Proceedings of the Society for Computation in Linguistics

Volume 3 Article 10

2020

Inflectional Networks: Graph-theoretic Tools for Inflectional Typology

Andrea D. Sims
The Ohio State University, sims.120@osu.edu

Inflectional networks: Graph-theoretic tools for inflectional typology

Andrea D. Sims
The Ohio State University
sims.120@osu.edu

Abstract

The interpredictability of the inflected forms of lexemes is increasingly important to questions of morphological complexity and typology, but tools to quantify and visualize this aspect of inflectional organization are lacking, inhibiting effective cross-linguistic comparison. In this paper I use metrics from graph theory to describe and compare the organizational structure of inflectional systems. Graph theory offers a well-established toolbox for describing the properties of networks, making it ideal for this purpose. Comparison of nine languages reveals previously unobserved generalizations about the typological space of morphological systems. This is the first paper to apply graph-theoretic tools to the goal of inflectional typology.

1 Introduction

Morphological typology has long classified languages in terms of how words are built out of morphemes. A typical formulation defines three or four types: isolating, agglutinative, fusional, and sometimes polysynthetic. More nuanced work seeks to break the types down into their component properties, with languages compared based on clusters of these (Plank, 1999). This newer approach is better able to capture cross-linguistic diversity, but it gives priority to the same aspects of morphological structure as the traditional classification scheme: syntagmatic relationships between formal elements (e.g. how many morphemes there are per word, known as the degree of synthesis (Comrie, 1981)), and the extent to which form-meaning mappings are isomorphic (e.g. as opposed to the language having inflection classes).

Morphological typologies built on these priorities fail to capture important aspects of morphological structure, corresponding to a distinction between two broad notions of morphological complexity that Ackerman and Malouf (2013) call Enumerative Complexity (E-complexity) and Integrative Complexity (I-complexity). E-complexity has to do with the size of a morphological system, e.g., the number of cells in lexemes' paradigms, the system's degree of synthesis, or the number of its inflection classes. I-complexity, on the other hand, has to do with the predictability of the inflected forms of lexemes. A morphological system is I-complex to the extent that the inflected forms of a newly encountered lexeme are unpredictable. This is a function of the distribution of elements in the system. Even systems with high E-complexity, such as a large number of inflection classes, may have low I-complexity, if morphological elements are distributed in ways that make them predictable (Ackerman and Malouf, 2013; Cotterell et al., to appear; Wurzel, 1989). I-complexity is thus oriented to the *internal organization* of inflectional systems, rather than their size. However, this organization is not captured by traditional typological measures.

In this paper I adopt metrics from graph theory, using them to describe and compare the internal organization of inflectional systems.[1] I analyze inflection classes as nodes in a network that are connected by the morphological structure that they have in common; two classes are connected if they use same exponent(s) to realize a set of morphosyntactic values. Conceptualized in this way, inflectional networks reflect the distribution of exponents in a language's inflectional system, and by extension, the internal organization of that system. Graph theory offers an established, widely applied toolkit for describing the properties of networks, making it a natural choice for application. While some interesting and previously unobserved generalizations emerge from comparison of different languages' inflectional networks, the primary goal of this paper is to demonstrate the usefulness of

[1] Data and code are available at https://github.com/sims120/inflectional-networks.

Proceedings of the Society for Computation in Linguistics (SCiL) 2020, pages 88-98.
New Orleans, Louisiana, January 2-5, 2020

	STOL	MESTO	KNIGA	KOST'
	'table'	'place'	'book'	'bone'
ACC.SG	stol	mesto	knigu	kost'
INS.SG	stolom	mestom	knigoj	kost'ju
DAT.PL	stolam	mestam	knigam	kostjam

Table 1: Partial inflectional paradigms of Russian nouns: three paradigm cells that differ in how informative they are about inflection class membership

applying graph-theoretic tools to inflectional data, and to outline some specific ways to quantify and compare inflectional systems.

Section 2 motivates an approach to typological comparison based on the paradigmatic distribution of exponents within an inflectional system. Section 3 gives a formal definition of an inflectional network. Section 4 discusses methodological choices. Section 5 introduces a variety of standard graph-theoretic measures, illustrating them using Russian noun inflection. Section 6 then compares nine languages' inflectional systems based on a couple of these measures, showing that their organization exhibits cross-linguistic diversity but also notable commonalities. Finally, Section 7 offers some conclusions and future directions.

2 Internal organization as a basis for inflectional typology

Work in the abstractive Word and Paradigm tradition (Blevins, 2006) emphasizes the paradigmatic or 'external' dimension of morphological structure: distributions of inflected word-forms within and across paradigms, and how these give rise to competition among inflectional exponents. In this view, word-internal/syntagmatic structure (e.g. stem-affix relations) is a byproduct of the ways in which words are paradigmatically related within and across inflectional paradigms (Ackerman et al., 2016; Blevins, 2016).

