Cognition 176 (2018) 232–247

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

The sentence wrap-up dogma

Laurie A. Stowe[a], Edith Kaan[b,*], Laura Sabourin[c], Ryan C. Taylor[d]

[a] Rijksuniversiteit Groningen, The Netherlands
[b] University of Florida, United States
[c] University of Ottawa, Canada
[d] University of British Columbia, Canada

ARTICLE INFO

Keywords:
Sentence wrap-up
Sentence processing
ERP
Closure

ABSTRACT

Current sentence processing research has focused on early effects of the on-line incremental processes that are performed at each word or constituent during processing. However, less attention has been devoted to what happens at the end of the clause or sentence. More specifically, over the last decade and a half, a lot of effort has been put into avoiding measuring event-related brain potentials (ERPs) at the final word of a sentence, because of the possible effects of sentence wrap-up. This article reviews the evidence on how and when sentence wrap-up impacts behavioral and ERP results. Even though the end of the sentence is associated with a positive-going ERP wave, thus far this effect has not been associated with any factors hypothesized to affect wrap-up. In addition, ERP responses to violations have not been affected by this positivity. "Sentence-final" negativities reported in the literature are not unique to sentence final positions, nor do they obscure or distort ERP effects associated with linguistic manipulations. Finally, the empirical evidence used to argue that sentence-final ERPs are different from those recorded at sentence-medial positions is weak at most. Measuring ERPs at sentence-final positions is therefore certainly not to be avoided at all costs, especially not in cases where the structure of the language under investigation requires it. More importantly, researchers should follow rigorous method in their experimental design, avoid decision tasks which may induce ERP confounds, and ensure all other possible explanations for results are considered. Although this article is directed at a particular dogma from a particular literature, this review shows that it is important to reassess what is regarded as "general knowledge" from time to time.

1. Introduction

In recent years it has become very difficult to publish an article in which event-related brain potentials (ERPs) are measured on the last word of the sentence, at least when syntactic variables are manipulated, because of the danger of sentence wrap-up effects. Readers are frequently assured that the authors have avoided problems with sentence-final position by using sentence medial targets, with no further discussion, assuming that the problem is general knowledge. In those cases where the response on a sentence-final word *is* measured, a section of the discussion is generally devoted to discussing why any effects being reported are *not* sentence wrap-up effects and can be generalized beyond this position. Why all the fuss? Reviewers are concerned about sentence-final wrap-up processes; these are assumed to have large enough effects to obscure or modify typical (i.e. sentence medial) effects, to the extent that an effect found at the end of the sentence may be uninterpretable and certainly cannot be generalized to sentence-medial positions.

The goal of this article is to raise the question of how and when sentence wrap-up impacts ERP results, in order to determine under what circumstances sentence-final measurement should be a concern in ERP experiments. One frequent assumption in the literature is that we can simply avoid the potential problems by using designs which avoid sentence-final positions. Unfortunately, sometimes it is just plain impossible to apply this avoidance strategy, as with strictly verb-final languages like Japanese (see Ueno & Kluender, 2003, for an example).

Even with languages in which the verb is not strictly sentence final, like Dutch and German, avoiding sentence-final placement sometimes requires the use of a non-preferred word order, impacting the acceptability of an out-of-the-blue sentence. This makes the outcome of the experiment hard to interpret for other reasons, clearly an undesirable result.

We will argue that the effects of sentence wrap-up are in fact not nearly as extensive as a reader of the ERP literature would be justified in assuming, given that so many authors avoid it. The actual experimental evidence does not support extensive modifications of ERP effects by sentence wrap-up. Further, we will argue that the "evil" status of sentence wrap-up effects in ERPs has become a dogma, a belief

* Corresponding author at: Department of Linguistics, University of Florida, Box 115454, Gainesville, FL 32611, United States.
E-mail address: kaan@ufl.edu (E. Kaan).

https://doi.org/10.1016/j.cognition.2018.03.011
Received 9 January 2017; Received in revised form 12 March 2018; Accepted 13 March 2018

system, rather than a potential concern which should be considered in evaluating a design, equivalent to any other concern, such as plausibility, the naturalness of the stimuli and effects of decision-making.

The structure of the article is as follows. The next Section 2 contains a basic description of several conceptualizations of wrap-up with some considerations of how each might impact ERP studies. The next two sections are parallel historical accounts of wrap-up in the behavioral psycholinguistic literature on the one hand (Section 3), and the ERP literature on the other hand (Section 4). The behavioral conceptualizations form an important check on the plausibility of the notion of wrap-up as used by ERP researchers. The final Section 5 provides a summary of the arguments in support of the sentence wrap-up dogma and the evidence which currently allows us to evaluate them.

2. Sentence wrap-up: What is it actually?

To begin, it is useful to examine the definition of sentence wrap-up. The theoretical concept of sentence wrap-up is generally not well-defined. Without a clear definition, what counts as evidence for—or against—sentence wrap-up is unclear. First, although the term sentence wrap-up is used widely, the evidence for wrap-up is in general also applicable to the ending of a clause, and a number of the theoretical constructs discussed below explicitly contain this assumption. We will continue to use the term *sentence wrap-up* because it is the common one in the ERP literature, based on Just and Carpenter (1980), but the point should be kept in mind that this concept also includes end-of-clause effects and this will be stressed at several points with regards to the behavioral evidence.

