For a Probabilistic and Multidisciplinary Approach to the Investigation of Morphological Processing

Davide Crepaldi[a], Marco Marelli[b,c] and Simona Amenta[d,e]

[a]International School for Advanced Studies (SISSA)

Via Bonomea 265, 34136, Trieste, Italy

davide.crepaldi@sissa.it

[b]University of Milano Bicocca

Piazza dell'Ateneo Nuovo 1, 20126, Milano, Italy

marco.marelli@unimib.it

[c]Milan Center for Neuroscience (NeuroMI), Italy

[d]University of Trento, Italy

Corso Bettini 31, 38068, Rovereto (TN), Italy

simona.amenta@unitn.it

[e]Ghent University, Belgium

Henri Dunantlaan 2, 9000 Gent, Belgium

This is an Editorial that introduces the Special Issue "Structure in words: The present and future of morphological processing from a multidisciplinary perspective", which was promoted in connection to the 10th International Morphological Processing conference (MoProc), held at SISSA, Trieste, Italy, June 22–24, 2017. The website for the conference is at https://indico.sissa.it/event/12/.

This Special Issue has received contributions from a number of labs, encompassing a vast range of methodologies and theoretical perspectives, and investigating diverse phenomena in different languages. Overall, 19 papers are gathered in the present volume. These contributions confirm a lively interest into morphology, and trigger a few considerations concerning the current status of the discipline, from both a theoretical and a methodological perspective.

On the theoretical side, an underlying theme seems to be the attempt to look at morphemes from a different perspective. Indeed, many papers move away, more or less subtly, from the typical characterization of morphemes in the psycholinguistic literature, and consider them in a wider diversity of contexts. Morphological processing is studied based on linguistically–informed notions (as in the experiment on zero-derivations by Wheedon et al.), considered in the context of wider issues such as literacy acquisition (as in the paper by Rastle), related to higher–level linguistic constructs (as in the work by Mancini et al., focusing on the interplay between syntax and discourse), and looked at from a learning perspective (as in the papers by Mirkovic et al. and Kimppa et al.). A certain degree of unease seems to emerge concerning the classic isolated–word reading approach focused on orthography and surface morphological complexity. In fact, this special issue presents a more "lively" representation of morphology, where the impact of morphemes is evaluated at the level of sentence and discourse (the work by Caffarra & Martin) or during spoken word processing (the papers by Bayersmann et al., Wilder et al., Oganyan et al., and Wang et al.).

It is also worth noting that this more general and diverse look on morphology is accompanied by an interest in the analysis of different languages: this special issue includes experiments on English, French, Italian, Mandarin, Basque, Spanish, Hebrew, Serbian, German, Finnish and an artificial language. Nearly 20 years later, it is nice that the community shows much sensitivity to cross–linguistic research, the core theme of the very first "MoProc" (Frost & Grainger, 2000). This is also nice with respect to the often too anglocentric character of psycholinguistic research (e.g., Frost, 2012). Even more interesting, most work on languages other than English is not simply motivated by cross-linguistic validation purposes. Rather, unique properties of the investigated languages are exploited to address theoretical issues that could not be otherwise evaluated (e.g., the association with different inflectional realizations of certain Italian quantifiers in the work by Arcara et al.; the affix subregularities of German in the study by Regel et al.).

On the methodological side, we note a marked multidisciplinarity in the approaches represented in the special issue: morphological processing is investigated by means of typical experimental psychology techniques, but also through cognitive neuroscience methods (as in Yablonski et al., and as discussed in the review paper by Leminen et al.) and computational modelling (e.g., Filipović-Đurđević & Milin). Multidisciplinarity even becomes the core topic in the work by Schmidtke & Kuperman, which draws attention to the apparent inconsistencies emerging when comparing different techniques (eye-movement data vs. ERPs). More generally, the choice of a given methodological approach is often related to clear theoretical purposes. Fitting examples in this respect are the papers employing computational methods. In these studies, the computational implementation is not only a way to precisely characterize and test pre–existing proposals, but actively informs theory by advancing, through simulations, novel interpretations of existing phenomena (e.g., the relationship between learning and information theory in the work by Filipović-Đurđević and Milin; cross-linguistic differences in morphological priming in the work by

