The role of attraction-repulsion dynamics in simulating the emergence of inflectional class systems

Erich R. Round
Surrey Morphology Group; Univ. of Queensland;
Max Planck Inst. for the Science of Human History

Sacha Beniamine
Surrey Morphology Group

Louise Esher
CNRS LLACAN

1   Introduction

Dynamic models of paradigm change can elucidate how the simplest of processes may lead to unexpected outcomes, and thereby can reveal new potential explanations for observed linguistic phenomena. Ackerman & Malouf (2015) present a model in which inflectional systems reduce in disorder through the action of an attraction-only dynamic, in which lexemes only ever grow more similar to one another over time. Here we emphasise that: (1) Attraction-only models cannot evolve the **structured diversity** which characterises true inflectional systems, because they inevitably remove all variation; and (2) Models with both attraction and repulsion enable the emergence of systems that are strikingly reminiscent of morphomic structure such as inflection classes. Thus, just one small ingredient — change based on dissimilarity — separates models that tend inexorably to uniformity, and which therefore are implausible for inflectional morphology, from those which evolve stable, morpheme-like structure. These models have the potential to alter how we attempt to account for morphological complexity.

2   Structure in inflectional systems

Inflectional classes ('rhizomorphomes', Round, 2015) constitute groups of lexemes which share inflectional exponents; they are a type of 'morphomic', morphology-internal structure, mediating the mapping between content and form in inflection. Within complex inflectional systems in natural language, such structures are common, and are demonstrably both productive and psychologically real for speakers (Enger, 2014; Maiden, 2018); they are claimed to limit the complexity of the inflectional system by offering a systematic, recurrent and predictable means of distributing exponents (cf. Carstairs-McCarthy, 2010; Blevins, 2016).

A matter of ongoing debate is what kind of dynamics could potentially lead to such structure (Maiden, 2018; Carstairs-McCarthy, 2010). In this paper, we use computational iterated learning models to reveal for the first time some of the simplest conditions under which stable inflectional class systems can emerge. The insights afforded are of value both for our theoretical understanding of morphomic structure and for the formulation of explicit mathematical models of paradigm evolution, essential to tasks such as robust quantitative historical inference (Kelly & Nicholls, 2017).

3   Attraction-only models cannot evolve stable inflection classes

An early iterated learning model, implementing a simple paradigm cell filling task (Ackerman et al., 2009) in which a lexeme can change only by becoming more similar to another, is described in Ackerman & Malouf (2015). The initial input to the model consists of a lexicon in which paradigms are populated with randomly distributed exponents. At each cycle, the model must predict a held out value which we term the *focus*, at the intersection of a focal cell

and lexeme. To predict the value of the focus, the model (i) picks a non-focal cell, which we call the *pivot*, (ii) selects all lexemes which share the exponent of the focal lexeme in the pivot cell, (iii) observes the exponents of these lexemes in the focal cell (exponents which jointly constitute *evidence*), and (iv) selects the most frequent exponent in the evidence to replace the held-out focus. Figure 1 illustrates one cycle. The result of one cycle is the input to the next.
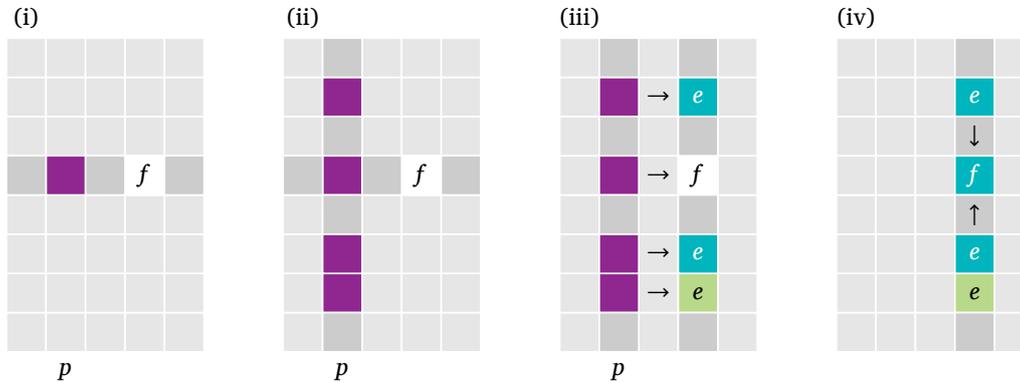

Figure 1: A cycle of Ackerman & Malouf's (2015) model. We label the focus f, the pivot cell p, and the evidence e.

