NLP for learning and teaching: challenges and opportunities

Georges Antoniadis[*] **— Piet Desmet**[**]

[*] *Université Grenoble-Alpes, LIDILEM, France*
Georges.Antoniadis@univ-grenoble-alpes.fr
[**] *itec-IMEC-KU Leuven, Belgium*
Piet.Desmet@kuleuven-kulak.be

Foreign Language Learning and Teaching, and in particular Computer-Assisted Language Learning (CALL), have been one of the first educational fields, from the early 1960s, to integrate insights and techniques from Natural Language Processing (NLP). Since then, various instructional methods from other fields and disciplines have also incorporated them into electronic learning environments for self-directed learning, blended learning or classroom teaching, as well as for the preparation of learning materials.

However, the use of NLP in CALL is hardly mainstream practice due to a mix of technological and pedagogical restraining factors: "The development of systems using NLP technology is not on the agenda of most CALL experts, and interdisciplinary research projects integrating computational linguists and foreign language teachers remain very rare" (Amaral & Meurers, 2011).

One of the main concerns is the technological readiness of NLP. Whereas Nyns (1989) still stated that he is "pessimistic about the possibility of intelligent computer-assisted language learning (ICALL)", one can only but observe that NLP has evolved and improved dramatically over the last twenty years, an evolution that is to a certain extent linked to the successful paradigmatic shift from rule-based NLP to statistical NLP integrating supervised and unsupervised machine-learning algorithms. However effective these algorithms may be for the analysis of native language, they face the challenge of handling erroneous and undesired output in the case of learner language.

Nevertheless, NLP in general and the broader field of intelligent computer-assisted language learning have made substantial progress, and some of the results

TAL. Volume 57 n° 3/2016, pages 7 à 13

of ongoing research and development are now making their way towards more widespread integration into L2 teaching and learning (Heift, 2017).

NLP has overall contributed to the improvement of learning environments, and to the development of research in the related fields. Today, online learning tools such as Massive Open Online Courses (MOOCs), Small Private Online Courses (SPOCs), Computer-Assisted Pronunciation Teaching (CAPT) systems, Computer-Assisted Instruction systems for mathematics, sign language learning applications or Intelligent Tutoring Systems (ITSs), among many others, are heavy "consumers" of NLP, or about to become some.

In language-learning applications, the aspects of NLP most frequently involved are analysis of learners' responses, feedback provision, automated generation of exercises, or the monitoring of learning progress. Other aspects related to learning and teaching also involve NLP, such as plagiarism detection, compilation of parallel corpora and annotated learner corpora, development of ontologies for the associated domains, etc. (Gamper & Knapp, 2002).

More generally, we see in language learning and teaching at least three components in which NLP is implied, involving, first, the exploitation of native language, second, the use of learner output and, third, intelligent tutoring. These three fields each bring to the table challenges, but likewise opportunities. Seven possible functions for ICALL applications exemplify further how the technology can be readily integrated.

A. Native language: enrich and exploit

   1. Resource generator: creating reference materials such as search engines on monolingual, bilingual corpora, or corpus-enriched learner dictionaries.
   2. Reading companion: helping students understand L2 materials through annotation layers, both on a formal and semantic level.
   3. Exercise and test generator: (semi-)automatic generation of exercise and test items based on the analysis of L2 text materials and/or on the analysis of learner errors.

B. Learner language: finding errors

   4. Error detector, feedback generator and automatic scoring tool: analysis of learner output using rule-based or data-driven statistical NLP-approaches in order to go beyond (more limited) approximate string matching techniques.
   5. Writing aid: supporting the second-language user in writing a functional, well-formed text.

C. Model learner behavior: measure complexity

6. Input provider: (semi-)automatic selection of comprehensible and authentic text material based on readability and formal complexity, analysis of meaning or text categorization.
7. Adaptive item sequencer: creating adaptive learning environments based on student modeling.

The generation of resources and access to resources through pedagogical criteria is certainly one of ICALL's very fruitful functions that provide learners and teachers with easy access to genuine use of the language through exploitation of annotated monolingual and/or parallel corpora (Nerbonne, 2003; Loiseau *et al.*, 2013; Leńko-Szymańska & Boulton, 2015; Chinkina & Meurers, 2016).