In the inter-paradigmatic direction, a central question has to do with how inflected forms cue inflection class membership – the so-called Paradigm Cell Filling Problem (Ackerman et al., 2009). Table 1 illustrates the issue using a subset of the inflected forms of Russian nouns. (For the moment I assume a typical, four-class description of Russian nouns, although I will ultimately employ a more robust representation in Sections 5 and 6.) In Russian, the accusative singular exponent *-u* (as in *knig-u* 'book-ACC.SG') is fully informative about inflection class membership, which is to say, about what the other forms of the same lexeme are. If a competent adult speaker encounters a neologism ending in *-u* and knows that it is accusative singular, all other forms of the noun are predictable (ignoring stress placement). However, inflected forms are not guaranteed to be fully (or at all) informative in this way. Instrumental singular *-om* is partially informative: the new word must belong to either the STOL class or the MESTO class, but the observed form does not resolve which. The dative plural exponent *-am* is uninformative, since it appears in every inflection class. The distributions of inflected forms across classes thus determine how and the extent to which allomorphs cue inflection class membership. They likewise define a pattern of relatedness among lexemes, and by extension inflection classes, and reflect the internal organization of the inflectional system.

This internal organization has been of particular interest in work that seeks to quantify inflectional complexity. From an I-complexity perspective, the Paradigm Cell Filling Problem is a significant issue because neither child (Lignos and Yang, 2016) nor adult (Bonami and Beniamine, 2016) speech input is sufficient to observe all inflected forms of all lexemes. Speakers must therefore be able to productively predict and generate unobserved inflected forms. The complexity of an inflectional system is a function of the difficulty of this task, given some partial knowledge of a lexeme (Stump and Finkel, 2013).

Estimates of the I-complexity of inflectional systems based on paradigmatic relations – essentially, proportional analogy – have been calculated in set-theoretic (Stump and Finkel, 2013) and information-theoretic terms (Ackerman et al., 2009; Ackerman and Malouf, 2013; Bonami and Beniamine, 2016; Mansfield, 2016; Parker and Sims, to appear; Sims and Parker, 2016; Stump and Finkel, 2013). Sequence-to-sequence neural network models for inflection have also been employed (Cotterell et al., to appear; Malouf, 2017). Using conditional entropy, Parker (2016) estimates the complexity of the Russian nominal system at between 0.5 and 0.6 bits, depending on how much detail about Russian inflectional outcomes is included in the analysis.

This notion of inflectional complexity has also

been extended to cross-linguistic comparison. Ackerman and Malouf (2013)[436] propose the Low Entropy Conjecture: "...enumerative morphological complexity is effectively unrestricted, as long as the average conditional entropy, a measure of integrative complexity, is low..." The Low Entropy Conjecture is posited to be a universal constraint on morphological I-complexity, driven by speakers' need to be able to solve the Paradigm Cell Filling Problem. Other work has suggested a trade-off between I-complexity and E-complexity (Cotterell et al., to appear). Importantly, however, both suggest that I-complexity reveals commonalities among languages' inflectional systems that are not captured by typological approaches focused on E-complexity.

As a basis for cross-linguistic comparison, the notion of I-complexity thus reflects something different about morphological structure than traditional measures do. It is also inextricably rooted in the internal organization of inflectional systems – in particular, the distribution of allomorphs across lexemes and classes. Yet tools for directly examining this organization are lacking.[2] Previous work largely boils the distributional properties of an inflectional system down to an estimate of its complexity as a whole (as with Parker's estimate for Russian nouns). While this is appropriate to some goals, single-value measures have the same problem found with all averages: many different distributions can produce the same average. As a basis for comparison across languages this offers an incomplete picture of the extent to which languages are similar or different (Elsner et al., submitted). Moreover, languages seem to differ in the extent to which paradigmatic relations (proportional analogy) are important to maintaining low I-complexity (Sims and Parker, 2016), suggesting the need to directly investigate a system's organization, and not only its resulting complexity.

These issues highlight the need to drill down on the distributional properties of individual morphological elements. Tools are needed for the description of individual systems at that level that offer a basis for meaningful cross-linguistic comparison.

3 Inflectional systems as networks

I define an inflection class system as an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where the set $\mathcal{V}$ of nodes consists of the inflection classes of the language and the set $\mathcal{E}$ of edges consists of unordered pairs of elements in $\mathcal{V}$. In particular, elements in $\mathcal{E}$ are defined by exponence shared among pairs of elements in $\mathcal{V}$. Taking the partial set of inflected forms from Table 1 as a simplified example, there are four inflection classes (thus, $\mathcal{V}(\mathcal{G})$ = {STOL, MESTO, KNIGA, KOST'}). The classes are distinct overall, but all four have the exponent *-am* in dative plural, the classes of STOL and KOST' both lack an overt accusative singular exponent, and STOL and MESTO both have *-om* in instrumental singular. These overlaps define six edges $\mathcal{E}(\mathcal{G})$ = {STOL-MESTO, STOL-KOST', STOL-KNIGA, MESTO-KOST', MESTO-KNIGA, KOST'-KNIGA}, as visualized in Figure 1.[3]

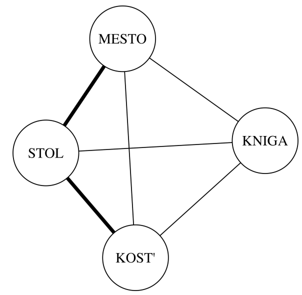

Figure 1: Network graph of the partial set of Russian noun forms shown in Table 1

Furthermore, the weight of an edge is defined as the number of cells in which two classes overlap. This is shown as a heavier line for the edges connecting nodes STOL and MESTO, and STOL and KOST'. Edge weight captures the observation that classes that overlap in more cells are more similar to each other. In language change, these are more likely to analogically influence each other. Edges can thus be thought of as paths of analogical reasoning— more specifically, the edges represent potential pivots for inflection class shift.