Guenther et al.). A theoretical reason also motivates the adoption of ERPs, which characterize several articles in this special issue. The popularity of this methodology likely depends on its time resolution, which fits well with the importance that was assigned to timing in the past decade of morphological processing research: ERP data, in principle, can inform on when a certain morphological effect emerges during processing. Finally, the present volume includes a few examples of work on special populations. In this respect it is interesting to note that, differently from classical neuropsychological studies mostly focusing on aphasia and acquired dyslexia, these papers report on patients rarely considered in the morphological processing literature (e.g., Mild Cognitive Impairment, Dementia and Down syndrome; Penke, and Nikolaev et al.).

We think that this collection of papers shows us a possible way forward for the field. The multiplicity of research tools, paradigms, experimental populations and languages that were taken up here has unveiled several different aspects of the same phenomenon; it is only by trying to integrate all these different aspects that we can hope to gain a better insight into how our mind/brain processes complex words. Experimental Psychology and Cognitive Neuroscience are still in their infancy; we study extremely noisy phenomena, with tools that are rather narrowly scoped. EEG is fantastic for its time resolution, but yields very limited spatial information. fMRI has complementary features, and its poor temporal precision hampers its contribution in a theoretical landscape where timing is critical. Response times, especially if analysed with novel approaches, can be extremely informative; but they don't tell much about how the morphological processing is reflected at the neural level. And it is not only about techniques; for example, isolated–word paradigms are fundamental to keep proper experimental control on fine–grained aspects of morphological processing, but at the same time have limited ecological validity, since in natural context individual words are always embedded in sentences, discourse and, more in general, communication. It is clear that no comprehensive understanding could ever come from only one of

these techniques, paradigms or experimental approaches—the papers collected in this Special Issue sends this message loud and clear.

We would add one final note, which perhaps emerges less clearly from the papers collected here, but is of fundamental importance, we think, to interpret the recent past of the field, and drive its future. It is rather shared opinion in the community that we didn't advance much in our understanding of how the mind/brain processes complex words in the last 20 years. New models were proposed (e.g., Baayen et al., 2011; Crepaldi et al., 2010; Grainger and Beyersmann, 2017; Grainger and Ziegler, 2011; Taft and Nguyen-Hoan, 2010; Marelli and Baroni, 2015), but none of them has clearly gained wider acceptance than the others, nor they, collectively and incrementally, built a clear pathway as a field. Experiments have been piling up nicely (e.g., Amenta and Crepaldi, 2012; Lemimen et al., this issue; Marelli, Burani and Traficante, in press), but this improved knowledge of experimental facts didn't fully translate into a clear theoretical advancement. We suspect that part of the reason behind this state of affairs is the still dominant view of morphemes as discrete units that undergo combinatorial rules; a discrete, black–and–white, compositional morphology. Many experimental facts we are called to understand defy such a discrete approach—data quite often depict blurred boundaries, seem to lie on a continuum, and are probabilistic in nature (e.g., Baayen et al., 2016; Davis and Rastle, 2009; Marelli et al., 2015; Smolka et al., 2014; Ulicheva et al., 2018). We would suggest that a more probabilistic, blurred, shades–of–grey view of morphology is what these data call for. There are surely cases where a morphemes–and–rules approach works nicely; but perhaps, rather than being considered as the standard for morphological description, these situations should be seen as special cases in a more general framework where form and meaning entertain a wide network of non–deterministic regularity, which the mind/brain captures probabilistically. We contend that such an approach would allow a better account for the experimental data, thus promoting a steeper theoretical

advancement (e.g., Baayen et al., 2019). Moreover, it would connect morphology to the more general field of learning and statistical information processing (e.g., Ramscar et al., 2018), thus characterizing it as a privileged window onto one of the core mechanisms in the cognitive system, rather than as a niche within the fields of word identification and language processing.

References

Amenta, S. and Crepaldi, D. (2012). Morphological processing as we know it: An analytical review of morphological effects in visual word identification. *Frontiers in Language Science*, 3, 232.

