

Author(s): Emmanuel Keuleers

University/Affiliation: Center for Psycholinguistics & CNTS, University of Antwerp

Email address(es): emmanuel.keuleers@ua.ac.be

Predicting exceptions may be harmful

In theories about the representation of inflected forms in the mental lexicon, emphasis is given either to computation or to storage. Theories that emphasize computation see the production of a complex form such as the regular past tense form *walked* as the result of a process that attaches the suffix *-ed* to the stem *walk*, while theories that emphasize storage consider that complex forms are stored in their entirety and that production simply involves retrieval of the full form. However, while theories may disagree on whether a past tense form such as *walked* is retrieved or computed, all theories agree that at least some exceptional forms (e.g., *be-was*, *go-went*) are retrieved rather than computed. To put it more generally, the more exceptional a form is, the more likely its production would be considered the result of a retrieval process.

Nonetheless, being able to *compute* exceptional forms which are assumed to be *retrieved* in ordinary language production seems to be a desirable characteristic for psychologically motivated computational models. In simulation tasks where part of the lexicon is treated as novel material for which complex forms are to be predicted on the basis of the remaining part of the lexicon (henceforth called *lexical reconstruction*), even the correct generation of the most exceptional form counts toward better performance. Similarly, a model that is able to correctly learn the mappings between stems and inflected forms in a particular domain is considered to have fully mastered a skill that is attributed to language users. For instance, Rumelhart and McClelland's (1986) pattern associator for the English past tense was evaluated on its ability to produce inflected forms of existing verbs through a feedforward network.

However, if we would know of a method to cause selective and reversible memory loss in a language user so that we could subject her to a lexical reconstruction experiment, it is highly plausible that her performance would not match that of such a computational model. Specifically, we can assume that she would not generate the attested forms of many exceptional items. Accordingly, it is very doubtful that each time an inflected form is produced, it is generated on the basis of a stem form, as is the case in a pattern associator.

This is not a serious problem if lexical reconstruction actually requires the same abilities as those used by speakers in producing novel complex forms. It may be the case that a model that performs well on lexical reconstruction also performs well on tasks where *pure generalization* ability is tested, i.e., where language users are asked to generate complex forms on the basis of nonce words (also known as *wug* testing, after Berko, 1958). In this paper, I present evidence that models that perform well in a lexical reconstruction task do not perform well in pure generalization task precisely because these models tend to predict *exceptional* patterns while human participants do so to a far lesser degree. This evidence comes from a large-scale simulation study in which

Keuleers and Daelemans (2007) specifically contrasted the performance of computational models on a lexical reconstruction task and on two pure generalization tasks involving Dutch noun plural production. In the lexical reconstruction task models had to predict plural forms for a random selection of 1/20th of the forms in a lexicon of more than 18 000 items. In the first pure generalization task, models had to predict which of two alternative plural forms of a list of nonce nouns was the most frequent choice of participants in an experiment by Baayen et al. (2002); In the second task models had to predict the plural forms for a list of nonce nouns used in an experiment by Keuleers et al. (2007).

The computational models that were used on these tasks were memory-based learning (MBL) models — so called because they make no abstraction from the learning material. These models can be seen as a more sophisticated version of the k nearest neighbors approach, in which the class of a novel item is based on the majority class of its k most similar neighbors. For instance, in an MBL model with $k = 7$, the plural suffix for a novel noun is based on the most frequent plural suffix among the 7 most similar sounding nouns (technically, all nouns at the 7 nearest distances). Similarity is computed on the basis of aligned phonological representations of these forms. Interestingly, the parameter k has a direct relation to the ability of a model to predict exceptional forms. All other things being equal, the lower the value of k , the better the model is able to predict inflectional patterns with a low frequency. Accordingly, the higher the value of k , the more *exceptional* patterns will be outvoted by more influential patterns.

Figure 1. Prediction accuracy of memory-based learning models with variable k on a lexical reconstruction task and on two pure generalization tasks.

