

Symposium on Quantitative Morphology

Petar Milin, University of Sheffield
Jim Blevins, University of Cambridge

10th International Morphological Processing Meeting
SISSA, Trieste, 2017.06.23

Quantitative approaches to morphological processing

Over the past couple of decades, quantitative approaches have progressed from occupying a segregated niche within the field of linguistics to forming an integral component of the emerging multidisciplinary paradigms in the language sciences. It is particularly appropriate to take stock of this progression in the domain of morphology, where quantitative approaches have, so far, had the greatest impact. This symposium aims to summarize what has been achieved by quantitative approaches, draw some of the wide-ranging implications for the study of form variation in language, and state the challenges to future progress that this work has identified.

The emergence of a quantitative paradigm

The successful application of quantitative methods and resources in morphology reflects the influence of a number of interacting factors. Foremost among these is the fact that the definition of morphology as ‘the study of word structure’ provided multiple quantifiable dimensions of variation. Quantitative analyses could be applied not only to individual words, but also to the paradigmatic structures defined by sets of words and to the syntagmatic sequences of recurrent sub-word units.

Antecedent traditions of statistical stylistics (including stylometry and stylo-statistics) had established models that could be extended to morphological patterns in general. The dramatic growth in corpora, concordances and other lexical resources supplied large-scale data sources that supported robust statistical analysis. External validation came from experimental methodologies that were well-adapted to measuring behavioural correlates of the wide spectra of morphological variation.

Over the past half-dozen years, the results obtained across different domains of morphology have converged to yield a coherent perspective on morphological analysis. The information-theoretic measures initially applied to the analysis of response latencies [1–3] have proven equally useful for modelling variation and structure in morphological systems [4, 5]. The discriminative learning models that help to explicate the course and outcome of morphological acquisition [6, 7] also support integrated models of comprehension and production [8] that avoid the combinatorial explosion faced by earlier exemplar-based accounts [9, 10], and challenged models that presuppose meaning-blind, obligatory decomposition at the morphemic level [11]. The central role of a cognitively plausible learning rule provides a point of contact with the learning processes proposed in neurally-motivated models of lexical and morphological knowledge [12, 13].

New challenges and perspectives

Until the emergence of this mutually-reinforcing cluster of approaches, experimental models tended to exhibit an overly conservative perspective, particularly with regard to notions carried over from other morphological traditions. This conservatism is clearly reflected in psycholinguistic ‘processing’ models that retained a basic split between ‘parsing’ and ‘storage’ strategies. More generally, models tended to adopt a range of traditional idealizations that had no experimental motivation. Language was described in terms of ‘units’ at different linguistic ‘levels’ which were assumed to be ‘stored’ in the ‘mental lexicon’ of a speaker. Generative claims about the modular ‘architecture’ of the language faculty also erected a fire wall between experimental evidence and theory construction. On the assumption that a speaker’s abstract ‘competence’ was represented by an abstract formal grammar, experimental evidence could only bear indirectly on that knowledge, insofar as it was ‘implemented’ in a separate ‘performance’ (production and comprehension) system. Some early studies also show a somewhat uncritical attitude towards the use of corpus data and experimental paradigms, notably those involving priming effects.

A natural but radical possibility suggested by a quantitative perspective is that the building blocks of a familiar morphological description are essentially pedagogical idealizations. The idea of a ‘mental lexicon’ corresponds closely to a dictionary. The ‘formal grammars’ that encapsulate a speaker’s combinatoric knowledge are likewise abstract versions of a reference grammar. The units into which the continuous speech stream is taken to be segmented similarly reflect the conventions of written forms of alphabetic languages. A measure of support for this perspective comes from the recalcitrance of many of the analytic tasks defined in terms of traditional notions. Conversely, a unit-agnostic approach permits a flexible conception of form variation in which patterns of different types are most efficiently described

over domains of varying size, ranging from phonemic alternations, through patterns involving the abstraction of word-sized sequences and beyond.

The evolution of quantitative approaches to morphology has now reached the point where the status of these assumptions can be reassessed. It is clear that the implicit assumption that corpora can be ‘too big to fail’ stands in need of explicit justification. The use of experimental or learning paradigms that are far removed from naturalistic language use or acquisition must also be confirmed by or related to more ecologically valid paradigms. The relevance of traditional notions must be motivated and not simply built into the formulation of questions and tasks. Above all, synergies between theoretical insights, computational modelling methods and robust statistical techniques must be harnessed to reestablish morphology, on an interdisciplinary basis, as the general study of *Formenlehre* or ‘form variation’ in language.

Schedule

The symposium will consist of four presentations that examine different aspects of the issues raised above, particularly as they relate to morphological processing.