This model is able to remove disorder from the system, because as lexemes change to be more like others, the dynamic is one of preferential attraction towards exponents that are already more frequent than their competitors. This rich-get-richer dynamic ensures that eventually, all lexemes converge on a single class (though discussion in Ackerman & Malouf (2015) focuses mainly on transitional states of the system just prior to ultimate uniformity). Ackerman & Malouf (2015) interpret this result as demonstrating the spontaneous emergence of self-organisational principles in morphological systems. We concur that the model exhibits self-organisation, but only of a radically homogenising kind. Here we investigate a family of minimally different dynamics (section 4); and their potential to generate outcomes which more closely resemble the morphological complexity of natural languages (section 5).

4   A family of dynamic, iterated learning models

As in Ackerman & Malouf's (2015) model, the initial input to our models consists of a lexicon in which paradigms are populated with randomly distributed exponents, and the basic task is again one of paradigm filling: at each cycle, the model must predict the *focus* based on *evidence* from multiple *pivots*, and the model's prediction is integrated into the lexicon input to the next cycle. All models in the family are similarly abstract and impoverished: the input provides indices for exponents rather than phonological forms; each cycle represents a general change in the system as a whole, with no method for capturing inter-speaker variation; and the models lack disruptive processes.

Our innovation is to alter the principles by which the paradigm filling task is accomplished, which we implement as modulable parameters to facilitate the controlled observation of their effects. We allow multiple pivot forms, reflecting the well-established observation that prediction based on multiple forms is more reliable than prediction based on pairings of forms or cells (Stump & Finkel, 2013; Bonami & Beniamine, 2016). In order to replicate the Zipfian frequency distribution observed for inflectional forms in natural language (Blevins et al., 2017), we introduce the option of frequency weighting in two ways: sampling foci in inverse proportion

to their lexeme frequency, and sampling pivots and evidence proportionally to their frequency. Furthermore, we allow the process to be influenced by *negative evidence* (Voorspoels et al., 2015), which introduces an evolutionary *repulsion dynamic*, which at the right strength can enable structured diversity to emerge. To do this, rather than looking only at lexemes which have the same exponent as the focal lexeme in the pivot, we also observe lexemes with different exponents, and use these lexemes to provide evidence of exponents which are expected to be different from the focus. The relative proportion of negative and positive evidence can be varied, and the incorporation of at least some negative evidence proves crucial to the emergence of morphomic structure.

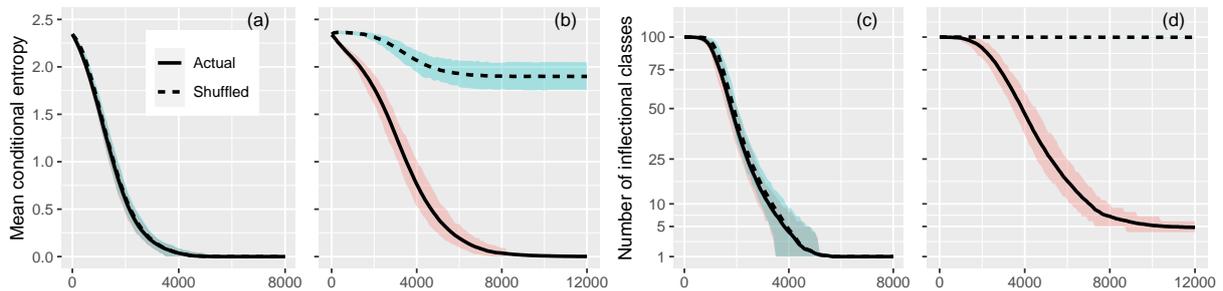## 5   Results and implications of the attraction-repulsion dynamic

Figure 2: Evolution of mean conditional entropy (a,b) and number of inflectional classes (c,d) for Ackerman & Malouf's (2015) model (a,c) and an attraction-repulsion model (b,d), initial conditions: 100 lexemes, 8 cells, 6 exponents. Lines show means of 100 runs, shading shows 90% variation. The horizontal axis measures evolutionary cycles.