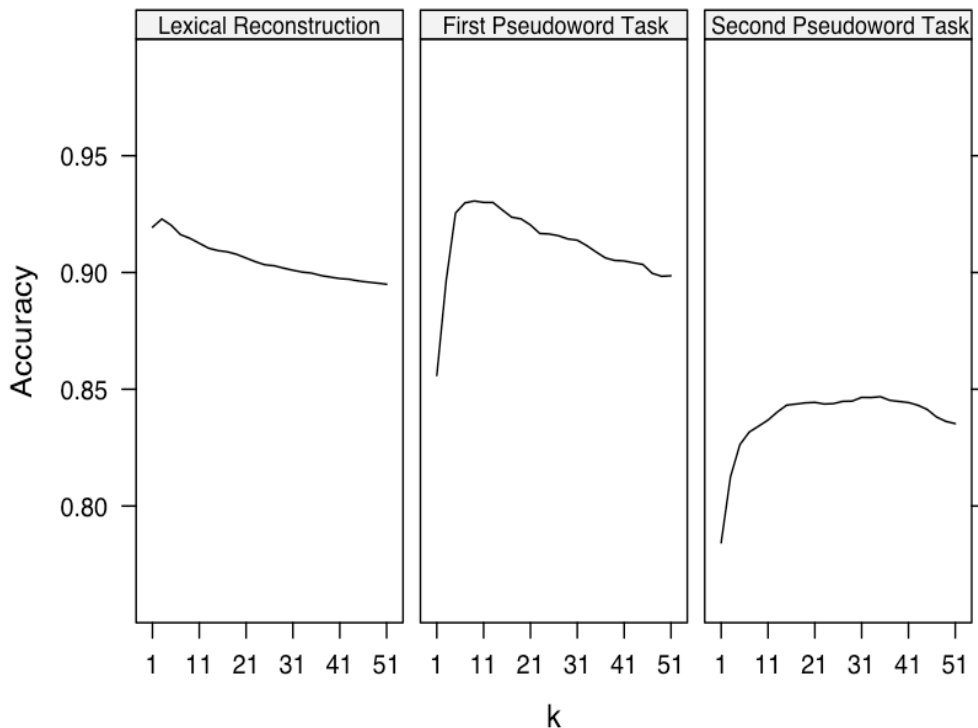


Figure 1 shows the optimal value for k for the three tasks described above. While the lexical reconstruction task benefits from a low value for k , the pure generalization tasks require a substantially higher value. The most dramatic finding, however, is that a nearly optimal value for the lexical reconstruction task ($k = 1$) is particularly unsuited for the pure generalization tasks. On the basis of these findings, *predicting exceptions* may be considered harmful in modeling language processes. Recent results from memory-based learning of past tense inflection in English (Keuleers & Sandra, submitted) and Dutch (Vandekerckhove, Keuleers, & Sandra, in preparation) support this conclusion. These findings may have consequences on developing theories of language learning. Particularly, they add value to the idea that good performance on the simulation of a linguistic processing task does not necessarily entail similarity to a language user.

References

- Baayen, R. Harald, Schreuder, Robert, De Jong, Nivja, and Krott, Andrea 2002. Dutch Inflection: The Rules That Prove the Exception. In *Storage and Computation in the Language Faculty*, Nootboom, Sieb, Weerman, Fred, and Wijnen, Frank (eds), 61–92. Dordrecht: Kluwer.
- Berko, Jean 1958. The child's learning of English morphology. *Word* 14: 150–177.
- Keuleers, Emmanuel and Daelemans, Walter 2007. Memory-Based Learning Models of Inflectional Morphology: A Methodological Case Study. *Lingue e Linguaggio* 6 (2): 151–174.
- Keuleers, Emmanuel and Sandra, Dominiek. Similarity and Productivity in the English Past Tense. Submitted for publication.
- Keuleers, Emmanuel, Sandra, Dominiek, Daelemans, Walter, Gillis, Steven, Durieux, Gert, and Martens, Evelyn 2007. Dutch plural inflection: The exception that proves the analogy. *Cognitive Psychology* 54 (4): 283–318.
- Rumelhart, David E. and McClelland, James L. 1986. On Learning the Past Tenses of English Verbs. In *Psychological and Biological Models* [Parallel Distributed Processing. Explorations in the Microstructure of Cognition 2], McClelland, James L., Rumelhart, David E., and The PDP Research Group (eds), 216–271. Cambridge, MA: MIT Press.
- Vandekerckhove, Bram, Keuleers, Emmanuel, and Sandra, Dominiek. The productivity of inflectional patterns in the Dutch past tense: The roles of distance and analogical support. Manuscript in preparation.