Baayen, R. H., Milin, P., Filipovic Durdevic, D., Hendrix, P., and Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review*, 118, 438-482.

Baayen, R. H., Shaoul, C., Willits, J., & Ramscar, M. (2016). Comprehension without segmentation: A proof of concept with naive discriminative learning. *Language, cognition and neuroscience,* 31(1), 106-128.

Baayen, R. H., Chuang, Y. Y., Shafaei-Bajestan, E., & Blevins, J. P. (2019). The discriminative lexicon: A unified computational model for the lexicon and lexical processing in comprehension and production grounded not in (de) composition but in linear discriminative learning. *Complexity*, 2019.

Crepaldi, D., Rastle, K., Coltheart, M. and Nickels, L. (2010). `Fell' primes `fall', but does `bell' prime `ball'? Masked priming with irregularly–inflected primes. *Journal of Memory and Language*, 63, 83-99.

Davis, M. and Rastle, K. (2009). Form and meaning in early morphological processing: Comment on Feldman, O'Connor, and Moscoso del Prado Martín (2009). *Psychonomic Bulletin and Review*, 17, 749-755.

Frost, R. (2012). Towards a universal model of reading. *Behavioral and Brain Sciences*, 35, 263-279.

Frost, R. and Grainger, J. (2000). Cross–linguistic perspectives on morphological processing: An introduction. *Language and Cognitive Processes*, 15, 321–328.

Grainger, J. and Ziegler, J.C. (2011). A dual-route approach to orthographic processing. *Frontiers in Psychology*, 2, 54.

Grainger, J. and Beyersmann, E. (2017). Edge-aligned embedded word activation initiates morpho-orthographic segmentation. *Psychology of Learning and Motivation*, 67, 285-317.

Marelli, M., Amenta, S. and Crepaldi, D. (2015). Semantic transparency in free stems: the effect of Orthography–Semantics Consistency in word recognition. *Quarterly Journal of Experimental Psychology*, 68, 1571–1583.

Marelli, M., & Baroni, M. (2015). Affixation in semantic space: Modeling morpheme meanings with compositional distributional semantics. *Psychological Review*, 122(3), 485–515.

Marelli, M., Traficante, D., and Burani, C. (in press). Reading Morphologically Complex Words: Experimental Evidence and Learning Models. In Pirrelli, V., Plag, I., and Dressler, W. U. (eds.) *Word Knowledge and Word Usage: a Cross-disciplinary Guide to the Mental Lexicon*. Berlin: De Gruyter.

Ramscar, M., Dye, M., Blevins, J., & Baayen, H. (2018). Morphological development. In Bar On, A. & Rabvit, D. (eds.), *Handbook of Communications Disorders: Theoretical, Empirical, and Applied Linguistic Perspectives*. Mouton de Gruyter.

Smolka, E., Preller, K. H., & Eulitz, C. (2014). 'Verstehen'('understand') primes 'stehen'('stand'): Morphological structure overrides semantic compositionality in the lexical representation of German complex verbs. *Journal of Memory and Language*, 72, 16-36.

Taft M., Nguyen-Hoan M. (2010). A sticky stick? The locus of morphological representation in the lexicon. *Language and Cognitive Processes*, 25(2), 277-296.

Ulicheva, A., Harvey, H., Aronoff, M. & Rastle, K. (2018). Skilled readers' sensitivity to meaningful regularities in English writing. *Cognition*. Epub ahead of print https://www.sciencedirect.com/science/article/pii/S001002771830249X.

Acknowledgements

A financial contribution from SISSA towards the organization of the conference is gratefully acknowledged. While writing this paper, the authors were supported by the European Research Council (ERC Starting Grant n. 679010, STATLEARN, awarded to Davide Crepaldi), the Italian Ministry of Education, University and Research (PRIN 2015 n. 2015PCNJ5F, awarded to Davide Crepaldi), and the Research Foundation–Flanders (FWO; Research Grant n. G011617N, supporting Marco Marelli and Simona Amenta).