The Sentence Wrap-up Dogma

Laurie A. Stowe[1]

Edith Kaan[2*], Laura Sabourin[3] and Ryan C. Taylor[4]

[1]Rijksuniversiteit Groningen; [2]University of Florida; [3]University of Ottawa; [4]University of British Columbia

*Corresponding author:

Edith Kaan

Department of Linguistics

University of Florida

Box 115454,

Gainesville FL 32611

United States of America

kaan@ufl.edu

1

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 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.

5

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

6

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, Sereno, Morris, Schmauder & Clifton, 1989; Rayner, Kambe, & Duffy, 2000). Eye

movement studies also show more regressions and longer saccades from clause-final than non-final words (Just & Carpenter, 1978; Camblin, Gordon & Swaab, 2007; Rayner, 1975; Rayner et al., 2000), which suggests that the clause is treated as a unit which is preferably not broken up during reading. Note that these effects are not limited to ends of sentences; clause boundaries show similar effects, although the magnitude of the effects are smaller.[1]

The most obvious interpretation for longer reading and fixation times is that there are extra processes at this point, which involve sentence wrap-up. The idea has an obvious intuitive appeal, to the extent that the existence of these effects has constituted perhaps the major argument for such processes in and of itself, without any need to show what the proposed processes *do*. Nevertheless, to prove that sentence wrap-up exists, it is necessary to demonstrate just what processes are involved by determining what factors affect the size of the clause-final lengthening of reading times. In the following four sections evidence for and against the most frequently discussed conceptions of the end-of-clause effects is summarized. We will appeal to both behavioral and ERP evidence in the discussion, which may seem counter-intuitive. Generally the ERP evidence comes later in time than similar behavioral evidence, but since the goal of this article is to investigate whether there are ERP effects that are limited to the end of the sentence, ERP evidence for and against the various views of sentence wrap-up will be given some priority in the discussion.

Unfortunately, the concept of sentence wrap-up is appealing enough as an explanation of sentence-final effects (i.e., longer RTs on the final word in self-paced reading and longer final fixation times in eye movement studies) that people

---

[1] The evidence for **clause-final** wrap-up invalidates a large portion of the sentence-final avoidance in ERP designs. In many experiments that are conducted in languages like German where verbs are obligatorily clause-final, sentence-final position is avoided by adding a subordinate clause after the target, but the target remains **clause**-final. The clause finality is frequently even explicitly marked by a comma, e.g. …*dass Maria Sängerinnen besucht, obwohl*… (litt:'that Maria singers visits, although….' "that Maria visits singers, although…"), where *besucht* is the word position of interest (Bornkessel, McElree, Schlesewsky & Friederici, 2004).

frequently refer to the effect itself as sentence wrap-up. We will use the term *sentence-final effects* throughout to refer to the phenomenon in order to separate it clearly from the theoretical construct *sentence wrap-up* which refers to the processes hypothesized to occur at the end of the sentence.

3.1   The clause as a unit in syntactic processing

The earliest studies of which we are aware that suggest special processes at the ends of clauses were carried out with click detection (Fodor & Bever, 1965; Garrett, Bever, & Fodor, 1966; Bever, Lackner & Kirk, 1969). Clicks presented near to clause boundaries were misremembered as being at the boundary instead, suggesting that clauses have a special status. Bever, Lackner and Kirk (1969) additionally reported that clicks were not mislocated to boundaries between units within a clause. At the time the importance of clause boundaries was thought to be due to the necessity of reconstructing the deep structure of the clause (Miller & Chomsky, 1963), in order to determine both the syntactic structure and the meaning of the sentence. In this view of syntactic structure, the interpretation of an element depended on its position in a deep structure in which the element was in its basic position. For example, "which book" can only be interpreted once it is clear whether it is the subject of the underlying sentence (1a), object of the main clause (1b), object of a preposition (1c) or some other underlying position.

1) a.   Which book was lying on the table?

A book was lying on the table

b.   Which book did Sharon give to Frans for Christmas?

Sharon gave a book to Frans for Christmas

c.   Which book did Sharon argue with Frans about?

Sharon argued with Frans about a book.

Since the deep structure depended on the entire clause, it could only be reconstructed completely as this point. Thus, clause boundary effects were taken as

9

evidence for the delayed reconstruction of the syntactic structure; to quote Fodor, Garrett and Bever (1968; p. 460): "…the clause […] provides the perceptual unit in speech (see Bever, Fodor, & Garrett, 1966). If this view is correct, it argues for the existence of a level of processing which provides just the sort of pre-analysis of the input string presupposed by the view of deep structure recovery we have been presenting."

The hypothesis that sentence-level structure building and interpretation is delayed until the end of the clause has been more or less universally rejected since that time. On-line techniques such as self-paced reading, eye tracking and ERPs have amply demonstrated that processing at multiple levels occurs immediately. Each of these methods is sensitive throughout the sentence to:[2]

● morphological processes (e.g. Bertram, Hyönä & Laine, 2000; Osterhout & Mobley, 1995)

● immediate syntactic integration, as seen in detection of ungrammaticalities, including morphosyntactic agreement processes across considerable structural distance (Friederici, 2002; Kaan, 2002; Osterhout & Holcomb, 1992)

● immediate resolution of structural ambiguity (Frazier & Rayner, 1982; Osterhout, Holcomb & Swinney, 1994; Stowe, 1986, 1991)

● anticipation of upcoming words or structures on the basis of semantic (Altmann & Kamide, 1999; DeLong, Urbach & Kutas, 2005; Van Berkum, Brown, Zwitserlood, Kooijman & Hagoort, 2005), morphosyntactic (Kamide, Scheepers, & Altmann, 2003; Wicha, Moreno & Kutas, 2004), or prosodic information (Ito & Speer, 2008; Nakamura, Arai & Mazuka, 2012).

One might consider an alternative in which local structure building occurs on-line, with the processes involved in establishing deep structure relations delayed until the end of the clause. Even this restricted type of sentence structure building as a

---

[2] The evidence for fast incremental processing at all levels of sentence comprehension is so overwhelming that it is difficult to decide what would count as the most basic references; those provided serve purely as examples.

process limited to sentence wrap-up is rendered untenable by the evidence for immediate parsing of WH-constructions (Garnsey, Tanenhaus & Chapman, 1989; Phillips, Kazanina & Abada, 2005; Stowe, 1986), since WH-movement is the paradigmatic example of a deep structure relationship (see examples in (1)).

