Title

Sources of variability in linguistic memory systems

Journal

Bilingualism Language and Cognition, 20(4)

ISSN

1366-7289

Author

WAGERS, MATTHEW W

Publication Date

2017-08-01

DOI

10.1017/s1366728916000997

Peer reviewed

Sources of variability in linguistic memory systems

Matthew W. Wagers

Department of Linguistics

University of California, Santa Cruz

Address for correspondence:

Stevenson College

University of California, Santa Cruz

1156 High St

Santa Cruz, CA, 95064 USA

mwagers@ucsc.edu

The target article makes a strong case that L2 comprehenders recover linguistic representations that are qualitatively similar to those recovered by L1 comprehenders. Moreover as they attempt to link non-adjacent elements, they do so using the same basic mechanism: cue-based retrieval in a content-addressable memory (Van Dyke & Lewis, 2003). In this commentary, I will not address the empirical adequacy of the argument, but instead consider some interesting theoretical challenges it poses for our understanding of working memory in sentence processing.

According to Cunnings's proposal, differences between L1 and L2 comprehenders are attributed to "variability in the ability to retrieve information constructed during online parsing from memory and in particular … L2 sentence processing is more susceptible to effects of similarity-based retrieval interference." Cunnings assumes a 'bi-partite' memory system (McElree, 2006): for tasks of even moderate complexity, most information is stored in long-term memory (LTM) while a small, restricted set of information is in the focus-of-attention. Information in the focus-of-attention is immediately available for incorporation into on-going processes, while extra-focal information must be retrieved and restored to the focus of attention – a costly, noisy operation. The probability that the "target" encoding is retrieved is a function of how distinctly it is identified by the retrieval cues.

Cunnings argues that L2 speakers are prone to retrieval interference in the domains of agreement and binding. Like L1 comprehenders, L2 comprehenders are susceptible to agreement attraction. Unlike L1 comprehenders, L2 comprehenders sometimes apparently consider binding-theoretic inaccessible antecedents. To explain the difference with L1 comprehenders, discourse-related cues like topic-hood are argued to be overweighted in L2 processing. This argument, however, shifts the burden of explanation from susceptibility to retrieval interference per se, to

how the retrieval structures differ between the two speaker groups. More wrong constituents are retrieved in L2 processing not because working memory is inherently noisier in L2, but because it is less well adapted to L2 sentence structures.

This account invites us to consider how such adaptation might improve with experience. A natural locus is in the cue sets used for retrieval: based on Cunnings's account, we might expect that learners get better by weighting diagnostic syntactic cues more highly than heuristic discourse cues – or, more generally, if they downweight cues that are not valid in the L2 grammar (though they may be typologically or statistically justified).

Another locus of adaptation worth considering is in how linguistic representations are parceled between focal and extra-focal states. For example, in ACT-R (Lewis & Vasishth, 2005), syntactic expressions are chunked exhaustively into non-overlapping maximal projections. These chunks are linked into a syntactic whole by virtue of mother/daughter features that point to other chunks. Although this modeling decision is theoretically well-motivated, we actually have little relevant data on how large compositional objects like a sentence are 'carved up' by the focus of attention (Wagers & McElree, 2013). Moreover, chunking at a very local grain can make it computationally awkward to constrain retrieval in a content-addressable memory by certain perspectival relations, like "c-command" or "clausemate" (Alcocer & Phillips, 2012).

Suppose we relax the assumption that the focus of attention can contain only one or two fully-specified maximal projections. Instead, suppose that it can span rather more global stretches of material, but that the depth at which it is represents that material trades off with breadth. Some nodes, like the most recent one, are likely to be represented with fully-elaborated feature structures; other nodes, less recent or non-local, are represented merely as labels or unique identifiers to a fuller representation in LTM. This conception of the focus of attention as a

sort of lightly-annotated and underspecified tree is inspired by the retrieval structures that Ericsson & Kintsch (1995) advocate to explain expert performance. Instead of maintaining just as much more information as capacity allows, it emphasizes preserving useful retrieval cues for navigating that information in LTM. An advantage of this approach is that it allows for locality constraints to be more easily implemented, since information about syntactic ordering is now present in the focus of attention itself and not recovered via retrieval. Wagers (2013) considers how reconceptualizing the focus of attention this way can help explain why *wh*-dependency processing is so interference-robust.

This leads to the following conjecture: L2 comprehenders may render themselves more susceptible to retrieval interference not only by using maladaptive cue-sets, but also by inefficiently allocating focus of attention in a way that hinders them from enforcing locality constraints retrospectively. It is potentially telling that an anti-locality constraint, Principle B, is on the whole, more faithfully implemented (Patterson, Trompelt & Felser, 2014) than a locality constraint, Principle A (Felser et al., 2009, Felser & Cunnings, 2012). It suggests that the L2 comprehenders may be better at blocking retrieval of recently encountered antecedents, about whose position in the tree they have relatively good information, than they are at blocking the retrieval of more distant antecedents, about whose position they have poorer information.

It is but a conjecture, and given our present understanding, Cunnings' proposal about L2 comprehenders' cue sets provides a compact and more useful explanation of their particular pattern of errors in variable binding and pronoun resolution. But cue sets are only one possibility for explaining how WM skills may diverge between L1 and L2 comprehenders. And Cunnings' broader attention to this issue shines a spotlight on just how much more we need to understand about the memory systems of all language users. *899 words*

Alcocer, P. & Phillips, C. (2012). Using Relational Syntactic Constraints in Content-Addressable Memory Architectures for Sentence Parsing. University of Maryland. Manuscript.

Ericsson, K. A., & Kintsch, W. (1995). Long-Term Working Memory. *Psychological Review*, *102*, 211–245.

Lewis, R. L., & Vasishth, S. (2005). An Activation-Based Model of Sentence Processing as Skilled Memory Retrieval. *Cognitive Science*, *29*, 375–419.

McElree, B. (2006). Accessing Recent Events. *The Psychology of Learning and Motivation*, *46*, 155–200.

Van Dyke, J., & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language*, *49*, 285–316.

Wagers, M. (2013). Memory mechanisms for wh-dependency formation and their implications for islandhood. In J. Sprouse & N. Hornstein (Eds.) Experimental Syntax and Island Effects. Cambridge, UK: Cambridge University Press.

Wagers, M. & Brian McElree. (2013). Working memory and language processing: theory, data and directions for future research. In C. Boeckx & K. Grohmann (Eds.) *The Cambridge Handbook for Biolinguistics.* Cambridge, UK: Cambridge University Press.