4 Segmentation and the definition of classes

The number of inflection classes a given language is analyzed as having is predicated on a segmentation of its words into stems and exponents. Mor-

[2] However, Beniamine (2018) is notable for the use of network visualization.

[3] All network graphs in this paper were plotted with the igraph package (Csardi and Nepusz, 2006) in R (R Core Team, 2019). This package was also used to calculate clustering coefficient, shortest path length, and betweenness centrality, as described in Section 5 below.

phological segmentation has long presented analytic challenges for description and typology (Beniamine et al., 2017a; Hockett, 1947; Nida, 1949), formal theory (Matthews, 1972; Spencer, 2012), and computational modeling (Goldsmith, 2001, 2010; Harris, 1970; Manning, 1998). Encoder-decoder neural models of inflection (Faruqui et al., 2016; Kann and Schütze, 2016; Malouf, 2017; Silfverberg and Hulden, 2018) have recently become popular in part because they are able to sidestep questions of how words should be segmented into morphological units and how to define discrete inflection classes. However, it is difficult to identify and interpret the latent representations that neural network models of inflection actually learn. The analyses below are instead based on manual segmentation, which has the advantage of being maximally linguistically interpretable.[4]

In what follows I use a global segmentation strategy (Beniamine et al., 2017b), in which the 'stem' is the maximal continuous string shared by all inflected forms of a lexeme. There are two exceptions to this principle: 1) Suprasegmental material (e.g. tone) is analyzed separately from segmental material, allowing globally shared segmental material to be identified as part of the stem, even when suprasegmental material is different from one inflected form to another. Suprasegmental material that is not shared by all inflected forms of a lexeme is assigned to the exponent. 2) Purely automatic phonology (e.g. of the type that is vowel harmony in Turkish, or vowel reduction in Russian) is ignored. This method results in bits of form that linguists often classify as stem allomorphy (morphophonological alternations, stem extensions, theme vowels, stress shift, etc.) being assigned to the exponent.[5]

Once a segmentation into stem and exponent is made, defining classes is a trivial matter: two words belong to the same inflection class if and only if the full sets of their exponents are identical. This method results in microclasses in the terminology of Beniamine et al. (2017b), which tend to be large in number, relative to classical descriptions. For example, descriptions of the Russian nominal system tend to posit either three (Vinogradov et al., 1952) or four (Corbett, 1982) (macro)classes, whereas the method used here produces 87 (micro)classes.[6]

Since this is a somewhat unusual analytic choice, it requires some justification. In defining inflection classes, linguists tend to abstract away from morphophonological alternations, especially if phonologically conditioned, preferring to define classes based (solely, ideally) on lexically-conditioned, suppletive exponents. This minimizes the number of inflection classes posited. However, there are at least four reasons to adopt a maximally inclusive definition of exponents, and a more robust number of classes.

First, returning to the Paradigm Cell Filling Problem and the notion of I-complexity, to 'solve' the PCFP speakers must predict entire word-forms. Limiting what counts as an exponent may lead to overestimation or underestimation of the I-complexity of inflectional systems (Elsner et al., submitted; Sims, 2015). This is important because the graph-theoretic approach to inflectional typology argued for in this paper is motivated exactly by a desire to better understand how I-complexity relates to the internal organization of inflectional systems, and the extent of cross-linguistic diversity in this respect.

Second, the line between morphology and phonology cannot always be drawn in a principled and pre-theoretic way. The choice to define exponents in a maximally inclusive way is not theory-neutral, to be sure – it is philosophically aligned with the Word and Paradigm framework. But to the extent that it errs, it does so consistently on the side of representing inflection classes as overly distinct. This is preferable to erring in the opposite direction because we can ask about the extent to which microclasses group into macroclasses, but if we abstract away from morphological differences and thus fail to distinguish two classes in the first place, we will never be able to detect any inter-

[4] A goal for the future is to expand the methods and code to include automatic segmentation of words into stems and exponents, e.g. through integration with the Qumin software package (Beniamine, 2018): https://github.com/XachaB/Qumin

[5] Multiple exponents are treated as a single, combined exponent. To the extent that each of multiple exponents has a separate distribution, an analysis in terms of multilayer networks (Bianchoni, 2018) would likely be needed to capture this. Multilayer network representations are more complex and I leave this extension for the future.

[6] As a reviewer observed, suppletive material is all assigned to the exponent, resulting in maximal differentiation from other classes and potentially increasing not only the number of classes, but the prevalence of disconnected subgraphs. Indeed, exactly this situation is encountered in Russian nouns (see Section 5), showing that segmentation choices affect the representation of the network to some degree. However, it is not clear that there is a 'right' or 'wrong' choice in this respect.