The clause-final syntactic processing view also more or less entails that on-line lexical semantic integration into the context should also be delayed, since the deep structure gives evidence about how the words are to be combined. However, lexical integration also proceeds incrementally across the sentence, as reflected for example by effects of cloze probability, contextual plausibility, and lexical ambiguity resolution (Kutas, Lindamood & Hillyard, 1984; Hagoort & Brown, 1994). The evidence for immediate and incremental parsing and comprehension makes the original idea that syntactic structure building is delayed until the end of the sentence, and therefore that sentence-final effects reflect primarily a stage of syntactic parsing, untenable.[3] This of course does not imply that readers and listeners always build a complete and detailed syntactic analysis. Interpretations are often underspecified or "good enough", depending on the reader's or listener's goals (Ferreira & Patson, 2007; Swets, Desmet, Clifton, & Ferreira, 2008). However, the evidence that detailed syntactic analyses *can* occur immediately is problematic for the clause-final sentence processing view.

3.2   The clause end as a locus for text integration

A second claim that has been adopted by a number of researchers is that processes which build inter-sentential coherence are primarily or solely carried out at the end of the clause. This suggestion is frequently attributed to Just and Carpenter (1980), who wrote that sentence wrap-up includes "search for referents that have not been

---

[3] Although this concept of sentence wrap-up is rather old-fashioned, we include it because at least Van Petten and Kutas (1991) suggest something of the sort, based on Friedman, Simson, Ritter & Rapin (1975), which will be discussed in Section 4.2.

assigned, the constructions of inter-clause relations (with the aid of inferences, if necessary), and an attempt to handle any inconsistencies that could not be resolved within the sentence" (p. 345). The suggestion that discourse processes are generally delayed until the end of the sentence is also related to the idea that propositions are based on clauses and thus, by extension, require the whole clause to be complete to be constructed. Van Dijk and Kintsch (1983) discuss several conceptions of propositions and adopt this form of meaning representation as the basis of making links in discourse[4]: they refer to "the strategies used to assign coherence relations between propositions in sequences, typically those underlying subsequent sentences" (p. 130).

It would be far too strong to claim that there is no realm of text and discourse processes which only elicits effects at the end of the clause; the field is too broad and the variety of processes involved in discourse processing are varied as well. For example, elaborative inferences, like adding a probable cause or consequence which has not explicitly been mentioned ("step on broken glass" → cut foot), may not be made at the first opportunity (e.g. *glass* in the example above). Indeed it has been claimed that they are not routinely made at all, as they are not necessary for coherence (e.g. Perfetti & Adlof, 2012). Even if true, however, this claim is independent of the current debate, since the end of the sentence does not have a special status either. While recognizing the limitations of the evidence about when inferences occur, the claim that discourse integration processes are by and large delayed until the end of the sentence seems to be incorrect.

A number of studies show effects of discourse context on comprehension well before the end of the sentence or clause, generally as soon as the relevant information becomes available. This has been demonstrated for lexical integration

---

[4] It should be noted that Van Dijk and Kintsch's use of proposition includes also atomic propositions (equivalent to "there is an x" for a noun and "x is young" for a modified noun) which clearly can be created before the end of the clause, and which could serve as the basis for referential bridging inferences. Complex propositions built of these atomic propositions are largely assumed to be equivalent to clauses, however.

into a discourse context (Van Berkum, Hagoort & Brown, 1999; Van Berkum, Zwitserlood, Hagoor & Brown, 2003), co-reference (e.g., Van Berkum, Koornneef, Otten & Nieuwland, 2007) and establishment of instrumental bridging inferences (e.g., Garrod & Terras, 2000). Van Berkum et al. (1999) examined sentence position explicitly, testing sentences containing a word which was neutral out of discourse context (e.g. *He had been very slow that morning*).[5] An extra-sentential context supporting the expectation of *quick* instead led to an immediate ERP response, whether the word was sentence-final or sentence medial (see Figure 5 and discussion on p. 665).

All the studies which we have found cited in the ERP literature as supporting sentence-final discourse processes are less clear-cut than they seem. For example, establishing co-reference, which Just and Carpenter (1980) cite as one of the processes that occurs in sentence wrap-up, clearly starts as soon as a referent is presented. This can be seen from effects on pronouns, both when no referent is available and when the referent of the pronoun is ambiguous. ERP studies show immediate responses to pronouns with no referent in the sentential or discourse context, as in *The old woman said that he…* (at least, in a sentence acceptability judgement task, Osterhout & Mobley, 1995). When there are several possible referents between which the comprehender can choose, there is an immediate frontal negativity, called the nRef, relative to cases in which the referent is unambiguous (Van Berkum et al., 2007). Self-paced reading and eye movement reading studies show evidence for immediate processing effects as well. For example, Ehrlich and Rayner (1983) found evidence that the first several fixations after a pronoun with a distant referent were longer than when the referent immediately preceded the pronoun; the fixations did not include the end of the sentence in their examples.

---

[5] Throughout we will use underscore in examples to indicate the target word(s) for the analyses to be discussed.

In a similar vein, McDonald and MacWhinney (1995) investigated interpretation of a pronominal subject of a second clause. They presented sentences auditorily with visually presented probe words at certain time points. Participants had to determine at these points whether they had heard the probe words or not. When implicit verb bias encourages co-reference with the object of the first clause (e.g. *John admired Steven because he...* which encourages people to assume that an admirable quality of Steven is about to be mentioned), as opposed to co-reference with the first clause subject, which is a default preference in two clause sentences, effects were seen at both the pronoun itself and at the final word in the sentence. Even the effects at the final word in this study do not really support a special role for this position, however, as the final word provides relevant information for confirmation or disconfirmation of the object co-reference interpretation in the sentences used in this experiment. In the continuation *because he is very stupid*, for example, *he* presumably refers to John, since stupidity is seldom considered admirable. Thus this experiment cannot clarify whether inference-related processing occurred because this is the final word or because inference-relevant information occurs at this position. Given the discourse integration effects earlier in the clause, inference making is in any case not limited to the end of the sentence, as suggested by a model which assumes that distinct processes take place at this point. Similar problems can be cited for other studies that report discourse effects at the end of the sentence (e.g. Just & Carpenter, 1978, for referential inferences; Pexman, Ferretti & Katz, 2000, for recognition of irony).

The summary to this point suggests that at least bridging inferences and referential inferences necessary for establishing coherence between clauses normally start as soon as information becomes available. Although the process may not be completed until later relevant information is also processed, it is not clear from any of these studies that this only occurs at the end of the sentence. There does not seem to be an obvious role for specifically sentence-final text-discourse integration processes here. If there are such differences between sentence-medial

14

processes and those at the end of the sentence, it should be possible to find an interaction between position and the presence of inference-relevant information.

