Investigating inflection as a complex system

Vito Pirrelli

Istituto di Linguistica Computazionale, Consiglio Nazionale delle Ricerche, Pisa Italy

From a cross-linguistic perspective, different inflection systems appear to apportion word processing costs differently, depending on when and where, in the full form, morpho-lexical and morpho-syntactic information is encoded. The resulting balance is the outcome of an interaction between form frequency and morphological productivity, responding to basic communicative requirements. Big families of stem-sharing inflected forms constitute the productive core of an inflection system. This core is easy to learn, as it requires memorization of one stem only, with all inflected forms being redundantly built upon it. Unsurprisingly, generalizable paradigms are less sensitive to token frequency effects, and tend to be located in the long, low-frequency tail of the Zipfian distribution of word forms. In contrast, the head of the Zipfian distribution mostly contains small families of alternating and possibly suppletive stems, which, however shorter, morpho-phonologically simpler and easier to process, require high token frequency to be learned and resist pressure towards regularization.

Of late, considerable converging computational evidence has accrued to support this picture (Ackerman and Malouf 2013, Balling & Baayen 2012, Blevins et al. 2017, Bonami & Beniamine 2016, Marzi et al. 2018, Marzi et al. in press, Pirrelli 2018). In the talk, I will show how the dynamic tension between ease of learning and ease of processing can shape and structure the inflection systems of typologically different languages. In the end, each language (and arguably each individual learner) is likely to strike a different balance, which nonetheless falls within a reasonably tight range of variation, bounded by a few learnability and processability constraints. This suggests that full investigation of morphological systems will likely benefit from the use of basic concepts from the toolkit of complexity theory in biological networks, such as emergence, non-linearity and self-organization.

References

Ackerman, F. and Malouf, R. (2013). Morphological organization: The low conditional entropy conjecture. *Language* 89, 429–464.

Balling, L. and Baayen, R.H. (2012). Probability and surprisal in auditory comprehension of morphologically complex words. *Cognition* 125, 80–106.

Blevins, J. P., Milin, P. and Ramscar, M. (2017). The zipfian paradigm cell filling problem. In F. Kiefer, J. P. Blevins and H. Bartos (Eds.) *Morphological Paradigms and Functions* (Leiden, Germany: Brill), 141-158.

Bonami, O. and Beniamine, S. (2016). Joint predictiveness in inflectional paradigms. *Word Structure* 9, 156–182.

Marzi, C., Ferro, M. and Pirrelli V. (in press). A Processing-oriented Investigation of Inflectional Complexity. *Frontiers in Communication*.

Marzi, C., Ferro, M., Nahli, O., Belik, P., Bompolas, S. and Pirrelli V. (2018). Evaluating Inflectional Complexity Crosslinguistically: a Processing Perspective, In: *LREC 2018 - Eleventh International Conference on Language Resources and Evaluation* (Miyazaki, Japan, 7-12/05 2018). Proceedings, article 745. European Language Resources Association (ELRA), 2018.

Pirrelli, V. (2018). Morphological theory and computational linguistics. In J. Audring and F. Masini (Eds.) *The Oxford Handbook of Morphological Theory* (Oxford, UK: Oxford University Press), 573-593.