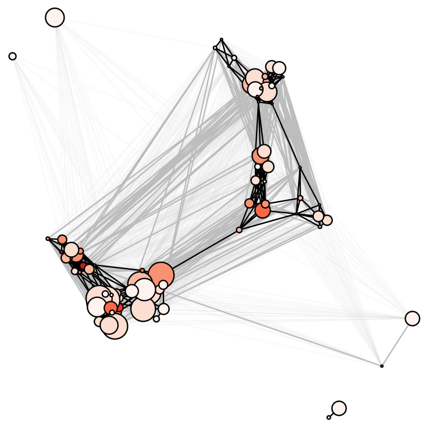

Figure 2: Inflection class system of Russian nouns (87 classes). Nodes size represents the log type frequency of the class. Node color reflects betweenness centrality (darker = more central). Edge color and thickness are according to weight: edges connecting nodes (classes) with the same exponents in more than half of cells are black (N $\geq$ 7); edges connecting nodes with the same exponents in exactly half of cells (N=6) are thick gray; weaker edges are thin gray.

esting aspects of inflectional organization that the abstracted-away-from differences constitute.

Third, as a practical matter, a global segmentation strategy can be applied in a uniform way across languages and requires a minimum of analytic/theoretical assumptions (Beniamine et al., 2017b), evading potential problems created by the use of different analytic methods for different languages.

Finally, and perhaps most importantly, different kinds of allomorphy tend to be found in different types of morphological systems (e.g. agglutinative vs. fusional) (Plank, 1999). Including some kinds of allomorphy and excluding others thus runs the risk of introducing systematic bias into cross-linguistic comparisons of inflection class organization.

In the following section I illustrate how standard measures for network description can be used to quantify the organizational structure of the Russian nominal inflectional system.

5 Network properties of Russian nouns

The inflection class network for Russian nouns is shown in Figure 2. Following Parker (2016), the underlying morphological analysis includes not

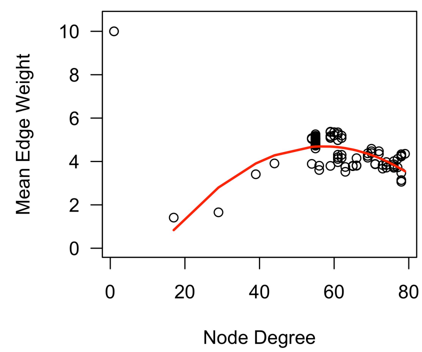

Figure 3: Correlation between node degree and mean edge weight for Russian nouns. The red line shows a quadratic regression fit.

just regular and productive inflectional suffixes, but also irregular suffixes, stress alternations, stem extensions, defectiveness (no inflected form for a given paradigm cell), and uninflectedness (only one form for all paradigm cells). Node size reflects the log type frequency of the class (i.e. the log number of lexemes it contains), based on 43,486 nouns in Zaliznjak (1977). Node color indicates betweenness centrality, discussed below. Edges are colored according to their weight.

5.1 Number of nodes, edges, and connected components

Basic descriptive statistics for the Russian nominal inflectional network include the number of its nodes ($|\mathcal{V}(\mathcal{G})| = 87$), the number of its edges ($|\mathcal{E}(\mathcal{G})| = 2660$), and how many connected components it has. A connected component is a subgraph containing all of the nodes that are connected via a path. The Russian noun system has two components. One has two nodes that differ from each other only in accusative (the result of animacy-conditioned allomorphy), exemplified by REBËNOK 'child, baby' (NOM.PL *rebjata*), which has a unique suppletive stem alternation *-onOk* $\sim$ *-at*.[7] The remaining 85 classes belong to the other connected component.

5.2 Degree distribution and edge weight

Node degree is the number of edges $\mathcal{K}$ that are connected to a node. In Russian, the large majority of classes have $|\mathcal{K}| > 50$.

[7] Capital O in -onOk indicates a fleeting vowel.

The relationship between node degree and edge weight is shown in Figure 3.[8] The quadratic nature of the distribution (R^2 = 0.55, $p < 0.0001$) probably partly reflects limitations on the extent to which classes can overlap but remain distinct. Classes with both high degree and high edge weight are likely targets for merger, which may explain the relative lack of such classes in Russian nouns. However, interestingly, there is no such restriction for low degree nodes, for which it is entirely possible to overlap with few other classes (low degree), but in many cells (high edge weight). The ways in which Russian nouns overlap thus do not appear to reflect random sampling from the full space of possibilities.[9]

5.3 Clustering coefficient

As is evident visually in Figure 2, Russian inflection classes form clusters: groups of nodes with high-density ties. This clustering is why Russian is typically described as having three of four classes: there are few general inflectional patterns, but many words with small deviations from these.

Clustering demonstrates one reason why node connectivity patterns affect system complexity. On the one hand, classes with high-density ties interfere with each other analogically. It might therefore seem that a greater density of edges in a network would lead monotonically to greater system complexity. However, when classes cluster, the interfering classes have mostly the same exponence. Strong clustering can thus actually lead to good interpredictability of forms for the majority of cells, even in a strongly connected network. It turns out there is no uniform relationship between the number of edges in a graph (or their weight) and the complexity of an inflectional system (Parker and Sims, to appear). This makes clustering an important network property for cross-linguistic comparison.

In an undirected network, the local clustering coefficient C_i of a node v_i with k neighbors is defined as:

$$C_i = \frac{2|\{e_{jk} : v_j, v_k \in N_i, e_{jk} \in E\}|}{k_i(k_i - 1)}$$

where N_i is the neighborhood of v_i, specifically, the set of nodes to which v_i is directly connected by an edge. The local clustering coefficient of v_i is thus the total number of edges among v_i's neighbors, divided by the total possible number of edges among neighbors. The global clustering coefficient of a system is the mean calculated over all C_i; values range between 0 and 1. The Russian nominal network has a global clustering coefficient of 0.816 (s.d. = 0.147).