Another richly explored function is that of the reading companion (Roosmaa & Prószéky, 1988; Liaw & English, 2017) helping learners to understand foreign-language input by providing annotation layers that offer formal, semantic and/or referential information going from lemmatization and part-of-speech tagging to named-entity recognition or topic detection. The automated generation of exercises and tests is yet another function linked to the exploitation and the enrichment of native language. This leads to the creation of mostly closed-form exercises on morpho-syntactic correctness, lexical restrictions or semantic appropriateness (Meurers *et al.,* 2010; Ziegler *et al.,* 2017).

When passing from the analysis of native to learner language, we come across the fourth function of ICALL, i.e., error detection, feedback generation and automatic scoring (Amaral & Meurers, 2011; Heift, 2003; Heift & Schulze, 2007; Nagata, 2009). Most of the tools focusing on this function are parser-based and include such techniques as mal-rules or constraint relaxation (Vandeventer, 2001). CALL is conceived as an intelligent tutor focusing on linguistic knowledge, integrating NLP and artificial intelligence techniques and exploiting larger resources (corpora, reference materials). Such systems focus not only on written language but also on speech, helping to identify pronunciation errors using automatic speech-recognition techniques (Eskenazi, 1999; van Doremalen *et al.*, 2016).

More recently, CALL is also seen as an effective facilitator focusing on language skills integrating tutorial components. So, for instance, writing aids support the learner in writing a functional well-formed text and offer not so much automated corrections, but rather suggestions to help the learner through on-the-fly prompts, query-and-answer interactive systems or post-writing checks (Cotos, 2011; Wanner *et al.*, 2013). To train speaking skills, dialogue-based CALL has proven to be quite effective (Bibauw *et al.*, 2015). Although recognition of free answers is possible, most dialogues constrain the learner's interactional turn on meaning — the degree of negotiability of the content of each message —, or on form, limiting the range of linguistic items that they can use. Form and meaning are not subjected to constraints in a dichotomous way: they rather follow a continuum going from totally constrained (pre-set form/meaning) to totally unconstrained production (free form/meaning) (Bailey & Meurers, 2008).

Finally, NLP is also quite effective in measuring complexity that allows CALL to become a reliable input provider offering, e.g., vocabulary or reading practice adapted to the current proficiency level and interest of the learner (Tack *et al.*, 2016). Complexity measurement opens also the way to adaptive item sequencing in which the selection of the items is a function of such parameters as the difficulty of the item and the learner profile (Wauters *et al.*, 2010).

The contribution of NLP to all of these functions is generally regarded as positive. It must be recognized, however, that only a handful of such applications have made it to the general public as commercial software. In most cases, the systems never leave the laboratory and have a limited range of use, sometimes only as a proof of concept. This might be due to such factors as the high production cost of NLP resources, the current quality of NLP results or the lack of adequate integration into the learning environment.

The goal of this issue dedicated to NLP for learning and teaching is to summarize the contribution of NLP to instructional systems, not only at a theoretical level (opportunities, limitations, integration methods) but also to the extent to which intelligent CALL environments (or part of them) can be realised.

The two first contributions are devoted to computer-assisted pronunciation training (CAPT). Sylvain Detey, Lionel Fontan and Thomas Pellegrini offer a synthetic overview of the added value of speech processing for pronunciation training. In this field, NLP should progressively lead to automatic error correction, including error diagnosis and remediation. The authors state that one of the main breakthroughs in CAPT is the analysis of oral-learner corpora, and they argue for a more intense scientific interaction between CALL researchers and speech engineers.

The second contribution, by Pierre Magistry, Murielle Fabre and Yoann Goudin, offers a concrete use case of CAPT for Chinese. Learning a new language having the kind of graphical system Chinese has represents a real challenge. The crucial point consists in grasping the right Orthographic-to-Phonology Correspondence (OPC) between sound and the graphical units of the sinogram. In this decoding procedure, the best candidates to focus on are graphical units corresponding to phonological cues. Going beyond the simple and traditional vocabulary list and frequency strategy, the authors propose a computational model which enables us to introduce and follow the learner into the graphic system and its rules by reinforcing the representation of the phonological cues and their reliability.

The following three contributions focus on written language rather than on speech. Ildikó Pilán, Elena Volodina and Lars Borin present a framework and its implementation relying on NLP methods, which aims at the identification of exercise item candidates from corpora. The hybrid system combining heuristics and machine-learning methods includes a number of relevant selection criteria. They focus on two fundamental aspects, i.e., linguistic complexity and the dependence of the extracted sentences on their original context. Previous work on exercise generation addressed these two criteria only to a limited extent, and a refined overall

candidate sentence selection framework appears also to be lacking. In addition to a detailed description of the system, they present the results of an empirical evaluation conducted with language teachers and learners, which indicate the usefulness of the system for educational purposes.