- | | |
|-------------|--|
| 9:00–9:05 | Opening remarks |
| 9:05-9:25 | <i>A reflection on psycholinguistics through its questions about morphology and the mental lexicon</i>
Emmanuel Keuleers (Tilburg University) |
| 9:25-9:45 | <i>The discriminative nature of morphology</i>
Michael Ramscar (Eberhard Karls Universität Tübingen) |
| 9:45-10:05 | <i>Storage vs processing in models of word inflection: A neuro-computational Hebbian perspective</i>
Vito Pirrelli (CNR, Pisa) |
| 10:05-10:25 | <i>Are baboons learning “orthographic” representations? Probably not</i>
Harald Baayen (Eberhard Karls Universität Tübingen/U of Alberta) |
| 10:25-10:30 | Closing remarks |

References

1. Kostić, A. Informational approach to processing inflectional morphology: Standard data reconsidered. *PSYCHOLOGICAL RESEARCH* 53, 62–70 (1991).
2. Moscoso del Prado Martín, F., Kostić, A. & Baayen, R. H. Putting the bits together: An information-theoretical perspective on morphological processing. *COGNITION* 94, 1–18 (2004).
3. Milin, P., Kuperman, V., Kostić, A. & Baayen, R. H. in *ANALOGY IN GRAMMAR: FORM AND ACQUISITION* (eds Blevins, J. P. & Blevins, J.) 214–253 (Oxford University Press, Oxford, 2009).
4. Ackerman, F. & Malouf, R. Morphological organization: The Low Conditional Entropy Conjecture. *LANGUAGE* 89, 429–464 (2013).
5. Blevins, J. P. *WORD AND PARADIGM MORPHOLOGY* (Oxford University Press, Oxford, 2016).
6. Ramscar, M., Yarlett, D., Dye, M., Denny, K. & Thorpe, K. The effects of feature-label-order and their implications for symbolic learning. *COGNITIVE SCIENCE* 34, 909–957 (2010).
7. Ramscar, M., Dye, M. & McCauley, S. M. Error and expectation in language learning: The curious absence of *mouses* in adult speech. *LANGUAGE* 89, 760–793 (2013).
8. Baayen, R. H., Milin, P., Filipović Đurđević, D., Hendrix, P. & Marelli, M. An Amorphous Model for Morphological Processing in Visual Comprehension Based on Naive Discriminative Learning. *PSYCHOLOGICAL REVIEW* 118, 438–481 (2011).
9. Baayen, R. H., Hendrix, P. & Ramscar, M. Sidestepping the combinatorial explosion: Towards a processing model based on discriminative learning. *LANGUAGE AND SPEECH* 56, 329–347 (2013).
10. Baayen, R. H., Shaoul, C., Willits, J. & Ramscar, M. Comprehension without segmentation: A proof of concept with naive discrimination learning. *LANGUAGE, COGNITION, AND NEUROSCIENCE* 32, 106–128 (2016).
11. Milin, P., Feldman, L. B., Ramscar, M., Hendrix, P. & Baayen, R. H. Discrimination in lexical decision. *PLOS ONE*, to appear (2017).
12. Chersi, F., Ferro, M., Pezzulo, G. & Pirrelli, V. Topological self-organization and prediction learning support both action and lexical chains in the brain. *TOPICS IN COGNITIVE SCIENCE* 6, 476–491 (2014).

13. Pirrelli, V., Ferro, M. & Marzi, C. in UNDERSTANDING AND MEASURING MORPHOLOGICAL COMPLEXITY (eds Baerman, M., Brown, D. & Corbett, G. G.) 141–166 (Oxford University Press, Oxford, 2015).

Abstracts

A reflection on psycholinguistics through its questions about morphology and the mental lexicon

Emmanuel Keuleers

Psycholinguistics has tended to formulate its fundamental questions according to the pattern “What is the psychological implementation of a linguistic notion?”, such as “What is the content of the mental lexicon?” or “How is morphology organized in the brain?”. I will argue that questions like these presuppose the lexical and morphological modularity of traditional linguistics. By definition, answering these questions through experimental research or computational modeling will by definition result in answers that preserve the traditional linguistic modularity. This self-perpetuation is further strengthened by psycholinguistics’ pervasive use of linguistic resources, which again reflects the traditional linguistic notions and linguistic modularity. I will discuss why these objections should not be confounded with objections to modularity or computability. Finally, I will explore how computational and experimental research questions can be phrased in terms of psychologically-grounded notions of language and communication.