5.4 Mean shortest path length

The path length between two nodes is the number of edges that must be followed to get from one to the other. Path length, like clustering coefficient, thus reflects patterns of network connectivity. Since edges in the inflectional network represent paths of analogical reasoning, the length of a path between a pair of nodes can be interpreted as being related to the likelihood of analogical interference between those classes, with low numbers indicating greater potential interference.

Since the Russian nominal network is not fully connected, the mean shortest path length for Russian nouns is here calculated within component. (Across components there are no paths, so shortest path length is infinite.) When calculated without edge weight (using a breadth-first search algorithm), the Russian network has a mean shortest path length of 1.249 (s.d. = 0.134) and when calculated taking edge weight into account (using the Dijkstra algorithm), the mean shortest path length is 8.929 (s.d. = 1.42).[10]

5.5 Betweenness centrality

We might also want to know which nodes are most central in the network. Central nodes are ones that are most likely to have shortest paths traverse them, often by virtue of them being connected to maximally separate parts of the network. As such, they are classes that are disproportionately likely to create pivots among classes that are more distinct, relative to other nodes in the net-

[8] The regression line excludes two nodes with degree of 1 and edge weight of 10. These are the same two nodes that belong to a separate component. If these are instead analyzed as a single class with a cross-cutting paradigm condition (Baerman et al., 2017), the merged class has degree of 0.

[9] Although there is not space in this paper to dive further into this issue, other languages show different degree-to-weight distributions.

[10] Shortest path length calculated over weighted edges seeks to minimize edge weight, treating edge weight as distance or cost. In the Russian nominal network, however, edge weight reflects similarity: more similar classes are connected by heavier edges. This would, oddly, result in the algorithm finding paths through maximally dissimilar classes. Edge weights were thus reversed for calculations of path length. Since Russian nouns have 12 cells, the maximum possible edge weight is 11. An edge weight of 11 was transformed to a value of 1, 10 was transformed to 2, etc.

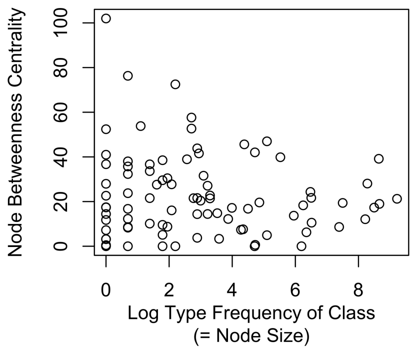

Figure 4: Correlation between node size and betweenness centrality for Russian nouns

work, putting those classes' exponents into potential analogical competition.

Betweenness centrality is calculated based on the set of shortest paths between v_i and v_j, for all possible values of i and j (where $i \neq j$). The betweenness centrality of a node v_k is the number of shortest paths in that set that include v_k, where $k \neq i, j$. In Figure 2 nodes are colored according to their betweenness centrality value, with darker red indicating more centrality. Figure 4 shows the betweenness centrality of classes as a function of their log type frequency.

Notice that low type frequency noun classes in Russian may be either high or low in centrality, but high type frequency classes have only low centrality. The nodes with the highest betweenness centrality turn out to be ones that are *mostly* regular but have irregularities that cross-cut the conventional classes in one or a few cells in the paradigm (especially, stress shift, vowel-zero alternation,[11] or an irregular nominative plural). Classes with the lowest betweenness centrality may also have low type frequency and exhibit irregularity, but in a different way: they are either uninflected or have unique stem extensions that serve to differentiate them from most other classes in most cells. Betweenness centrality thus reveals two different kinds of irregularity in Russian nouns, with different connectivity profiles within the network.

The distribution in Figure 4 is consistent with the observation by Sims and Parker (2016) that low type frequency classes contribute disproportionately to the unpredictability (complexity) of the Russian nominal system; Stump and Finkel (2013) make a similar generalization based primarily on Icelandic verbs. However, it is seems likely that the true underlying issue has to do with how classes are embedded in their network – the effect is driven by classes with high betweenness centrality, which are themselves likely to have low type frequency.

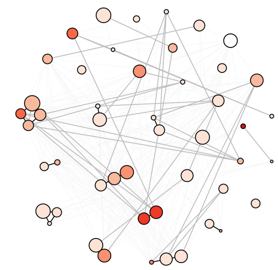

Figure 5: Inflection class system of Greek nouns

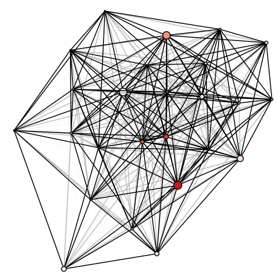

Figure 6: Inflection class system of Nuer nouns

6 Cross-linguistic comparison

I now turn to look at how these network measures might be used as a basis for typological comparison. Table 2 gives summary information

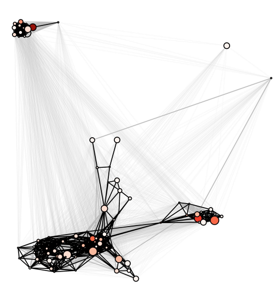

Figure 7: Inflection class system of Palantla Chinantec verbs

[11] E.g. NOM.SG *otec* 'father', GEN.SG *otc-a*.