With the contribution of Yves Bestgen, we enter the promising field of the automatic evaluation of written production. The author focuses on the automatic measurement of phraseological complexity. Rather than limiting himself to lexical measures, he proposes phraseological indices of writing proficiency that prove to present a quite important degree of generalizability.

The final contribution, by Marie-Josée Hamel, Nikolay Slavkov, Diana Inkpen and Dingwen Xiao, reviews some long-standing issues in the literature on written corrective feedback, discusses the potential of technology to support some of the tasks involved in the essay marking process, and then presents a new error annotation tool, *MyAnnotator,* developed by the authors with the purpose of facilitating technology-mediated corrective feedback. They offer an overview of different types of electronic tools that can be used in the teaching of writing, including editors, correctors and annotators, and then draw brief comparisons between *MyAnnotator* and other similar tools.

ICALL is definitely an interesting field allowing at least for a moderate optimism. We intend the contributions to this issue to be clearly illustrative of the recent advances in NLP-based CALL. Although the best is yet to come, there is no reason anymore for the more general scepticism that is still widespread within the field of language learning and teaching.

Acknowledgements

We want to thank the *TAL journal* editors and committee as well as the specific scientific committee. We are particularly grateful to the reviewers for their time and effort to improve this special issue.

Specific scientific committee (by alphabetical order):

Véronique Aubergé, LIG, Université Grenoble-Alpes, France; Yves Bestgen, IPSY, Université catholique de Louvain, Belgium; Éric Bruillard, STEF, ENS Cachan, France; Cristelle Cavalla, DILTEC, Université Sorbonne-Nouvelle, France; Thierry Chanier, LRL, Université Blaise-Pascal de Clermont-Ferrand, France; Françoise Demaizière, Université Paris-Diderot, France; Philippe Dessus, LSE, Université Grenoble-Alpes, France; Sylvain Detey, Waseda University, Japan; Walt Detmar Meurers, Universität Tübingen, Germany; Maxine Eskenazi, Carnegie Mellon University, USA; Cédrick Fairon, CENTAL, Université catholique de Louvain, Belgium; Dan Flickinger, LinGO Laboratory, Stanford University, USA; Nuria Gala, LIF, Université Aix-Marseille, France; Sylviane Granger, CECL, Université catholique de Louvain, Belgium; Natalie Kübler, CLILLAC-ARP, Université Paris-Diderot, France; Jean-Marc Labat, LIP6, Université Pierre-et-Marie-Curie, France;

References

Amaral, L., Meurers, D., "On using intelligent computer-assisted language learning in real-life foreign language teaching and learning", *ReCALL*, 23, 2011, p. 4-24.

Bailey, S., Meurers, D., "Diagnosing meaning errors in short answers to reading comprehension questions", *Proceedings of the 3rd Workshop on Innovative Use of NLP for Building Educational Applications,* 2008, p. 107-115, Columbus, Ohio: Association for Computational Linguistics.

Bibauw, S., François, T., Desmet P. "Dialogue-based CALL: an overview of existing research". *In*: Helm, F. (Ed.), Bradley, L. (Ed.), Guarda, M. (Ed.), Thouësny, S. (ed.), *Critical CALL – Proceedings of the 2015 EUROCALL Conference,* Padova, Italy. EUROCALL, 26-29 August 2015, p. 57-64, Dublin: Research-publishing.net.

Chinkina, M., Meurers, D., "Linguistically-aware information retrieval: Providing input enrichment for second language learners". *In*: *Proceedings of the 11th Workshop on Innovative Use of NLP for Building Educational Applications,* 2016, p. 188-198, San Diego, CA: ACL.

Cotos, E., "Potential of automated writing evaluation feedback", *CALICO Journal* 28:2, 2011, p. 420-459.

Eskenazi, M., "Using automatic speech processing for foreign language pronunciation tutoring: some issues and a prototype", *Language Learning & Technology,* 2:2, 1999, p. 62-76.

Gamper, J., Knapp, J., "A review of intelligent CALL systems", *Computer Assisted Language Learning* 15, 2002, p. 329-342.