In fact, one would predict that the investigation of such interactions would become a hive of research intended to determine what the differences are. Instead, there seem to be a small handful of studies which report such a direct comparison. Rayner et al. (2000) demonstrate that the same word is read for a longer time when it is clause-final than when it is not final (position-related lengthening), but this did not interact with their manipulation of discourse integration difficulty, which consisted of a difference between good and poor exemplars (e.g. *football* vs. *curling*) of a category (e.g. *sport*) which was used to refer back to it. The manipulation was strong enough to produce a main effect, but did not interact with the position effect, suggesting that an equivalent reference integration takes place in both positions, rather than something special at the end of the clause.

Kuperberg, Paczynski and Ditman (2011) carried out the only study of which we are aware which actually shows an interaction between the inference manipulation and position. They tested sentences like *On Monday she had <u>sunburn</u>/She had <u>sunburn</u> on Monday* in a short context that started with a sentence like *Jill had very fair skin*. In between were sentences that provided a cause (*forgot to put <u>sunscreen</u> on*), allowed a causal inference (*usually remembered to wear <u>sunscreen</u>*) or blocked a causal inference (*always put <u>sunscreen</u> on*). Both strongly cued and intermediately cued inferences showed evidence of being easier to integrate, in terms of a diminished N400 when *sunburn* occurred at the end of the sentence, but in the middle of the sentence only the strongly cued inferences showed this effect. As we pointed out earlier, there has been considerable debate over whether elaborative inferences are always made or only when this is necessary to maintain coherence. Here the cause of the sunburn is necessary for coherence, but the optionality of causal inferences may account for why position interacts with the strength of inference support in this study. Despite this one exception, specific discourse-related clause-final effects do not seem to be rampant, contrary to the supposition of the

15

hypothesis that this is a privileged position for discourse processes. Certainly they do not appear to be common enough to explain why sentence-final effects are so pervasive.

3.3   *Wrapping the clause up?*

If most clause-internal processing is carried out incrementally as the sentence is heard or read, the issue arises: what function does the extremely common behavioral clause-final effects serve, or to put it another way, what remains to wrap up? A number of researchers have been concerned about this issue over the years. Just and Carpenter (1980) suggest that there is a stage of processing which includes finalizing any incomplete processes. This might include a check of which processes are not complete and need to be finalized, although they do not explicitly mention such a process.

Taking this suggestion at a fairly general level, first, it predicts that the size of the sentence-final effect should correlate with (some) measures of complexity. Hirotani, Frazier and Rayner (2006) provide a thorough overview of the eye-tracking literature. In their own experiments, they found that the complexity of the sentence has no effect on the size of the clause or sentence length effects. In fact, the gaze duration is virtually the same following a vocative form (as in *John, …*) as following a full clause, although the syntactic structure and semantic complexity of the two must clearly differ considerably. They suggest that gaze time mimics prosody, rather than any form of sentence integration. We will return to this possibility below. In the same line, Hill and Murray (2000), although their goal was not specifically to examine wrap-up, demonstrate that the presence of a comma reliably elicits a brief lengthening which is not affected by the variables they manipulated (early vs. late closure), although later effects on eye-movements were found. They suggest that this reflects a low level perceptual process rather than higher level cognitive processing.

16

Hirotani et al.'s (2006) evidence is not incompatible with the claim that the increased processing load later in a sentence affects reading times. Warren, White and Reichle (2009) report that complexity of the preceding sentence has a measurable effect late in the clause; however, the size of the complexity effect is the same regardless of whether the target word is actually at the end of the clause or not. That is, complex is complex, whether clause-final or not, and is additive with the clause-final effect. Again, this suggests that clause-final processes are not specific to wrapping up the clause.

A hypothesis which does not necessarily assume that sentence wrap-up reflects sentence complexity in general is that there are more specific processes which cannot be completed immediately and these must be dealt with at the end of the clause. This possibility is most commonly attributed to Just and Carpenter (1980), who, as mentioned above, incorporated sentence wrap-up as one of the stages in their model of reading. Repeating the quote above for convenience, Just and Carpenter suggest that sentence wrap-up includes "search for referents that have not been assigned, the constructions of inter-clause relations (with the aid of inferences, if necessary), and an attempt to handle any inconsistencies that could not be resolved within the sentence" (p. 345).

As we saw in the preceding section, the first two processes, which relate to higher level textual integration processes, do not seem to be left to the end of the clause, but are dealt with immediately, even if sometimes incorrectly. In their last point, inconsistencies that have not been resolved, Just and Carpenter apparently refer to lexical ambiguity resolution, given the citations and evidence they provide. In this and several other articles they cite a study in which ambiguous lexical items are processed in contexts which do not support a choice of one meaning for the ambiguous word. The full version of one of the studies cited as supporting this claim is Carpenter and Daneman (1981). The study indeed shows longer reading times when the most likely meaning of the word (*tears* = water in eyes) must be revised to the less frequent meaning (*tears* = rip) due to later evidence from a word *dress*

17

which is inconsistent with the dominant meaning (e.g. *He saw tears in her brown dress*). The experiments reported in this article form a large part of the evidence for Just and Carpenter's sentence wrap-up, cited and recited through the literature.

Unfortunately, the word leading to revision was also the final word in the sentence, so that any sentence-final effect cannot be separated from immediate revision as soon as evidence becomes available. In a related study (Daneman & Carpenter, 1983) it appears that longer reading times occur as soon as inconsistent information appears, as in sentences like *There is a sewer near our home who makes terrific suits*. Clear effects of inconsistency are seen immediately after the word *who*, which suggests that the dominant meaning of sewer is not correct, not at the end of the sentence. The entire pattern of results (immediate preference for one resolution followed, if necessary, by a change of interpretation) is in line with the evidence cited above that semantic integration occurs immediately on-line. It is clear that comprehenders do not wait for the end of the clause to choose the meaning; if they did wait, the later information would not be inconsistent. Seidenberg, Tanenhaus, Leiman and Bienkowsky (1982) and Swinney (1979) both explicitly tested whether comprehenders wait until the clause or sentence is complete to decide on a meaning. They found clear evidence for a choice soon after presentation of the ambiguous word, even when no contextual information is available to support a particular choice and even when both meanings of the word are equi-frequent. Again, it seems that the evidence offered for wrap-up of the sentence does not really support the hypothesis when it is examined more closely.