Language	Family	Cells	Classes	Lexemes	Sources
Chinantec verbs	Oto-Manguean	24	101	838	(Merrifield and Anderson, 2007)
French verbs	Indo-European	49	65	6,485	(Stump and Finkel, 2013)
Greek nouns	Indo-European	6	48	25,370	(Sims, 2015; Idryma Manoli Triantafyllidi, 1998)
Icelandic verbs	Indo-European	30	146	1,034	(Stump and Finkel, 2013; Jörg, 1989)
Kadiwéu verbs	Mataco-Guaicura	5	57	364	(Baerman et al., 2015; Griffiths, 2002)
Nuer nouns	Nilotic	6	25	252	(Baerman, 2012)
Russian nouns	Indo-European	12	87	43,486	(Parker, 2016; Zaliznjak, 1977)
Seri verbs	Isolate	4	254	952	(Baerman, 2016; Moser and Marlett, 2010)
Võro verbs	Uralic	9	23	4,668	(Baerman, 2014; Iva, 2007)

Table 2: Summary properties of the languages under investigation. Where more than one data sources is listed, the first is the direct source; the second is the original source

and sources for nine inflectional systems investigated here: Palantla Chinantec verbs, French verbs, Greek nouns, Icelandic verbs, Kadiwéu verbs, Nuer nouns, Russian nouns, Seri nouns, and Võro verbs. See Sims and Parker (2016) for further information about these data sets. This represents an opportunistic sample; it is not genetically or geographically balanced. This section focuses on comparing mean shortest path length and global clustering coefficient across these languages. A comparison based on the other metrics is left to future work for reasons of space, but the example is illustrative of how graph-theoretic measures can lead to new generalizations about the typological space of morphological systems.

Impressionistically, the diversity of the nine languages is striking. In addition to differing substantially in how many paradigm cells and classes they have, Figures 5 through 7 show the inflectional networks for Greek, Nuer, and Palantla Chinantec. The Greek nouns are connected by relatively fewer and weaker edges whereas the Nuer nouns are robustly connected. Additionally, nodes clusters into distinct groups in Palantla Chinantec, like in Russian.

Interestingly, however, when we turn to measures of shortest path length and clustering coefficient, an emergent pattern is evident. For shortest path length and clustering coefficient, direct comparison across languages is not meaningful because the sizes of the inflectional systems (number of nodes and edges) differ. More meaningful is a comparison between the inflectional systems and randomized versions of those systems. Simu-

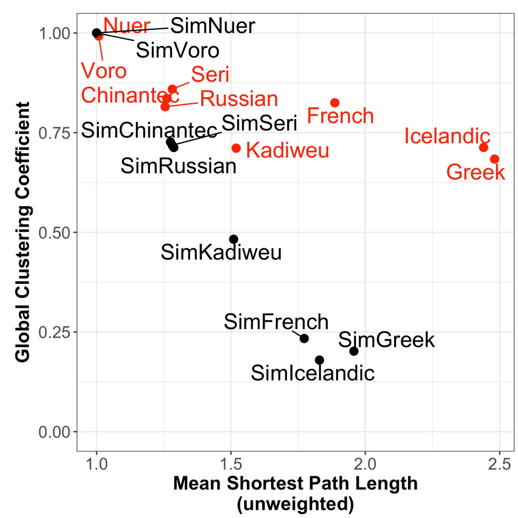

Figure 8: Comparison of real and simulated (resampled) inflection class systems according to mean shortest path length and global clustering coefficient

lated languages were generated by randomly sampling with replacement from the set of exponents for each paradigm cell, assigning them to classes. The exponents for each paradigm cell were sampled separately. The resulting simulated systems have the same number of allomorphs and classes as the real systems, but the paradigmatic relations that define the internal organization of the system have been randomly shuffled.

The results are shown in Figure 8.[12] (For the simulated languages, mean values from 100 ran-

[12] A version based on weighted edges, in which the distribution of weights from each real language was sampled with replacement and assigned at random to edges, produced qualitatively similar results.

domizations are shown.) The real systems differ from the simulated systems primarily in clustering, with the real languages exhibiting relatively more clustering as path length increases. Notably, for Nuer and Võro there is no meaningful difference between the real and simulated versions in either clustering or path length. This is equivalent to saying that Nuer and Võro lack (non-random) inflection class structure.

The closer the mean shortest path length of a network is to a value of 1, the closer that network necessarily is to forming a single large cluster, since every node is directly connected to every other node. This is what we see in Nuer and Võro. In contrast, networks with relatively long average path length values are relatively sparsely populated with edges (compare Figure 5 to Figure 6). In inflectional terms, this translates to classes that are more distinct. This sparsity gives more opportunity for (non-random) clustering. At the same time, it is not true that these networks *must* cluster to a significant degree, as the divergence between the real and the simulated languages shows.

The fact that in many languages, microclasses can be grouped into successively larger macroclasses is not a new observation (Brown and Hippisley, 2012; Dressler et al., 2006), but the generalization that some types of languages (i.e. ones whose networks are relatively sparsely populated with edges) are more likely to have this property is a new typological observation. But *why* do languages with greater average path length also employ significant amounts of clustering? Here it is not possible to do more than speculate in a broad way, but one possibility is that inflection classes that are more distinct are more likely to fracture over time as a result of independent changes (e.g. sound change), leaving groups of closely related but not identical classes. When classes are more distinct to begin with, such changes may be more likely to result in clustering. Further work would be needed to examine this possibility. But whatever the reason for the emergent pattern in Figure 8, it shows the ability of graph-theoretic measures, when applied to inflectional typology, to unearth new empirical generalizations about the internal organization of inflectional systems.

