

**A syntax-based query system
adapted to language learning and teaching**

Ilaine Wang

MoDyCo, Université Paris Ouest Nanterre La Défense
`i.wang@u-paris10.fr, sylvain@kahane.fr, isabelle.tellier@univ-paris3.fr`

American Association for Corpus Linguistics 2016

14th Annual Technology for Second Language Learning (TSSL) Conference

Category 3: application

A syntax-based query system adapted to language learning and teaching

When a language learner encounters an unknown grammatical construction, one solution is to look it up in textbooks or directly in grammars, where a definition as well as several examples of canonical uses are provided. However, in some cases, explicit rules and a few number of uses are not sufficient to fully comprehend a grammatical construction, especially if the learner's native language is typologically distant from the target language.

Next step could be to search more examples, perhaps in authentic corpora to observe and analyse what is considered as natural and usual in the target language. Learners would therefore be actors of the construction of their own knowledge, which was encouraged by John's Data-driven learning approach. However, using a grammatical construction as a query may not be as easy as using plain words to obtain concordances. Indeed, learners would need to provide a description of the construction, which is not self-evident for non-specialists.

In this study, we present our efforts to provide the missing link between examples taken from textbooks to illustrate grammatical constructions and subsidiary instances of those constructions that can be found in context in native corpora. The process simply works like a syntax-based search engine: from a sentence in input containing the targeted grammatical construction, our tool provides other sentences in context, ranked by the similarity of their construction. As an illustration, we could retrieve hundreds of relevant examples of a given construction based on a few examples displayed in a textbook, including similar constructions which are not mentioned in grammars as possible variations. The focus of our study is on Korean language learners, but the methodology could be extended to any language and teachers are the other evident target as this method can be useful in the preparation of teaching materials.