Figure 2 compares a replication of Ackerman & Malouf's (2015) model with a model which attends to negative evidence as well as positive (with a 30%-70% weighting). Both models were run 100 times and throughout their evolutions we track two measures: firstly, the complexity of the paradigm cell filling task measured as mean conditional entropy between cells (Ackerman et al., 2009; Ackerman & Malouf, 2015); and secondly, the number of inflection classes (unique rows in the lexicon). As Figure 2 shows, there is a dramatic reduction in conditional entropy over time in both models. However, we emphasise that this entropy metric has the undesirable property of conflating two kinds of systemic order: (1) order simply due to **lack of variation**, and (2) order due to **structured variation**. Our second metric, the number of inflection classes present, differentiates these two 'orderly' scenarios from one another, and shows that the attraction-repulsion model not only lowers entropy, but also does so while preserving distinct, stable inflectional classes, numbering 4.8 on average.

We also verify this finding by a second method. If low entropy, or a low number of inflectional classes, is due overwhelmingly to a *lack of variation* at any stage in the evolutionary process, then taking the exponents of any given cell and shuffling them among the language's lexemes should have little effect. This is what we see in the A&M model (Figure 2a,c), where the entropy and number of classes of this 'shuffled' version of the evolving system barely differ from the actual system itself. In contrast, if these properties are due to *structured variation*, then shuffling should cause both to be disrupted and both metrics will rise. This is precisely what we observe in the attraction-repulsion model (Figure 2b,d).

More generally, we found that a system's progression to uniformity could be slowed by incorporating additional parameters: reducing the amount of available pivots and evidence,

or sampling these according to a Zipfian distribution to simulate frequency effects. However, stable inflectional structure still did not emerge in these enriched models, because the dynamic was still one of pure attraction. It is only once negative evidence is attended to, and thus a repulsion dynamic introduced into the system, that the models will consistently develop stable paradigmatic structure.

6 Conclusion

We have presented a new and fundamental mechanism for the spontaneous emergence of self-organising structure, which closely resembles what we find in natural language morphologies. Our findings have several topical ramifications. We clarify that morpheme-like structure is able to emerge via a dynamic process consisting merely of piecemeal individual changes, within a system which does not explicitly represent morphomic structure (e.g. by means of morphomic indices). The simplicity of the developments involved indicates that, contrary to the prevalent characterisation of morphomic structure as 'unnatural' morphology, it is entirely plausible to view such structure as a *natural phenomenon* liable to arise spontaneously within inflectional systems.

References

Ackerman, Farrell, James P. Blevins & Robert Malouf. 2009. Parts and wholes: implicative patterns in inflectional paradigms. In James P. Blevins & Juliette Blevins (eds.), *Analogy in Grammar*, 54–82. Oxford: OUP.

Ackerman, Farrell & Robert Malouf. 2015. The No Blur Principle Effects as an Emergent Property of Language Systems. In *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, vol. 41, doi:10.20354/B4414110014.

Blevins, James P. 2016. *Word and Paradigm Morphology*. Oxford: OUP.

Blevins, James P., Petar Milin & Michael Ramscar. 2017. The Zipfian Paradigm Cell Filling Problem. In Ferenc Kiefer, James P. Blevins & Huba Bartos (eds.), *Morphological paradigms and functions*, 141–158. Leiden: Brill.

Bonami, Olivier & S. Beniamine. 2016. Joint predictiveness in inflectional paradigms. *Word Structure* 9(2). 156–182. doi:10.3366/word.2016.0092.

Carstairs-McCarthy, Andrew. 2010. *The Evolution of Morphology*. Oxford: OUP.

Enger, Hans-Olav. 2014. Reinforcement in inflection classes: Two cues may be better than one. *Word Structure* 7(2). 153–181. doi:10.3366/word.2014.0064.

Kelly, Luke J. & Geoff K. Nicholls. 2017. Lateral Transfer in Stochastic Dollo Models. *The Annals of Applied Statistics* 11(2). 1146–1168. http://www.jstor.org/stable/26362220.

Maiden, Martin. 2018. *The Romance verb: Morphomic structure and diachrony*. Oxford: OUP.

Round, Erich R. 2015. Rhizomorphomes, meromorphomes, and metamorphomes. In Greville Corbett, Dunstan Brown & Matthew Baerman (eds.), *Understanding and Measuring Morphological Complexity*, 29–52. Oxford: OUP.

Stump, Gregory T. & Raphael Finkel. 2013. *Morphological Typology: From Word to Paradigm*. Cambridge: CUP.

Voorspoels, Wouter, Daniel J Navarro, Amy Perfors, Keith Ransom & Gert Storms. 2015. How do people learn from negative evidence? Non-monotonic generalizations and sampling assumptions in inductive reasoning. *Cognitive Psychology* 81. 1–25. doi:10.1016/j.cogpsych.2015.07.001.