7 Conclusions

While traditional approaches to inflectional typology have focused on the size of inflectional systems, this does not capture their internal organization, particularly as related to the predictability of inflected forms (also called the system's I-complexity). I have argued for thinking of inflectional systems as networks in which the nodes are classes and the edges are exponents that two classes have in common. This allows for tools from graph theory to be applied to the task of describing the internal organization of inflectional systems in their full richness.

The cross-linguistic comparison in section 6 highlighted the possibility of using graph-theoretic measures to compare the network structure of inflection class systems. The measures employed here offer a fundamentally different basis for typology than in traditional approaches and revealed novel generalizations about the typological space of morphological systems. In particular, clustering emerged as a common property.

Future work should focus on identifying which graph-theoretic measures are most useful for cross-linguistic comparison of morphological systems. Additionally, as has already been demonstrated in other domains (e.g. transportation networks), node connectivity profiles not only define classes of networks, but affect the dynamics of a network differently (Guimerà et al., 2007). This hints at the possibility of better predicting inflectional change. Ultimately, graph theory offers a promising basis for inflectional typology, and more.

Acknowledgments

Preliminary versions of this work were presented at the Future of Language Science event (Northwestern University) and Linguistic Institute Sunday Poster Sessions (University of Kentucky), and to the Laboratoire de Linguistique Formelle (Université de Paris and CNRS). I thank the audiences in those places for their helpful feedback. Conversations with Jeff Parker have shaped my thinking and I thank him for his work compiling some of the data sets. I also thank Matthew Baerman, Raphael Finkel and Greg Stump for making data sets available. This paper has benefited from the comments from three anonymous reviewers; all errors remain my own.

References

Farrell Ackerman, James P. Blevins, and Robert Malouf. 2009. Parts and wholes: Implicative patterns

in inflectional paradigms. In James P. Blevins and Juliette Blevins, editors, *Analogy in grammar: Form and acquisition*, pages 54–82. Oxford University Press.

Farrell Ackerman and Robert Malouf. 2013. Morphological organization: The low conditional entropy conjecture. *Language*, 89(3):429–464.

Farrell Ackerman, Robert Malouf, and James P. Blevins. 2016. Patterns and discriminability in language analysis. *Word Structure*, 9(2):132–155.

Matthew Baerman. 2012. Paradigmatic chaos in Nuer. *Language*, 88(3):467–494.

Matthew Baerman. 2014. Covert systematicity in a distributionally complex system. *Journal of Linguistics*, 50(1):1–47.

Matthew Baerman. 2016. Seri verb classes: Morphosyntactic motivation and morphological autonomy. *Language*, 92(4):792–823.

Matthew Baerman, Dunstan Brown, and Greville G. Corbett. 2017. *Morphological complexity*. Cambridge University Press.

Matthew Baerman, Dunstan Brown, Roger Evans, Greville G. Corbett, and Lynne Cahill. 2015. Surrey morphological complexity database.

Sacha Beniamine. 2018. *Classifications flexionnelles. Étude quatitative des structures de paradigmes*. Ph.D. thesis, Linguistique, Université Sorbonne Paris Cité – Université Paris Diderot (Paris 7).

Sacha Beniamine, Olivier Bonami, and Joyce McDonough. 2017a. When segmentation helps: Implicative structure and morph boundaries in the Navajo verb. In *Abstracts from the First International Symposium of Morphology, University of Lille*, pages 11–15.

Sacha Beniamine, Olivier Bonami, and Benoît Sagôt. 2017b. Inferring inflection classes with description length. *Journal of Language Modelling*, 5(3):465–525.

Ginestra Bianchoni. 2018. *Multilayer networks: Structure and function*. Oxford University Press.

James P. Blevins. 2006. Word-based morphology. *Journal of Linguistics*, 42:531–573.

James P. Blevins. 2016. *Word and paradigm morphology*. Oxford University Press.

Olivier Bonami and S. Beniamine. 2016. Joint predictiveness in inflectional paradigms. *Word Structure*, 9:156–182.

Dunstan Brown and Andrew Hippisley. 2012. *Network Morphology: A defaults-based theory of word structure*. Cambridge University Press.

Bernard Comrie. 1981. *Language universals and linguistic typology: Syntax and morphology*. University of Chicago Press.

Greville G. Corbett. 1982. Gender in Russian: An account of gender specification and its relationship to declension. *Russian Linguistics*, 6:197–232.

Ryan Cotterell, Christo Kirov, Mans Hulden, and Jason Eisner. to appear. On the complexity and typology of inflectional morphological systems. *Transactions of the Association for Computational Linguistics*.

Gabor Csardi and Tamas Nepusz. 2006. The igraph software package for complex network research. *InterJournal*, Complex Systems:1695.

Wolfgang U. Dressler, Marianne Kilani-Schoch, Natalia Gagarina, Lina Pestal, and Markus Pöchtrager. 2006. On the typology of inflection class systems. *Folia Linguistica*, 40:51–74.

