

### Intra-paradigmatic predictability and morphological base identification in Texistepec Popoluca

This paper describes and analyzes the inflectional paradigms of verbs and nouns in Texistepec Popoluca, a Zoquean language with numerous morphophonemic alternations. I ask the question of what lexical representations Texistepec Popoluca speakers store in their lexicons in the face of these alternations. I consider possible answers to this question from the perspectives two prominent models of morphological paradigms; the storage of only abstract underlying representations, as in Generative phonology (Kenstowicz and Kisseberth, 1977), and the storage of a single surface form in the paradigm as the unique base (Albright, 2002). In general, Albright's proposal holds numerous advantages over abstract underlying representations in terms of learnability, but I find that the morpho-phonology of Texistepec Popoluca poses a particularly dire problem for the selection of a unique surface form as base.

Texistepec Popoluca does not have any small sub-classes of one or more irregularly inflected words, so the correct segmental representation of a given word is uniformly adequate for predicting all inflectional variants. However, a number of near mergers and marginal phonemic contrasts may obscure the correct segmental representations of words in this language. For example, the categories /i/ and /ɛ/ have undergone an almost complete merger to /ɛ/. The distinction is preserved only in syllable nuclei following the 3rd person ergative agreement prefix /j-/, as shown in (1). The 3rd person ergative form is therefore the crucially informative form for learning which category a particular word belongs to. This is the form selected as base by the Minimal Generalization Algorithm for unique base selection (Albright and Hayes, 2002). The Texistepec Popoluca learner must also struggle with the neutralization of the contrast between /i/ and /i/, both of which surface as [i] under prefixation with /j-/, as shown in (2). This contrast is easy to preserve, as long as we *do not* adopt the 3rd person ergative as the morphological base. Therefore, under unique base selection, to preservation of the /i/ vs. /ɛ/ contrast is in direct conflict with the preservation of the /i/ vs. /i/ contrast. A similar problem affects the nasal vs. oral contrasts between /m/ vs. /b/ and /n/ vs. /d/. In these cases, only the 3rd person absolutive and imperative forms of these words reveal a contrast that is neutralized elsewhere.

The table in (3) show partial paradigms for three pairs of words. Neutralized contrasts are indicated by a dotted line, and forms that preserve contrasts are shaded. It is obvious from this table, that no single form can be uniformly selected as the base for all paradigms. Unique base selection is guaranteed to neutralize a pervasive segmental contrast, so an adequate system of lexical representations in Texistepec Popoluca must either use abstract underlying representations or store more than one form from each paradigm as a base.

It is possible to distinguish between these two remaining possibilities in terms of the predictions they make about speakers' intuitions for novel words, and the predictions they make about the direction of historical changes. In Texistepec Popoluca, there are many words that simply never occur in the 3rd person ergative, or never occur in the absolutive or imperative, for semantic and pragmatic reasons. The correct representation of these words is under-determined by speakers' linguistic experience. Interestingly, speakers seem to be aware of this fact, and are ambivalent about how to treat novel derivatives of these words for which the relevant inflections become possible. Furthermore, a number of words that derived from both historical /i/ and historical /ɛ/ have come to have free variation with respect to the vowel. The same is attested found for the nasal vs. oral contrast. I discuss several problems that these facts pose for a strictly UR-based system of representations, and I conclude that Texistepec Popoluca speakers must have at least some knowledge of surface forms, beyond the putative UR. Phenomena like these are best accounted for by a model like the Minimal Generalization Algorithm, which easily and accurately represents gradient intuitions, and predicts the direction of historical changes.

These results suggest that an adequate representation of Texistepec Popoluca paradigms is one with **multiple bases**. I propose a modified version of the Minimal Generalization Algorithm, in which base selection is restricted but not unique. I show how the model I propose learns an adequate and thorough grammar of Texistepec Popoluca, while accurately predicting the kinds of historical changes and native speaker intuitions that I report.

## Examples

- (1) a.  $\widehat{\eta g \epsilon j}$  'fly' 1ERG  
 $k \epsilon j$  'fly' 3ABS  
 **$k j \epsilon j$**  'fly' 3ERG
- b.  $\widehat{\eta g \epsilon j}$  'signal' 1ERG  
 $k \epsilon j$  'signal' 3ABS  
 **$k i j$**  'signal' 3ERG
- (2) a.  $\widehat{m b i \eta}$  'fall' 1ERG  
 $p i \eta$  'fall' 3ABS  
 **$p i \eta$**  'fall' 3ERG
- b.  $\widehat{m b \epsilon \eta}$  'paint' 1ERG  
 $p \epsilon \eta$  'paint' 3ABS  
 **$p i \eta$**  'paint' 3ERG

(3)

Putative UR	1ERG <b>n-</b>	1ABS <b>k-</b>	3ERG <b>j-</b>	3ABS <b>Ø-</b>
/naj/	[naj]	[kdaj]	[d <sup>h</sup> aj]	[naj]
/daj/	[naj]	[kdaj]	[d <sup>h</sup> aj]	[daj]
/piη/	[ $\widehat{m b i \eta}$ ]	[kp $\dot{i}$ η]	[p $\dot{i}$ η]	[p $\dot{i}$ η]
/piη/	[ $\widehat{m b \epsilon \eta}$ ]	[kp $\epsilon$ η]	[p $\dot{i}$ η]	[p $\epsilon$ η]
/kεj/	[ $\widehat{\eta g \epsilon j}$ ]	[xkεj]	[k $\dot{j}$ εj]	[kεj]
/kij/	[ $\widehat{\eta g \epsilon j}$ ]	[xkεj]	[k $\dot{i}$ j]	[kεj]

## References

- Albright, Adam. 2002. The identification of bases in morphological paradigms. Doctoral Dissertation, UCLA.
- Albright, Adam, and Hayes, Bruce. 2002. Modeling English Past Tense Intuitions with Minimal Generalization. Paper presented at *Sixth Meeting of the ACL Special Interest Group in Computational Phonology*, Philadelphia.
- Kenstowicz, Michael, and Kisseberth, C. 1977. *Topics in Phonological Theory*. New York: Academic Press.