Finding structure in form and time: discriminative explorations in quantitative morphology

Michael Ramscar

Traditional studies of language assume an atomistic model in which linguistic signals comprise discrete, minimal form elements associated with discrete, minimal elements of meaning. Since linguistic production has been seen to involve the composition of messages from an inventory of form elements, and linguistic comprehension the subsequent decomposition of these messages, researchers in linguistic morphology have focused on attempting to identify and classify these elements, along with the lossless processes of composition and decomposition they support, a program that has raised more questions than answers, especially when it comes to the nature of form-meaning associations.

By contrast, behavioral and neuroscience research based on human and animal models has revealed that “associative learning” is a lossy, discriminative process. Learners acquire predictive understandings of their environments through competitive mechanisms that tune systems of internal cue representations to eliminate or reduce any uncertainty they promote. Critically, models of this process better fit empirical data when these cue representations do not map discretely onto the aspects of the environment learners come to discriminate. In this talk, I will briefly describe the basic principles of learning, along with the empirical basis for the belief that human communication is subject to the constraints these principles impose, and describe how, from this perspective, languages should be seen as probabilistic communication systems that exhibit continuous variation within a multidimensional space of form-meaning contrasts.

This systematic picture of communication indicates that discrete descriptions of languages at an individual (psychological) or community (linguistic) level must necessarily be idealizations. Idealizations inevitably lose information, and I will then describe how the development of a discriminative, information theoretic approach to language both leads in turn to the appreciation of the vast array of socially evolved structure that serves to underpin human communication, and explains why the overly abstract models of language of the 20th inevitably led to this structure being ignored.

Are baboons learning “orthographic” representations? Probably not.

Harald Baayen, Maja Linke, Franziska Broecker and Michael Ramscar

The ability of Baboons (*Papio papio*) to distinguish between English words and non-words [1] has been modeled using a deep learning convolutional network model that simulates a ventral pathway in which lexical representations of different granularity develop [2]. However, given that pigeons (*Columba livia*), whose brain morphology is drastically different, can also be trained to distinguish between English words and nonwords [3], it appears that a less species-specific learning algorithm may be required to explain this behavior. Accordingly, we examined whether the learning model of [4], which has proved to be amazingly fruitful in understanding animal and human learning [5–7] could account for these data. We show that a discrimination learning network using gradient orientation features as input units and word and nonword units as outputs succeeds in predicting baboon lexical decision behavior – including key lexical similarity effects and the ups and downs in accuracy as learning unfolds – with surprising accuracy. The performance of this model, in which words are not explicitly represented, is remarkable because it is generally assumed that lexicality decisions, including the decisions made by baboons and pi-

geons [2, 3] are mediated by explicit lexical representations. Our results suggest that in learning to perform lexical decision tasks, baboons and pigeons do not construct a hierarchy of lexical units, but rather they make optimal use of low-level information obtained through the massively parallel processing of gradient orientation features. Accordingly, we suggest that skilled fluent reading of both simple and morphologically complex words in humans may involve a transition from a high-level system building on letter representations acquired during explicit instruction in literacy to the use of a similar strategy of exploiting massively parallel processing from low-level visual features to semantics. Methodological implications of these results for theories of lexical access will be discussed.

References

1. Grainger, J., Dufau, S., Montant, M., Ziegler, J. C. & Joël, F. Orthographic processing in baboons (*Papio papio*). *SCIENCE* **336**, 245–248 (2012).
2. Hannagan, T., Ziegler, J. C., Dufau, S., Joël, F. & Grainger, J. Deep learning of orthographic representations in baboons. *PLOS ONE* **9**, e84843 (2014).
3. Scarf, D. ET AL. Orthographic processing in pigeons (*Columba livia*). *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES* **113**, 11272–11276 (2016).
4. Rescorla, R. A. & Wagner, A. R. in *CLASSICAL CONDITIONING II* (eds Black, A. H. & Prokasy, W. F.) 64–99 (Appleton-Century-Crofts, New York, 1972).
5. Miller, R. R., Barnet, R. C. & Grahame, N. J. Assessment of the Rescorla-Wagner Model. *PSYCHOLOGICAL BULLETIN* **117**, 363–386 (1995).
6. Ramscar, M., Yarlett, D., Dye, M., Denny, K. & Thorpe, K. The effects of feature-label-order and their implications for symbolic learning. *COGNITIVE SCIENCE* **34**, 909–957 (2010).
7. Baayen, R. H., Milin, P., Filipović Đurđević, D., Hendrix, P. & Marelli, M. An Amorphous Model for Morphological Processing in Visual Comprehension Based on Naive Discriminative Learning. *PSYCHOLOGICAL REVIEW* **118**, 438–481 (2011).