014) describes a pattern from Dutch, for example, in which the segment [ɹ] is possible in certain loanwords when they appear in underived forms (e.g. *Flo[ɹ]ida* 'Florida'), but replaced by [ʀ] in derived words, including diminutives and derived adjectives (*Flo[ʀ]ida-tje* 'Florida-DIM'). If understood in terms of faithfulness constraints indexed to exceptional roots, or markedness constraints indexed to affixes, such alternations appear to require non-local interactions between affix and non-adjacent root segments.

Proposal: We propose that apparently non-local effects can be captured in terms of local constraint evaluation, but only if we allow constraints to be indexed not only to individual morphemes, but also to complex morphological constituents such as stems and morphological words. We share with many others the view that indexed constraints can be sensitive to both morpheme type (e.g. root, affix) and to arbitrary lexical specification (McCarthy & Prince, 1993; Itô & Mester, 1995, 2001; Beckman, 1998; Pater, 2000; Flack, 2007; Gouskova, 2007; Jurgec, 2010). Our extension is that indexed constraints must further be specified for the morphological domain to which they apply, whether this is a single morpheme or a constituent consisting of a root plus zero or more affixes (i.e. a stem or word). Assuming that morphosyntactic notions of headedness are not visible within the phonological component, however, we propose that if a constraint indexed to stems or words is further restricted to some arbitrary class L, it will apply only if all morphemes in the stem or word are equally specified as L.

This predicts a pattern in which marked structures are preserved in stems or words that contain a single morpheme (i.e. a root belonging to an indexed class), but not in stems or words that are complex (i.e. containing at least one non-exceptional affix). We present corroborating data from Tagalog and 15 other languages.

Consider the Tagalog labial alternations in (1). Tagalog allows f in bare loanword roots, but not in prefixed or suffixed words, in which case p surfaces. Note that the segmental content of the triggering affix does not matter. We account for this pattern using constraints that can be indexed to both morphological properties (i.e. morpheme type, lexical exceptionality) and to a

morphological domain (i.e. morpheme, stem, word).

(1) Tagalog MDEE: f → p (Zuraw 2006; Jurgec 2014)

BARE ROOT $\boxed{\text{f}}$		PREFIXED $\boxed{\text{p}}$		SUFFIXED $\boxed{\text{p}}$	
<u>f</u>ilipino	'Filipino'	mag-<u>p</u>ilipino	'language'	<u>p</u>ilipino-ŋ	'DEF'
<u>f</u>iesta	'feast'	pam-<u>p</u>ista	'INSTR'	<u>p</u>ista-han	'festival'

For Tagalog, the active constraint is defined in (2):

(2) IDENT$_{\text{L,Word}}$
No change in any segment that is part of the phonological exponent of a *Word* specified as *L[oanward]*. (A constituent is treated as specified for some property P iff all morphemes within that constituent are specified as P.)

This constraint does not apply in suffixed (3) or prefixed words, because the word-level domain contains non-L-marked morphemes (assuming that no prefixes or suffixes are L-marked in Tagalog). It does apply in non-affixed words (4), where the word consists of the L-marked root morpheme only. As a result, f cannot surface in affixed words, but can surface in bare roots.

(3) *Root ≠ Word*: IDENT$_{\text{L,ω}}$ does not apply

/filipino$_{\text{L}}$-ŋ/	IDENT$_{\text{L,ω}}$	*f	IDENT$_{\text{L}}$
filipino$_{\text{L}}$-ŋ	d.n.a.	*!	
☞ pilipino$_{\text{L}}$-ŋ	d.n.a.		*

(4) *Root = Word*: IDENT$_{\text{L,ω}}$ applies

/filipino$_{\text{L}}$/	IDENT$_{\text{L,ω}}$	*f	IDENT$_{\text{L}}$
☞ filipino$_{\text{L}}$		*	
pilipino$_{\text{L}}$	*!		*

The same type of constraint can account for the Dutch loss of [ɹ] discussed above, though to indexation to the *stem* rather than to the word. We show that it can extend further, to account for long-distance MDEEs across 15 other languages, including Russian, Polish, English, Catalan, Hungarian, Basque, and Slovenian. These languages all exhibit exceptional patterns that appear to be overridden in more complex morphological environments, a type of morphological sensitivity that has been challenging for most rule- or constraint-based approaches.

Discussion: Our proposal maintains locality of constraint evaluation by attributing non-local MDEEs to the interaction of phonology with complex morphological constituents. Gouskova & Linzen (to appear) propose a different account of these effects within a Maximum Entropy grammar that uses weighted constraints to model probabilistic outputs. They account for long-distance MDEEs by proposing that morphemes can be associated with non-locally evaluated regularization factors, alongside locally evaluated scaling factors for individual constraints. By maintaining local evaluation in one aspect of constraint evaluation, but requiring its *absence* in another, their proposal weakens the overall role of locality in the phonological grammar.

A question that remains for our account is whether markedness constraints can be indexed to stems in the same way that we have proposed faithfulness constraints are. Such constraints would predict cases where marked structures can occur only in affixed words, and to our knowledge such patterns are unattested. This leads to the interesting possibility that indexation to complex morphological constituents may be possible only for faithfulness constraints, echoing earlier proposals that indexation more generally is possible only in the domain of faithfulness (Itô & Mester, 1995, 1999; Inkelas & Zoll, 2007). We leave this for further investigation.

Conclusions: Morphologically derived environment effects constitute an unusual case of long-distance interactions arising from the interplay of exceptional phonological patterns and morphological structure of words. These patterns have constituted a serious challenge for theories of locality and exceptionality in phonology, but we show that a simple extension of lexical indexation so that it refers to morphological domains can successfully account for them.