In the sections below we will discuss several different views. The first issue is whether there are processes that are carried out at the end of the sentence which are different in nature than those carried out within the sentence. This is the case under views of sentence wrap-up in which (1) there is a specific stage or type of processing which involves literally wrapping up the sentence, or in which (2) there are certain stages of linguistic processing that can only occur at the end of the clause or sentence.

The most common suggestions of the latter type are that syntactic processing (3.1) or the integration of a proposition within the larger context (3.2) occur at the end of the clause. Under either of these views, certain predictions follow: there are certain processes that can only occur at this point, they should not be seen earlier in the clause, and they should make use of a different set of neural resources than any (similar) processes that do occur earlier. The view that there is a distinct stage of processing which is literally responsible for wrap-up has not received much formalization though there is a stage with this name in the model of Just and Carpenter (1980). They explain this stage as follows "A special computational episode occurs when a reader reaches the end of a sentence. This episode, called sentence wrap-up, is not a stage of processing defined by its function, but rather by virtue of being executed when the reader reaches the end of a sentence" (p. 345). If taken as a stage that is separate from the processes carried out at other points in the sentence, wrap-up may, for instance, be a checking process to make sure that processing is complete before the memory representation of lower-level information is discarded at the end of a sentence. A process which performs this check would be specific to the end of the sentence and would presumably involve neural resources that are not made use of before the clause is completed. Although theoretically possible, this view of sentence wrap-up has not been explicitly employed, as far as we are aware, and is not the view espoused by Just and Carpenter.

A different view of wrap-up is that it involves the completion of processes which were not or could not be carried out earlier in the sentence for some reason, for instance, the assignment of referents to pronouns, establishing inter-clausal connections, or dealing with semantic inconsistencies (see Sections 3.2 and 3.3). This is more in line with Just and Carpenter's (1980) explanation of their wrap-up stage. There does not seem to be any reason to assume that those processes are carried out in a different way at the end of the sentence than earlier in the sentence; to assume that different resources are used seems unwieldy. The corollary is that the neural underpinnings of these processes can be assumed to be the same as if the process had occurred earlier. This does not mean that they might not differ from similar processes carried out earlier in the sentence. For example, their timing may be different under pressure of a completion deadline. The cases mentioned by Just and Carpenter mainly involve cases where no information was available in the clause to allow the process to take place, for example resolution of referential or lexical ambiguity. In such cases, we might speculate that the processes are carried out on the basis of less information, which might lead to some quantitative differences. The evidence for this sort of process is discussed in Section 3.3.

Before beginning the history of what we know about "wrap-up", it is possibly even more important to recognize what it is ***not***. Some authors seem to include decision processes, as in acceptability or grammaticality judgments, with sentence wrap-up effects, but this confuses two separate concepts. Decisions about sentence acceptability, plausibility or grammaticality clearly do affect final words differently than earlier words, because the decision becomes definitive at this position. This sort of effect can be seen in behavioral measures as well as ERPs. For example, Kuperberg, Kreher, Goff, McGuire, and David (2006) report that when participants carried out a post-sentence acceptability judgment task, the final word reading times for *acceptable* sentences were longer than for sentences with a semantic or syntactic incongruity earlier in the sentence. Presumably this is because the decision still had to be made for this condition, but not for the sentences containing an earlier unacceptability. For this reason, one should indeed be suspicious of the generality of effects at the final word of the sentence if a decision task is being carried out. The obvious solution is to avoid using this sort of task when interested in sentence-final words. Most ERP effects can be seen even without a judgment task (e.g. with a simple question to engage comprehension or a relatedness judgment on a word presented after the sentence instead), although the effects may be slightly smaller. A very clear confound with final position is easily avoided in this manner.

3. A history of "sentence wrap-up" effects and their interpretation

There is substantial behavioral evidence that *something* occurs at the end of the sentence or clause. An early form of evidence is that verbatim memory declines very quickly after the end of the sentence (Jarvella, 1971; Kintsch, Welsch, Schmalhofer, & Zimny, 1990). The evidence from Jarvella's study is particularly striking. He showed that verbatim memory for words in the preceding pair of sentences dropped approximately 30% across a sentence boundary, even when the number of intervening words was held constant. This suggests that the literal information from the clause can be dismissed in favor of a higher level representation from the end of the sentence. Although this counts as evidence for sentence wrap-up, it should be noted that a similar discontinuity was seen at a clause boundary, particularly if a stricter definition of recall was used (p. 412). This raises the issue of whether the term "sentence wrap-up" is accurate.

Evidence from measures of reading time provide the most pervasive evidence that *something* interesting occurs at sentence and clause boundaries. Self-paced reading typically shows longer reading times at the end of the sentence (e.g. Just, Carpenter, & Woolley, 1982; Mitchell & Green, 1978) and the end of the clause (Hill & Murray, 2000). Reading studies making use of eye-tracking technique shows the same pattern: increased fixation times at the end of the clause or sentence (Rayner, Kambe, & Duffy, 2000; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989). Eye-movement studies also show more regressions and longer saccades from clause-final than non-final words (Camblin, Gordon, & Swaab, 2007; Just & Carpenter, 1978; Rayner, 1975; Rayner et al., 2000), which suggests that the clause is treated as a unit which is