3.4   Perceptual and prosodic accounts of sentence-final effects

Since effects related to language comprehension do not seem to explain most sentence-final effects, much recent work within the eye movement research community has focused more on perceptual and prosodic accounts of the time spent at various boundaries. Hirotani et al. (2006), for example, take the evidence

discussed in section 3.2 to be largely compatible with the hypothesis that the longer reading times reflect a covert prosodic boundary. In particular, they cite the fact that pause times are approximately as long when a comma follows a vocative, like *John, …* as when it follows a complex clause. Warren et al. (2004) report longer first fixation durations on nouns that are sentence-final and followed by a period, than on clause-final nouns followed by a comma. In turn, the latter had longer first fixations than non-final nouns. This pattern did not interact with the syntactic complexity of the sentence. This work is consistent with ERP evidence that suggests that commas elicit a centro-parietal positivity, which is a weaker form of the positivity found in response to prosodic phrase boundaries in auditory stimuli (Steinhauer, 2003; Steinhauer & Friederici, 2001). Other researchers have suggested other perceptual mechanisms to account for clause- and sentence-final effects (Hill & Murray, 2000), or a combination of perceptual and prosodic mechanisms, and integrative mechanisms (Warren et al., 2009). Again, none of these mechanisms can be claimed to be specific to final positions.

3.5   Unacceptability and the end of the clause

Before turning to the ERP literature, we will discuss one sentence-internal factor that *does* seem to influence sentence-final reading times specifically: the presence of anomaly earlier in the sentence. This factor was not discussed earlier because the discussion was restricted to those aspects of language comprehension which would affect normal well-formed sentences found in text, since these normal sentences show end-of-the-sentence lengthening of reading times which cannot be attributed to anomaly. Anomalies do, of course, occur in real life as well as in experiments. People make mistakes and the listener has to work around them. However, the anomalies experimenters make use of are not the bread-and-butter of daily communication. Braze, Shankweiler, Ni and Palumbo (2002) report that sentence-medial anomalies elicit responses which differ depending on the nature of the

anomaly. When the error is a syntactic error, readers regress from the point of the anomaly, but not later in the sentence. For semantic errors, there are relatively few immediate regressions, but a large proportion of regressions are initiated from the last phrase in the sentence. In a sense, this seems to be a selective sentence wrap-up effect for semantically incorrect sentences (wait for all the data and see if it can be fixed), which is not applied to syntactic errors. Readers attempt to fix those immediately instead.

It is important to note that subjects were not carrying out a grammaticality or acceptability judgment task in this experiment; they simply answered content questions, so this pattern appears to result when readers attempt to understand sentences like these. De Vincenzi et al. (2003) show similar effects for self-paced reading; other studies have shown some variability regarding which kind of syntactic and semantic information is noticed and definitely dealt with immediately and which is not (e.g., Boland, 2004). Unfortunately most of these studies were not particularly concerned with end-of-sentence effects. These studies use designs similar to the ERP studies which will be discussed below; the Braze et al. (2002) and De Vincenzi et al. (2003) studies were in fact based on an ERP study, so this pattern of results is interesting for the discussion of ERP studies and wrap-up below.

3.6   Summary: views of "wrap-up"

To summarize, historically, "wrap-up" has been attributed to syntactic construction being clause-based, the integration of information within or across sentences, completion of syntactic analyses, prosodic or perceptual processes, or a combination thereof. We have argued that none of these proposed mechanisms are either tenable or unique to a clause or sentence final position. This leaves the account of the sentence-final negativity observed in the ERPs for sentences that contain a mid-sentence anomaly. We will discuss this in detail in the next section.

4. ERPs and sentence wrap-up

The goal of this section is to present the elements of the sentence wrap-up dogma as it developed over time. This history is not presented entirely chronologically, as each of the successive claims about ERP and sentence wrap-up will be evaluated before going to the next claim, and the sources of evidence concerning different elements of the argumentation overlap in time. Most importantly, the experiments which are cited as providing strong support for the sentence wrap-up effect will be presented last, rather than in their proper time. The reason for this is to allow the reader to evaluate the strength of the evidence provided, not as it seemed at the time when it was originally presented, but in relation to the other evidence which is now available.

Despite the evidence that sentence wrap-up does not form a clearly differentiable stage of processing found in every sentence, as initially claimed by Just and Carpenter (1980), the basic concept has received a great deal of credence in the ERP community, based primarily on the intuitive appeal of wrap-up as an explanation for sentence-final effects. There are mainly two effects that have been claimed to reflect some form of sentence wrap-up: a sentence-final positivity, and a sentence-final negativity.

The curious point about references to sentence wrap-up in the ERP literature is that this has not been a locus of research in itself. It is discussed in terms of avoiding the sentence-final position to prevent contamination by wrap-up, which makes it difficult to assess to what extent any effects reported at the end of the sentence reflect processes that can be characterized as wrap-up. In this section we will describe the evidence for a sentence-final positivity and a sentence-final negativity, both of which unquestionably exist. More centrally, we will discuss the limitations of the frequently cited evidence for interaction between these effects and ERP responses to language processing, which parallels the lack of evidence for a

special status of sentence-final position for language processing in the behavioral literature.

4.1   Sentence-final positivity

When discussing sentence wrap-up effects, reference is sometimes made to an increased positivity which is elicited by the final word of sentences. This effect is obvious when multi-word ERPs are reported; it can be seen for example in Kutas and Hillyard (1980), see Figure 1. Note that the sentence-final word, whether it shows a font-size related P560 or not, has a considerably more positive movement in the waveform than the preceding words. The extent of the positivity is easily seen in the peak to peak measure, indicated by the arrows next to the penultimate and the ultimate word.

[insert Figure 1 around here]

The positivity has a centro-parietal distribution, and typically starts in the N400 time window (300-500ms), and is long-lasting. This positivity might be the ERP equivalent of the increased reading/fixation times found in the behavioral literature. If so, we would expect the positivity to respond to the same manipulations as in eye tracking and self-paced reading—not much, in other words. Alternatively it may be what it has been claimed to be, a reflection of sentence wrap-up processes. In that case, it ought to show a clear response to complexity of semantic or syntactic processing earlier in the sentence or to discourse variables of some sort.