Micha Elsner, Andrea D. Sims, Alex Erdmann, Antonio Hernandez, Evan Jaffe, Lifeng Jin, Martha Johnson, Shuan Karim, David King, Luana Lamberti Nunes, Nathan Rasmussen, Cory Shain, Symon Stevens Guille, Stephanie Antetomaso, Kendra Dickinson, Noah Diewald, Michelle McKenzie, and Byung-Doh Oh. submitted. Models of morphological learning: Implications for typology and change.

Manaal Faruqui, Yulia Tsvetkov, Graham Neubig, and Chris Dyer. 2016. Morphological inflection generation using character sequence to sequence learning. In *Proc. of NAACL*.

John A. Goldsmith. 2001. Unsupervised learning of the morphology of a natural language. *Computational Linguistics*, 27:153–198.

John A. Goldsmith. 2010. Segmentation and morphology. In Alexander Clark, Chris Fox, and Shalom Lappin, editors, *The handbook of computational linguistics and natural language processing*, pages 364–393. Wiley-Blackwell.

Glyn Griffiths. 2002. *Dicionário da lingua Kadiwéu*. Summer Institute of Linguistics.

Roger Guimerà, Martha Sales-Pardo, and Luís A.N. Amaral. 2007. Classes of complex networks defined by role-to-role connectivity profiles. *Nature Physics*, 3:63–69.

Zellig S. Harris. 1970. From phoneme to morpheme. In Zellig S. Harris, editor, *Papers in structural and transformational linguistics*, pages 32–67. Springer.

Charles F. Hockett. 1947. Problems of morphemic analysis. *Language*, 23:321–343.

Idryma Manoli Triantafyllidi, editor. 1998. *Lexiko tis koinis neoellinikis*. Aristoteleio Panepistimio Thessalonikis, Institouto Neoellinikon Epoudon.

Sulev Iva. 2007. *Võru kirjakeele sõnamuutmissüsteem*. University of Tartu.

Christine Jörg. 1989. *Isländische Konjugationstabellen / Icelandic conjugation tables / Tableaux de conjugaison islandaise / Beygingatöflur islenskra sagna*. H. Buske.

Katharina Kann and Hinrich Schütze. 2016. MED: The LMU system for the SIGMORPHON 2016 shared task on morphological reinflection. *ACL 2016*, page 62.

Constantine Lignos and Charles Yang. 2016. Morphology and language acquisition. In Andrew Hippisley and Gregory T. Stump, editors, *The Cambridge handbook of morphology*, page 743764. Cambridge University Press.

Robert Malouf. 2017. Abstractive morphological learning with a recurrent neural network. *Morphology*, 27(4):431–458.

Christopher D. Manning. 1998. The segmentation problem in morphology learning. In D.M.W. Powers, editor, *Proceedings of NeMLaP3/CoNLL98 Workshop on Paradigms and Grounding in Language Learning*, pages 299–305. Association for Computational Linguistics.

John Mansfield. 2016. Intersecting formatives and inflectional predictability: How do speakers and learners predict the correct form of Murrinhpatha verbs? *Word Structure*, 9:183–214.

P.H. Matthews. 1972. *Inflectional morphology: A theoretical study based on aspects of Latin verb conjugation*. Cambridge University Press.

William R. Merrifield and Alfredo B. Anderson. 2007. *Diccionario Chinanteco de la diáspora del pueblo antiguo de San Pedro Tlatepuzco, Oaxaca*. Instituto Lingüístico de Verano.

Mary B. Moser and Stephen A. Marlett. 2010. *Comcaac quih yaza quih hand ihiip hav = Diccionario Seri-Español-Inglés*. Plaza y Valdés Editores y Universidad de Sonora.

Eugene A. Nida. 1949. *Morphology: The descriptive analysis of words*. University of Michigan Press.

Jeff Parker. 2016. *Inflectional complexity and cognitive processing: An experimental and corpus-based investigation of Russian*. Ph.D. thesis, Department of Slavic and East European Languages and Cultures, The Ohio State University.

Jeff Parker and Andrea D. Sims. to appear. Irregularity, paradigmatic layers, and the complexity of inflection class systems: A study of Russian nouns. In Peter Arkadiev and Francesco Gardani, editors, *Complexities of morphology*. Oxford University Press.

Frans Plank. 1999. Split morphology: How agglutination and flexion mix. *Linguistic Typology*, 3:279–340.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Miikka Silfverberg and Mans Hulden. 2018. An encoder-decoder approach to the paradigm cell filling problem. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2883–2889.

Andrea D. Sims. 2015. *Inflectional defectiveness*. Cambridge University Press.

Andrea D. Sims and Jeff Parker. 2016. How inflection class systems work: On the informativity of implicative structure. *Word Structure*, 9:215–239.

Andrew Spencer. 2012. Identifying stems. *Word Structure*, 5:88–108.

Gregory T. Stump and Raphael Finkel. 2013. *Morphological typology: From word to paradigm*. Cambridge University Press.

V.V. Vinogradov, E.S. Istrina, and S.G. Barxudarov. 1952. *Grammatika russkogo jazyka 1: Fonetika i morfologija*. Akademija Nauk SSSR.

Wolfgang U. Wurzel. 1989. *Inflectional morphology and naturalness*. Kluwer.

A.A. Zaliznjak. 1977. *Grammatičeskij slovar' russkogo jazyka: Slovoizmenenie*. Russkij jazyk.