

Li8: Morphology/Lent 2018

Abstraction and extrapolation

Jim Blevins (jpb39)
M11-12 / LB9 / 12.02

Intramorphological information

- ❖ The Segmentation Problem and the Interpretation Problem are both artifacts of a method of analysis, i.e., “[T]he theory creates a problem which it is then unable, or only partly able, to resolve.” (Matthews 1972: 74).
- ❖ It is futile to search for the ‘best’ trade-off in Latin or German, or to attempt to adjudicate between general subject and general 2pl markers in Georgian.
- ❖ The best strategy for avoiding these analytic traps is not to deny that words exhibit internal structure (à la ‘Whole Word Morphology’) or to treat ‘non-denotational’ variation as functionless (à la ‘Autonomous Morphology’).
- ❖ Instead, it is useful to recognize the way that form variation may express ‘intramorphological’ information by (i) **discriminating** words and larger units and by (ii) encoding **information about the shape of other forms**.

What about lossless decomposition?

- ❖ We want to guard against alarmist conclusions from isolated examples, and avoid a situation in which an outlier tail wags a statistically predominant dog.
- ❖ Regular plurals in English appear to be segmentable into stems and suffixes, and the suffixes do not show the variation that motivates inflection classes.
- ❖ The properties of the plural seem to be amenable to an analysis as the sum of the lexical properties of the stem and the number properties of the suffix /z/.
- ❖ The surface form of the suffix is conditioned by the final segment of the stem. For forms like *pans* and *foxes*, it appears possible to isolate the stems *pan* and *fox* and the suffix /z/, and reconstitute the original forms from these parts.

Is decomposition *ever* lossless?

- ❖ Even in this case, it is unclear that decomposition is lossless.
- ❖ Disassembling *pans* and *foxes* into stems and suffixes loses the information that the noun stems occur with the regular marker.
- ❖ Neither the grammatical properties of the nouns nor the form of the noun stems predict the fact that they occur with the regular suffix rather than following the patterns exhibited by *men* or *oxen*.
- ❖ ‘Morphologically conditioned allomorphy’ is invoked to govern the selection of a plural strategy, where this is not predictable from the properties or form of the stems *pan* and *man*, *fox* and *ox*.

Decomposition and frequency

- ❖ Given that few nouns follow residual plural patterns in English, a speaker can learn them as exceptions and assume that other nouns follow the regular pattern.
- ❖ However, this distinction brings in considerations of frequency, and frequency is another type of information that is sacrificed by a decompositional analysis.
- ❖ In particular, the frequency of a plural form is not generally recoverable from the frequency of the stem and the frequency of the suffix.
- ❖ Frequency information could be recovered from disassembled parts under certain conditions, notably if parts were not the 'same' element in all of their occurrences.
- ❖ Yet under these conditions, the link between recurrence and redundancy would be broken altogether, calling into question the assumptions and motivations underlying the strategy of disassembling and interpreting of recurrent units.

From systems to items

- ❖ The intuition behind atomistic decomposition is that recurrent structure is redundant and that the goal of morphological analysis is to eliminate redundancy by reducing systems to inventories of minimal units and combinatorial rules.
- ❖ Yet there is little evidence that morphological systems are genuinely organized in ways that facilitate disassembly and reassembly of words.
- ❖ It may be possible to disassemble a word form into minimal units.
- ❖ However there are often no effective criteria for selecting 'correct' segmentations into stable roots, stems and inflections. Moreover, in all but the simplest systems it is not possible to reconstitute word forms from inventories of minimal units, because disassembly has lost information about the distribution of units.
- ❖ Hence, atomistic accounts become preoccupied with the analysis of individual forms and never get back to analyzing the properties of the morphological system.

Deviant irregularity

- ❖ Atomistic treatments of stem alternations in Estonian (or ablaut in English) reveal a fundamental analytic bias.
- ❖ The goal of these analyses is to assimilate seemingly ‘irregular’ formations to more regular affixal patterns.
- ❖ This goal reflects the view that regular formations are ‘normative’, and that irregular forms are deviations from the uniform patterns that a system strives to maintain.

Discriminative irregularity

- ❖ A 'discriminative' perspective inverts this perspective.
- ❖ By enhancing the discriminability of forms, irregular forms aid effective communication. Ironically, suppletive patterns approach the ideal of a 'one form-one meaning' principle espoused by approaches such as Natural Morphology.
- ❖ From this perspective, regulars do not get a 'free ride' as normative patterns. Instead, regulars also serve a specific function by providing patterns that support generalization.

The herd immunity of irregulars

- ❖ Recognizing the function that irregular formations performing a system does not justify an inverted classification on which irregulars are treated as ‘communicatively optimal’.
- ❖ Irregulars can be seen to enjoy a ‘herd immunity’ within a system. The enhanced salience and discriminability of irregulars is offset by a commensurate reduction in generalizability.
- ❖ Given the Zipfian structure of the input, irregulars can only function in a system that is either (i) small enough to be acquired from directly encountered forms, or (ii) regular enough to allow speakers to deduce the shape of unencountered forms.

The regular-irregular equilibrium

- ❖ From a learning-based implicational perspective, neither regular nor irregular patterns are normative; they merely serve different, broadly complementary, functions in a system.
- ❖ The coexistence of patterns reflects an opposition between a pair of communicative pressures, one that enhances discriminability and another that promotes generalizability.
- ❖ Different languages can be expected to reach different states of equilibrium between these pressures, leading to different proportions of regular and irregular patterns.