The first report of the sentence-final positivity is attributed to Friedman, Simson, Ritter and Rapin (1975). They presented spoken words with spaces between them, to investigate how context is used to "fill in" a missing sound (e.g. *The –eel was on the shoe*); the ability to do this is called the phoneme restoration effect. The ERP to the final word was more positive than to other words in the sentence, even in control conditions without phoneme restoration. There are clearly multiple

22

possibilities for interpretation of this effect. Friedman et al. themselves attribute this effect to completion of a syntactic structure, since the positivity occurred regardless of whether the sentence-final word was meaningful or allowed phoneme restoration to occur (cf. Van Petten and Kutas, 1991). Participants also had to indicate which word was intended, from a limited set of possibilities (*heel, wheel, peel*), so a decision was made and a response was prepared at the sentence-final point as well.

It is important to determine exactly when a sentence-final positivity occurs and what variables influence its onset and/or amplitude, to determine what sort of process it reflects. As we pointed out in Section 1, different ideas of what wrap-up involves make quite different predictions of how the wrap-up processes will impact other processes occurring at the end of the sentence. The evidence relevant to deciding what neuronal process the sentence-final positivity reflects will be addressed in the next section.

4.2   *Sentence-final positivity and sentence comprehension*

The primary evidence that this positivity is a phenomenon related to sentences or clauses comes from a comparison between sentences and word strings carried out by Van Petten and Kutas (1991). They compared final words in word strings, normal coherent sentences and syntactic sentences (in which each syntactic position is randomly filled by a word for the appropriate category). Participants were asked to indicate whether a probe presented after the sentence had appeared in the sentence. Sentences show the sentence-final positivity compared to word strings (Figure 5, p. 104). This comparison provides evidence that this is a sentence- or clause-final positivity, rather than a string-final positivity.

As with behavioral measures, if the sentence-final positivity reflects sentence wrap-up processes, we would expect it to respond to some of the factors which make it more difficult to construct the syntactic or semantic representation of a sentence. With regard to semantic processing, the sentence-final positivity appears

to be purely additive. In a between-item study examining the effects of cloze probability in sentence medial and sentence-final positions, Kutas and Hillyard (1983) show that the positivity rides on top of the N400 effect elicited by unexpected or anomalous words, which was also visible in Figure 1 above (see their Table 1, p. 542). This suggests that increased semantic integration costs do not impact whatever process is reflected in the sentence-final positivity.

Additional evidence to this effect is provided by Hagoort and Brown (2000). Using auditory materials with no task other than listening, they tested the effects elicited by sentence-final semantic anomalies and sentence-internal versions of the same anomaly formed by adding material after the target word. This comparison was made precisely because of the possibility of sentence wrap-up effects in auditory presentation, since intonation would provide a strong cue to the location of the end of the clause. They found a similar scalp distribution and amplitude for the N400 effects, independent of clause position (see their Figures 2 and 4). The sentence-final positivity also appears to be present as a main effect, but Hagoort and Brown's comparison was made across experiments, and no direct statistical comparison was made. In fact, the sentence-final positivity was not discussed at all. Kuperberg, Sitnikova, Caplan and Holcomb (2003), using visual presentation and a judgment task, also found evidence for a sentence-final positivity starting in the N400 time window and continuing into the P600 time window. They examined the effects of violating thematic role constraints, which elicited a P600 effect, and pragmatic constraints, which elicited an N400 effect. The N400 did not interact with sentence position, nor did the P600; this can be seen in their Figure 2 (p. 123).

Obviously there is a much larger literature out there, and other semantic or pragmatic factors may cause alteration in the sentence-final positivity. We will return to some of this literature in the next section. Nevertheless, the prospects for an account of the final positivity as a sign of sentential *semantic* wrap-up seem pretty slim. Van Petten and Kutas (1991) examined one of the most extreme semantic differences possible. They examined word position effects over two kinds of

24

sentences as well as the word strings mentioned in the last section: high cloze congruent sentences and syntactic prose sentences (grammatical strings where every open class word is randomly filled in with a word of the same class, producing grammatical nonsense). Predictably, the N400 effect is much more negative for nonsense than for highly predictable sentence endings. However, as in the Kutas and Hillyard (1983) study, after the N400 is complete, fully congruent sentences converge with the nonsensical sentences in the form of a sentence-final positivity, but in this case, unlike the low cloze sentences Kutas and Hillyard used, readers will most likely fail to construct the meaning of the sentence, since the entire sentence contains words with no coherent message. The convergence of the positivities can be clearly seen in Van Petten and Kutas' Figure 5 (p. 104), in which the final words of both sorts of strings show a very similar positivity once the divergence in waveforms in the time window of the N400 (approx. 300-500 msec) is complete.

This leaves syntactic wrap-up as a possible explanation for the positivity. The fact that Van Petten and Kutas's (1991) sentences differed from word lists would be compatible with an explanation involving syntactic structure, and they do indeed assume this type of explanation. Their conclusion forms one of the bases of the sentence wrap-up dogma. However, there is no evidence that we are aware of suggesting that the syntactic complexity of the preceding sentence affects the amplitude or latency of the sentence-final positivity, which would provide strong support of a syntactic wrap-up effect. The weight of the evidence comes from studies involving syntactic violations. These studies will be discussed separately in Sections 4.3-6, as most of the evidence regarding syntactic violations has had to do with a sentence-final negativity, rather than the positivity.

4.3   Sentence-final negativity

Although the sentence-final positivity is sometimes cited as evidence for sentence wrap-up, the sentence-final negativity is the effect which is most frequently

associated with the ERP sentence wrap-up dogma. The following sections discuss the stages of the development of the dogma, not all of which are taken into account in its current version.

This development began in 1992, when Osterhout and Holcomb (1992) reported on ERPs elicited by sequences like *The broker hoped/persuaded to sell the stock.,* contrasting a condition with a temporarily ambiguous sentence containing a reduced relative (*persuaded*) and an unambiguous main clause control (*hoped*), see Table 1 for an overview of target sentences and word positions within the sentences. Evidence from ERPs time-locked to *to,* which indicates that the *persuade* variant cannot be a well-formed main clause, elicited the now-classic P600 effect. This effect can readily be interpreted as a response to the ill-formedness of the preferred main clause parse when the verb is obligatorily transitive; its presence suggests that participants noticed at this point that the sentence was not a grammatical main clause and revised the structure. The results were of great interest, when taken together with the results of Neville, Nicol, Barss, Forster and Garrett (1991), because the P600 appeared to be typically present for syntactic difficulties and ill-formedness, exactly as the N400 is typically present for semantic difficulties and ill-formedness. If it indeed were the case that manipulations of syntactic difficulty affect an entirely different ERP component than manipulations of semantic difficulty it would provide strong evidence for the psycholinguistic view that semantic and syntactic information are processed in quite different ways.

