Inferential Organization in Grammar Systems: Word Structure. Paradigm Organization and Learnability

Speakers of languages with complex morphology confront a significant learning task This task is formulated as the Paradigm Cell Filling Problem (PCFP) in Ackerman et al. (2009):

> PARADIGM CELL FILLING PROBLEM: Given exposure to an inflected
> wordform of a novel lexeme, what licenses reliable inferences about the
> other wordforms in its inflectional family?

The challenge in the PCFP, is not new (Paul 1891, Hockett 1967, Paunonen 1976, Bybee 1985, Anttila 1989, Thyme 1993, Wurzel 1989, see also Fertig 2013): analogical inferences from (incomplete sets of) forms belonging to known inflectional patterns permit reasonable guesses concerning likely candidates for unknown forms. Observations about implicational organization have been quantified in recent word-based approaches to morphological analysis (see overview in Blevins 2016), where two interdependent dimensions of part/whole relations are developed: the internal structure of words interpreted in terms of discriminability among related words and the external relations among words as reflected in paradigm organization. These dimensions are exemplified in the Fur patterns (Nilo-Saharan) Waag 2010) for 1st person singular completive versus the 3rd person singular completive for the verb `to speak' in (1), where the combinatorics of affixes and stem variation are reflected in segment length, tonal melodies and metathesis which distinguish related words.

1a. ʔ-**írs**íŋɔ́   1b. **rì**sìŋɔ̀

   1SG-spoke   spoke
   `I spoke'              `s/he spoke'

The tonal melodies exhibit opposite values for person contrasts. 1st singular is associated with a prefix i.e., ʔ , while its stem represents a methathetic variant of the 3rd singular stem form, i.e., **ír**X versus **rì**X. Knowing the form for 1sg is to know the form of the 3sg, and vice versa. The mutual inferential relations in (1) are trivial, but become more complex when the whole system of Fur inflection is considered.

Calculations of informativity concerning patterned ingredients and meanings associated with them is the object of Information-Theoretic measures in word-based formal models. The L(ow) C(onditional) E(ntropy) C(onjecture) (Ackerman and Malouf 2013) is a cross-linguistic hypothesis about morphological organization: it displays low conditional entropies, reflecting high predictability between known words and their unknown variants. The LCEC is a way of solving the PCFP, providing learners with cues to facilitate good guesses about previously unencountered words.

The PCFP is emphasized in research showing inflected forms display Zipfian distributions: small numbers of inflected words are heard frequently providing partial paradigm information, while increasing the corpus size does not provide exposure to the "missing" words from the complete paradigm, but merely reinforces the distributions found in smaller samples. (Chan 2008, Yang 2010, to appear, Bonami and Beniamine 2015, Ramscar and Blevins 2015).

Following Bonami and Beniamine 2015, Ramscar and Blevins 2015 and Blevins 2016, we argue that the learning paradox of skewed and incomplete paradigms necessitates something like the LCEC and correlatively, the type of word-based morphological model within which it operates. We explore empirical data that confirm and/or challenge the LCEC, suggesting that both help to refine the possible scope of the conjecture and the nature of its extensions to new data sets.