Heift, T., "Multiple learner errors and meaningful feedback: A challenge for ICALL systems", *CALICO Journal* 20, 2003, p. 533-548.

Heift, T., "History and key developments in intelligent computer-assisted language learning (ICALL)". *In*: Thorne, S., May, S. (eds)*, Language, Education and Technology,* Cham: Springer International, 2017, p. 1-12.

Heift, T., Schulze, M., *Errors and Intelligence in Computer-Assisted Language Learning. Parsers and Pedagogues,* New York & London: Routledge, 2007.

Meurers, D., Dickinson, M., "Evidence and interpretation in language learning research: opportunities for collaboration with computational linguistics", *Language Learning* 67, 2017, p. 66-95.

Meurers, D., Ziai, R., Amaral, L., Boyd, A., Dimitrov, A., Metcalf, V. *et al.*, "Enhancing authentic web pages for language learners". *In*: *Proceedings of the 5th Workshop on*

Innovative Use of NLP for Building Educational Applications, 2010, p. 10-18, Los Angeles, CA: ACL.

Nagata, N., "Robo-sensei's NLP-based error detection and feedback generation", *CALICO Journal* 26, 2009, p. 562-579.

Nerbonne, J., "Computer-assisted language learning and natural language processing". *In*: Mitkov R. (ed), *The Oxford Handbook of Computational Linguistics*, Oxford: Oxford University Press, 2003, p. 670-698.

Roosmaa, T., Prószéky, G., "GLOSSER – Using language technology tools for reading texts in a foreign language". *In*: Jager, S., Nerbonne, J.A., Van Essen A. (eds)*, Language Teaching and Language Technology*, Lisse: Swets & Zeitlinger, 1998, p. 101-107.

Ziegler, N., Meurers, D., Rebuschat, PE., Ruiz, S., Moreno Vega, J., Chinkina, M., Li, W., Grey, S., "Interdisciplinary research at the intersection of CALL, NLP, and SLA: methodological implications from an input enhancement project", *Language Learning* 67, 2017, p. 209-231.

Liaw, ML., English, K., "Technologies for teaching and learning L2 reading". *In*: Chapelle, C.A., Shannon S. (eds), *The Handbook of Technology and Second Language Teaching and Learning*, 2017, Chichester: Wiley-Blackwell.

Leńko-Szymańska, A., Boulton, A., *Multiple Affordances of Language Corpora for Datadriven Learning*, 2015, Amsterdam: John Benjamins.

Loiseau, M., Antoniadis, G., Ponton, C., "Facets and prisms as a means to achieve pedagogical indexation of texts for language learning". *In*: Cordeiro J., Virvou M., Shishkov B. (ed.), *Software and Data Technologies,* vol. 170, 2013, p. 253-268, Springer Verlag.

Tack A., François T., Ligozat A., Fairon C., "Modèles adaptatifs pour prédire automatiquement la compétence lexicale d'un apprenant de français langue étrangère". *In*: *Actes de la conférence conjointe JEP-TALN-RECITAL 2016, volume 2 : TALN. Conférence sur le Traitement Automatique des Langues Naturelles (TALN),* Paris, France, 4-8 July 2016, p. 221-234, Paris: ATALA.

Van Doremalen, J., Boves, L., Colpaert, J., Cucchiarini, C., Strik, H., "Evaluating automatic speech recognition-based language learning systems: a case study", *Computer Assisted Language Learning*, 29:4, 2016, p. 833-851.

Vandeventer, A. 2001. "Creating a grammar checker for CALL by constraint relaxation: A feasibility study", *ReCALL* 13:1, 2001, p. 110-120.

Wanner, L., Verlinde, S., Ramos, M.A., "Writing assistants and automatic lexical error correction: word combinatorics". *In*: Kosem, I., Kallas, J., Gantar, P., Krek, S., Langemets, M., Tuulik, M. (eds.), *Electronic Lexicography in the 21st Century: Thinking Outside the Paper. Proceedings of the eLex 2013 Conference,* 17-19 October 2013, Tallinn, Estonia. Ljubljana/Tallinn: Trojina, Institute for Applied Slovene Studies/Eesti Keele Instituut, p. 472-482.

Wauters K., Desmet P., Van Den Noortgate W., "Adaptive item-based learning environments based on the item response theory: possibilities and challenges", *Journal of Computer Assisted Learning* 26:6, 2010, p. 549-562.