Given the desirability of this conclusion, it was a disappointment that the end of syntactically unacceptable sentences in the Osterhout and Holcomb study sometimes showed a negativity resembling the N400 instead of the P600 effect. In *The broker persuaded to sell the stock.,* the sentence is incomplete, because there is no main clause verb phrase. Rather than a positivity, similar to the one found at *to*, there was a negativity. Osterhout and Holcomb also presented versions of the materials like *The broker hoped/persuaded to sell the stock was sent to jail.* In these, the variant with *hope*, which contains an extra, unaccounted-for verb phrase, also

elicited a sentence-final negativity relative to the grammatical version with *persuade*. Taken together, these results suggest that a negativity is associated with ungrammaticality in the sentence-final position in some way. This conclusion was further supported by the fact that *stock* elicited identical effects in the *hoped* and *persuaded* variants when it was not a sentence-final word, as in the longer sentences above. Osterhout and Holcomb (1993) report a replication of the 1992 study using auditory presentation, which showed very similar results.

Table 1

Summary of Effects in Osterhout and Holcomb (1992, 1993) and Osterhout, Holcomb and Swinney (1994)

Stimulus	Control word	Positivity/P600 present?	Negativity/N400 present?
Osterhout and Holcomb (1992,1993)			
The broker persuaded <u>to</u> sell the stock.[1]	hope	√	x
The broker persuaded to sell the <u>stock.</u>	hope	x	√
The broker hoped to sell the <u>stock</u> was sent to jail.	persuade	x	x
The broker hoped to sell the stock <u>was</u> sent to jail.	persuade	√	x
The broker hoped to sell the stock was sent to <u>jail.</u>	persuade	x	√
Osterhout, Holcomb and Swinney (1994)			
The doctor charged/forced the patient <u>was</u> lying.	hoped	√	x
The doctor forced the patient was <u>lying.</u>	charged/hoped	x	√

[1] Effects found at specific target words marked with an underscore, relative to sentence containing the control word in the first

verb position (i.e. *hope ←→ persuade*).

The authors concluded that their sentence-final negativities should be classified as N400 effects, given their scalp distribution and latency of onset. They suggested that the N400 effects were caused by processing difficulties at the message level, rather than by difficulties in syntactic processing, due to the sentence-final position. This explanation preserved the distinction between syntactic and semantic processing effects in ERPs. However, the negativity went on for longer than the typical N400 time window, and the authors also speculated that the later portion of the negativity might reflect a blocking of the sentence-final positivity.

If this characterization is correct, it would support the idea that the sentence-final positivity responds to sentential factors like well-formedness and is indeed a sentence wrap-up effect, contrary to the conclusion drawn at the end of section 4.2. This would be an important conclusion, since it would support the concept of a sentence wrap-up effect that affects the final position. However, it is useful to consider the circumstances under which the sentence-final N400 or the late negativity (if they are actually separable) occur, as well as evidence about when they do **not** occur before accepting this conclusion. This issue will come up again later in Sections 4.4.2 and 4.4.3.

The first type of sentence which elicited a sentence-final negativity in Table 1, where the sentence is incomplete, is the odd man out. This type of structure has not received much attention. The second type of sentence-final negativity, a negativity following an ungrammaticality earlier in the sentence, which itself elicited a P600, has been replicated very frequently. This effect has been found under at least the following circumstances (earlier ungrammatical element which elicited a P600 is marked with an asterisk):

● At the end of a sentence containing an extra verb phrase (Osterhout and Holcomb, 1992, 1993; Osterhout et al., 1994: *The broker hoped to sell the stock *was sent to <u>jail.</u>)*

● At the end of a sentence following a verb that has no grammatical attachment site (Osterhout, et al., 1994; *The doctor forced the patient *was <u>lying.</u>)*

29

- At the end of a sentence containing an earlier phrase structure order violation (Hagoort Brown & Groothusen, 1993: *…the emotional *somewhat reaction of his wife.*)

- At the end of a sentence containing an earlier subject verb agreement violation (Hagoort et al., 1993; De Vincenzi et al., 2003; Osterhout and Mobley, 1995: *The men *hopes to succeed.*)

- At the end of a sentence containing an earlier auxiliary main verb mismatch (Osterhout and Nicol, 1999; Ditman, Holcomb, & Kuperberg, 2007: *The cats won't *eating the food on the porch.*)

- At the end of a sentence containing an earlier reflexive number or gender agreement error (Osterhout and Mobley, 1995: *The man helped *themselves/herself to the food.*)

- At the end of a sentence containing an earlier grammatical gender violation (Molinaro, Vespigiani and Job, 2008: *La vecchina con *il/*la scialle cammina lentamente per la salita.* Trans: The old-woman with an[wrong phonological form]/la[wrong gender] shawl walks slowly up the slope)

- At the end of a sentence containing a semantic error (Bohan, Leuthold, Hijikata Sanford, 2012; De Vincenzi et al., 2003; Osterhout and Nicol, 1999: *The cats won't bake the food on the porch.*)

The negativity in these studies is typically observed starting 300-500ms after onset of the last word, although it can appear at earlier word positions (see Hagoort et al., 1993; De Vincenzi et al., 2003, see section 4.4.3). It often lasts until the end of the epoch recorded. The negativity either is broadly distributed or has a central/parietal maximum, except when the syntactic violation is on the penultimate word (Osterhout et al., 1994); in the latter case the negativity has more of an anterior distribution. This convergence on the existence of a sentence-final negativity in violation sentences is clear and unarguable. What most of these sentence-final negativities have in common is that they are at the end of a sentence containing an earlier ungrammaticality. This should bring to mind the evidence cited earlier that

sentence-final lengthening responds in size depending on an earlier problem in the sentence discussed in Section 3.5. There are also studies that have failed to find the sentence-final negativity following syntactic errors. We will return to both these points in section 4.4.3, where we will attempt to characterize what does and does not happen at the end of the sentence in ERP experiments.

For the minute, it is important to remember that in general sentence wrap-up is assumed to be a general phenomenon, not one that occurs only when a sentence does not make sense or is ill-formed. From this viewpoint, sentence-final negativity is only a "real" wrap-up effect if it can be related to processing of well-formed sentences in some way. To put it a different way, a concept of sentence wrap-up that is limited to ungrammatical sentences is a totally different beast than the notion of wrap-up that is consistently appealed to in the ERP literature.

4.4   Sentence-final position as a confounding factor

The combination of one fact and one undemonstrated possibility led to the idea that sentence-final position acts as or *could* act as a confound in ERP experimental design. The fact is that there is a sentence-final negativity associated with (some sorts of) unacceptable sentences. The undemonstrated possibility was that this component would also be present when the word which introduces the ungrammaticality is presented at the end of the sentence, leading to qualitatively different effects than on intermediate words in the sentence.