The 'Paradigm Cell Alignment Problem (PCAP)'

- ❖ The observation that speakers never encounter all of the inflected forms of their language entails they must be able to solve the PCFP on the basis of the forms that they do directly encounter.
- ❖ Regularity clearly facilitates prediction from partial and biased samples.
- ❖ Yet in order for a collection of partial samples to guide the prediction of unattested forms, the forms that speakers do know must be organized in such a way that they collectively exhaust the inflectional variation.
- ❖ Speakers must recognize which sets of inflected forms can be 'pooled' to cover the variation that characterizes each 'class' of paradigms.

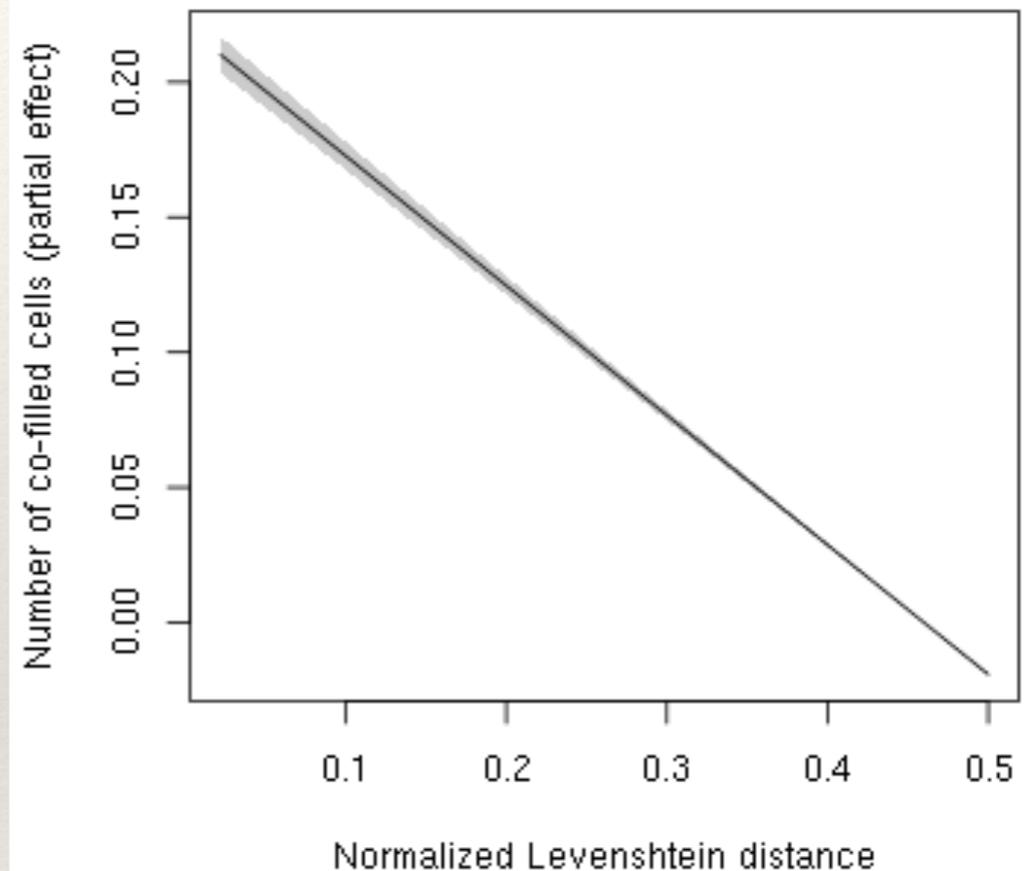
It takes a neighborhood?

in his analogizing ... [t]he native user of the language ... operates in terms of all sorts of internally stored paradigms, many of them doubtless only partial; and he may first encounter a new basic verb in any of its inflected forms. (Hockett 1967: 221)

- ❖ One possibility is that **lexical neighbourhoods** (Baayen et al. 2006; Gahl et al. 2011) provide an organization that allows a learner to extrapolate general inflectional patterns on the basis of local form similarities.
- ❖ An implicational model can address the PCAP if sparse paradigms that belong to a common 'class' exhibit sufficient cell overlap to provide an analogical base (via the antecedents of a proportion), and sufficiently diverse cell coverage to exhaust the inflectional patterns of a language.
- ❖ Neighbourhoods would permit the prediction of the full system from partial patterns via analogical extrapolation from a small set of nearest neighbors.

Neighborhood bootstrapping

- ❖ An initial study in Blevins Milin & Ramscar (2017) found that the form proximity of wordform pairs in German, gauged by normalized Levenshtein distance, predicts the number of co-filled paradigm cells, once effects due to the frequencies of the words are partialled out.
- ❖ What remains to be established is whether neighbours also exhibit sufficiently non-overlapping inflectional gaps to exhaust the variation within an inflectional class.
- ❖ Then studies of dense acquisition corpora can probe the learning trajectories associated with the individual informativity of words and collective informativity of neighbourhoods.



Partial effect of the normalized Levenshtein distance between pairs of nouns for the number of co-filled paradigm cells (Blevins, Milin & Ramscar 2017)

Why words?

- ❖ Words are the smallest free forms in a language, i.e., the smallest units that exhibit an independent distribution, can occur as utterances, (and FWIW characterize the initial production stage of learners).
- ❖ There are well-known discrepancies between what are sometimes termed ‘grammatical’ and ‘phonological’ words (Bloomfield 1914).
- ❖ But these discrepancies cannot be overcome by considering smaller units: there is no discrepancy between the ‘grammatical’ and ‘phonological’ morpheme because there is no notion of a ‘phonological’ morpheme.
- ❖ The word-level discrepancies might also be interpreted as evidence that the diagnostic properties adopted in linguistic descriptions are in fact just secondary cues that reinforce a fundamentally statistical structure.