Computational methods for descriptive and theoretical morphology

Workshop, 17th International Morphology Meeting

Convenors: Olivier Bonami and Benoît Sagot

While computational morphology is a respected and well-established subfield of computational linguistics with important applications in NLP, there has long been a lack of cross-fertilization with work in descriptive and theoretical morphology. This led to situations of mutual misunderstandings (see e.g. the discussions of theoretical approaches to morphology in Karttunen 2003 and Roark and Sproat 2007) and missed opportunities. The situation has changed radically over the last decade, with important work in 4 directions.

1. Implemented morphological fragments provide a mean to confirm the validity of analyses. This approach, which is familiar from work in syntax and semantics since the mid-1980s, has started becoming more common, in particular within Network Morphology (Corbett & Fraser 1993; Brown & Hippisley 2012) and Paradigm Function Morphology (Stump 2001); see the *Cats CLAW* online tools: http://www.cs.uky.edu/~raphael/linguistics/claw.html), but also through ad-hoc implementations not tied to a specific theoretical framework.
2. Quantitative explorations have started to uncover previously unstudied aspects of the structure and properties of morphological systems. Three main lines of research can be identified, focusing on implicative structure (Finkel & Stump 2007; Ackerman, Blevins & Malouf 2009; Sims 2010; Ackerman & Malouf 2013; Stump & Finkel 2013; Bonami & Boyé 2014), on the inference of inflection classes from raw paradigms (Brown & Evans 2012; Lee 2014) or on the relative information-theoretic compactness of alternate descriptions of a system (Walther & Sagot 2011; Walther, Jacques & Sagot 2014).
3. Such studies rely on the availability of large-scale electronic morphological lexica, which can be developed using lexicographic and/or corpus-based approaches. Such lexica constitute a way to formalize lexical knowledge, enable quantitative linguistic studies of morphology and the lexicon, and pave the way for natural language processing applications. When freely available, they allow for a better mutualisation of efforts and reproducibility of the experiments (see e.g. Le*fff* (Sagot 2010) and Flexique (Bonami, Caron & Plancq 2014) for French).
4. The development of large-scale resources can benefit from computational morphology, especially when dealing with under-resourced languages. One line of work applies unsupervised learning of morphology (e.g. Goldsmith 2001; see Hammarström & Borin 2011 for a recent overview) to bootstrap morphological description (Hammarström 2009); another line attempts to derive automatically implemented grammars and lexica from existing resources (Bender, Schikowski & Bickel 2012; Bender, Crowgey *et al.* 2014).

This workshop is meant as a forum for presentation of work in these four areas or any other area where computational methods are put to use to address descriptive or

theoretical issues in morphology. Submissions are welcome that present a computational method or electronic resource, use such a method or resource in original research, or both.

Submission

We invite submissions in the form of 4-page abstracts, including figures and references.
Abstracts should be submitted via easychair using the following link:
https://easychair.org/conferences/?conf=cmdtm2015
Deadline for abstract submissions: September 20, 2015
Notification of acceptance: October 31, 2015

Invited speaker

Rob Malouf (San Diego State University)

Program committee

Delphine Bernhard (Université de Strasbourg, France)
Emily Bender (University of Washington, USA)
Olivier Bonami (Université Paris-Sorbonne, France)
Dunstan Brown (University of York, United Kingdom)
Roger Evans (University of Brighton, United Kingdom)
Nabil Hathout (CNRS, France)
Rob Malouf (San Diego State University, USA)
Fiammetta Namer (Université de Lorraine, France)
Benoît Sagot (Inria, France)
Andrea Sims (Ohio State University, USA)
Gregory Stump (University of Kentucky, USA)
Géraldine Walther (CNRS, France)

References

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

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

Bender, Emily M., Joshua Crowgey *et al.* 2014. Learning grammar specifications from IGT: A case study of Chintang." *Proceedings of the 2014 Workshop on the Use of Computational Methods in the Study of Endangered Languages, ACL 2014*. 43–53.

Bender, Emily M., Robert Schikowski & Balthasar Bickel. 2012. Deriving a lexicon for a precision grammar from language documentation resources: A case study of Chintang. *Proceedings of COLING 2012*. 247–262.

Bonami, Olivier & Gilles Boyé. 2014. De formes en thèmes. In Florence Villoing, Sarah Leroy & Sophie David (eds.), *Foisonnements morphologiques. Etudes en hommage à Françoise Kerleroux*, 17–45. Presses Universitaires de Paris Ouest.

Bonami, Olivier, Gauthier Caron & Clément Plancq. 2014. Construction d'un lexique flexionnel phonétisé libre du français. In Franck Neveu *et al.* (eds.), *Actes du quatrième Congrès Mondial de Linguistique Française*, 2583–2596.

Brown, Dunstan & Roger Evans. 2012. Morphological complexity and unsupervised learning: Validating Russian inflectional classes using high frequency data. In Ferenc Kiefer, Mária Ladányi & Péter Siptár (eds.), *Current issues in morphological theory: (Ir)regularity, analogy and frequency*, 135–162. Amsterdam & Philadelphia: John Benjamins.

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

Corbett, Greville G. & Norman M. Fraser. 1993. Network Morphology: A DATR account of Russian nominal inflection. *Journal of Linguistics* 29. 113–142.

Finkel, Raphael & Gregory T. Stump. 2007. Principal parts and morphological typology. *Morphology* 17. 39–75.

Goldsmith, John A. 2001. Unsupervised learning of the morphology of a natural language. *Computational Linguistics* 27. 153–189.

Hammarström, Harald. 2009. *Unsupervised learning of morphology and the languages of the world*. Göteborg: Chalmers University of Technology and University of Gothenburg PhD dissertation.

Hammarström, Harald & Lars Borin. 2011. Unsupervised learning of morphology. *Computational Linguistics* 27. 309–350.

Karttunen, Lauri. 2003. Computing with realizational morphology. In Alexander Gelbukh (ed.), *Computational linguistics and intelligent text processing*, 205–216. Heidelberg: Springer.

Lee, Jackson. 2014. *Automatic morphological alignment and clustering*. Tech. rep. University of Chicago.

Roark, Brian & Richard Sproat. 2007. *Computational approaches to morphology and syntax*. Oxford: Oxford University Press.

Sagot, Benoît. 2010. The *Lefff*, a freely available and large-coverage morphological and syntactic lexicon for French. *Proceedings of LREC 2010*.

Sims, Andrea. 2010. Probabilistic paradigmatics: Principal parts, predictability and (other) possible particular pieces of the puzzle. Paper presented at the 14th International Morphology Meeting, Budapest.

Stump, Gregory T. 2001. *Inflectional morphology. A theory of paradigm structure.* Cambridge: Cambridge University Press.

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

Walther, Géraldine, Guillaume Jacques & Benoît Sagot. 2014. The opacity-compactness trade-off: Morphomic features for an economical account of Khaling verbal inflection. Paper presented at the 16th International Morphology Meeting, Budapest.

Walther, Géraldine & Benoît Sagot. 2011. Modélisation et implémentation de phénomènes flexionnels non-canoniques. *Traitement Automatique des Langues* 52(2). 91–122.