The speculation that the final position in the sentence is problematic has several early articulations, with no supporting data further than those in Table 1, although they are frequently used in citations as if they offer much firmer evidence than is actually the case. When Osterhout and Holcomb (1995) discussed negativities which had been reported as responses to ungrammatical sequences (two word sequences in the case of Münte, Heinze, and Mangun, 1993; sentences in the case of Rösler, Putz, Friederici, and Hahne, 1993), they point out: "But perhaps most

31

importantly, the critical word in both studies was the *last stimulus presented* [original authors' emphasis] prior to the subject's response, and in the Rösler et al. study the critical word appeared in sentence-final position. There are a number of reasons for suspecting that such a placement introduces the possibility of confounding the response to the anomaly with sentence wrap-up, decision, and response factors" (p. 196). Note that as in many versions of the dogma, several potential effects of the sentence-final position were named, not just wrap-up. Some of these can easily be avoided by not using a task requiring a decision at the end of the sentence. That leaves a single study, Rösler et al. (1993), as potentially providing evidence for the sentence-final confound.

4.4.1   The sentence-final negativity does not replace the P600

The citation in the previous paragraph, like other early citations, concentrates on syntactic violations that elicit *only* negativities when tested in sentence-final position. Even Osterhout, Allen, McLaughlin and Inoue (2002) formulate the problem with the end of the sentence in this manner. "For example, most studies reporting an anterior negativity, rather than a P600-like positivity, have placed the anomalous word at the end of the sentence. This potentially confounds the effects of the anomaly with end-of-sentence wrap-up effects (see Osterhout, 1997, for evidence in support of this possibility)" (p. 1310). This formulation suggests that all sentence-final ungrammaticalities should elicit a negativity with no P600, as they would have in medial positions.

However, that prediction was not born out. A negativity *instead* of a positivity on a sentence-final ungrammatical word is certainly not common and is limited, as far as we know, to the single sentence-final study cited by Osterhout and Holcomb (1995) in the passage above. Instead a large number of studies have reported a negativity *in addition* to a P600 in sentence-final position. Probably the best-known study is that of Friederici, Pfeiffer and Hahne (1993) and its various replications. Friederici et al. (1993) reported a biphasic negative-positive response to auxiliary-

main verb mismatches (e.g. …*wurde bohnere*, "was *polish [infinitive]") as well as to the better-known syntactic category violations (…*wurde im gegessen*, "was in the *eaten"). This far from the only instance, however; below is a summary of a number of studies which showed a biphasic negative positive response when ungrammaticalities were presented at the end of the sentence.

- Phrase structure violation (Friederici et al., 1993, Hahne and Jescheniak, 2001; *Die Birne wurde im gepflückt.* "The pear was in the *picked")

- Subject verb agreement violation (Hinojosa, Martin-Loeches, Casado, Munoz and Rubia, 2003; *La prueba ocultada por el fiscal aparecí.* "The proof hidden by the public prosecutor *appeared [1st person]."[6]

- Nonfinite verb mismatch (Hahne and Friederici, 1999; Gunter, Stowe & Mulder, 1997; *De kleine drenkeling werd redden.* "The drowning child was *save.")

- Reflexive number and gender agreement error (Osterhout, 1997; *The man helped themselves/herself.*)

- Grammatical gender violation (Koelsch, Gunter, Wittfoth, Sammler, 2005; *Sie befährt den *Land.* "she travels through the[masc] country[neuter]")

These are the same sorts of syntactic violations which, when placed medially in the sentence, were followed by a sentence-final negativity, as can be seen in the list in section 4.3. None of these cases show only a negative response. The negativity seen in these studies did not typically appear to be an N400 as the effect frequently had a left and/or anterior distribution.

4.4.2 *The late negativity disappears if the violation occurs sentence-finally*

The evidence cited above demonstrates that the P600 is not absent at the end of the sentence. The same studies also demonstrate that there is no sign of a late sentence-final negativity either. Since multi-word ERPs are not available in these studies, this is

---

[6] This study is analogous to the others, but person agreement is manipulated, rather than number agreement. We see no reason to expect principled differences under the sentence wrap-up hypothesis however.

difficult to establish for sure in most of these studies. However, the waveforms themselves are more consistent with the presence of the sentence-final *positivity* instead. The presence of the P600 makes it difficult to quantify the sentence-final positivity as it might obscure a relative negativity, but it seems to be at least equivalent to the well-formed control conditions, since the P600 does not seem to be smaller than usual.

Fortunately, the size of the P600 can be tested. Two such studies were already discussed when evaluating the possibility that the sentence-final positivity reflects sentence wrap-up in section 4.2. In fact, even Osterhout (1997), one of the two primary citations for the sentence-final confound, noted that "syntactic anomalies in the present study elicited a robust P600-like positivity regardless of the position and word class of the anomalous word" (p. 514) and reported a main effect of position consistent with an equivalent sentence-final positivity whether the sentence is ungrammatical or not. This conclusion is supported by the study from Kuperberg et al. (2003) which was also described earlier. They demonstrated that violations which elicited a P600 medially, elicited just as large a P600 finally (see their Figure 2, p. 123). Since similar syntactic and semantic anomalies in the middle of the sentence are followed by a sentence-final negativity at the end of the sentence, the difference between these studies does not appear to be due to the type of anomaly. It is thus hard to maintain that the late negativity represents a blocking of wrap-up as reflected in the sentence-final positivity. This is unfortunate for the wrap-up argument, since an effect which is limited to ungrammatical sentences is quite possibly not a real wrap-up effect.

4.4.3 Sentence-final negativity is not even final

One necessary assumption of the proposal that the sentence-final (late) negativity represents a wrap-up effect is that the effect, like the sentence-final positivity, is, in fact, sentence-final. In fact there is little evidence that the negativity is limited to the final word. This impression is created by the fact that very few studies present multi-

word ERPs, which makes it difficult to tell when the negativity actually starts, but what evidence there is suggests that the negativity may start as soon as the P600 response to the sentence medial ungrammaticality is finished and may continue until the end of the sentence. There are two studies that do show the intervening period, Hagoort et al. (1993; see Figures 3 Exp 1, p. 457; Figure 9, Exp 3, p. 466) and De Vincenzi et al. (2003; see Figure 3, p. 288). Both studies show this pattern. The sentence-final negativity can probably best be interpreted as the end of a slow wave elicited by the original ungrammaticality, rather than a response to the final word per se. The results from Hagoort et al. (1993) make the point even more clearly. Hagoort et al. (1993) carried out analyses to examine the end of the sentence negativity, and found the effect was significant at the penultimate words as well as the final word (see their Figure 9, Exp 3, p. 466).

Whether or not we accept the concept of a slow negativity beginning immediately after the violation, the evidence that the negativity is already present earlier than the last word entails that it is not a modification of the sentence-final positivity. The sentence-final positivity is clearly elicited by the last word in the sentence, not the last several words in the sentence.

This slow wave negativity deserves further research. The functional significance of the long-lasting negativity deserves more investigation in its own right. Similar slow waves have been found when there is an additional memory load (Fiebach, Schlesewsky, & Friederici, 2001; King and Kutas, 1995; Ruchkin, Johnson, Canoune & Ritter, 1990), although these have typically had a centro-frontal scalp distribution; in this respect it is interesting that in fact De Vincenzi et al. (2003)'s slow wave was rather frontal from their figure. An ungrammaticality could well initiate such a long negativity since it cannot be stored in a fully analyzed form, but cannot be dismissed entirely as later input may give an indication of how the sentence can be correctly understood.

4.4.4   Task effects

Another potential explanation for the effect is suggested by the observation that the sentence-final negativity seems to be modulated by task effects. The vast majority of studies that reported sentence-final negativities for mid-sentence violations used an acceptability judgment task, in which participants were asked to give a response at the end of each sentence; the only exceptions are Hagoort et al. (1993) and De Vincenzi et al. (2003), which used no task, and comprehension questions after 10% of the sentences, respectively. Osterhout & Mobley (1995) tested sentences with medial violations in an acceptability judgment task (Experiments 1 and 2), and in a simple reading task (Experiment 3). The sentence-final negativity disappeared for all syntactic violations in the latter task. This suggests that the slow negativity may be related to maintaining information relevant to the decision task, or at least, that it is boosted by the decision task (see also Sabourin and Stowe, 2004 where an extended time window was analyzed). This slow negativity may need some time to develop: with the exception of Friederici et al. (1993) and Gunter et al. (1997), all experiments listed in 4.4.1 that tested sentence-final violations used acceptability judgment tasks, but none report a late negativity.

Regardless, Osterhout and Mobley's (1995) finding that the final negativity disappears when there is no decision strongly supports the full version of the sentence-final warning: avoid "sentence wrap-up, *decision, and response factors*". Fortunately, there is an easy fix to avoiding this problem: use a task that does not require a sentence-final decision. Numerous studies have shown that, although smaller, N400 and P600 effects can be reliably found for most effects without such a task.

4.4.5   Summary to this point

To sum up, it is clear that a syntactic error at the end of the sentence or clause does not produce only a negative response in the place of a positivity (4.4.1), except possibly under very specific conditions. Exactly what those conditions may be is at

36

this time unclear, although it may include incomplete sequences like those in the first example in Table 1. Moreover, all the evidence suggests that the late negativity does not occur when the ungrammaticality actually occurs at the end of the sentence, so the evidence for a late negativity summarized in Section 4.3 does not count as evidence for a sentence wrap-up effect per se nor for a sentence wrap-up confound.

4.5   Overlapping components, the N400 and the frontal positivity

Returning to the main twists in the development of the wrap-up dogma, once it became clear that biphasic responses were common in final position (Section 4.4.1), the argument for the sentence-final dogma changed direction. The focus switched to the N400 time window on the final word and the claim that this response is limited to the final word.

This next twist relates to a suggestion made by Osterhout, et al. (1994). As discussed above, they found that sentence-final words in sequences like *The doctor persuaded the patient was lying.* showed a clear negativity at the final verb *lying*, in comparison to variants with verbs that can take a sentential complement. In line with the conclusion of Osterhout and Holcomb (1992), they considered that this was probably an N400 response, as the ungrammatical sentence cannot be fully interpreted. However, in this experiment, the negativity showed a distinctly anterior scalp distribution. Since the final verb *lying* followed *was,* which had elicited a P600 which continued into the final word, they suggested that the overlap between these two components might explain the frontal distribution (see also Brouwer & Crocker, 2017, for a discussion of component overlap).

Because the N400 typically shows up on much the same electrodes as the P600 does, the two components could cancel each other out in this region. What Osterhout, et al. suggest is that the number of components can be reduced. Rather than positing the P600, the N400 and a distinct negative effect with a more anterior

distribution in the same time window as the N400, they suggest that the left anterior negativity (LAN) is a crypto-N400 disguised by a positivity that overlies and masks it.

Osterhout (1997) extended this argument to the combination of an N400 and P600 that are both elicited by the final word itself, when he found a biphasic frontal negativity plus P600: "Importantly, the presence of effects with opposite polarities within the same epoch could obscure the actual distribution of these effects, due to component overlap. Most notably, the anterior distribution of the 'syntactic' negativity has been taken to indicate that it is neurally distinct from the posteriorly distributed N400 elicited by semantic anomalies. However, it is possible that these anterior negativities reflect a modulation of the N400 component, but that the distribution of this effect is obscured due to its overlapping spatially and temporally with a posteriorly distributed positive wave" (p. 497). In Osterhout's argument the overlapping distributions were linked to the idea that there is a sentence-final message-level N400, as in "Most studies reporting an anterior negative-going effect have placed the anomalous word in sentence-final position." (p. 517). Note that the emphasis in the argument here has moved from the slow negativity to a sentence-final N400, a necessary move if the sentence-final late negativity simply does not occur when the ungrammaticality is at the end of the sentence, as shown in 4.4.2. Hagoort and colleagues also endorsed this hypothesis as an explanation of the frontal negativities (Hagoort and Brown, 1999; Hagoort, 2003). However, it should be noted that the component overlap argument can be and has been separated from the sentence-final position.

Two separate issues are twined together here. One is whether the (frontal) negativity is only found in final position. That claim is related to the idea that syntactic violations only elicit a message-level negativity at the end of sentences (the wrap-up dogma). The second issue at stake is whether anterior negativities (LAN) really exist or are only an artifact of the combination of an N400 (whether wrap-up effect or not) and a syntactic effect. These important issues are discussed in turn in Sections 4.5.1 and 4.5.2 below.

4.5.1 Biphasic responses are not limited to sentence-final positions

The issue most relevant to the evaluation of the existence of sentence wrap-up is thus whether syntactic negativities are limited to the final position. This is very clearly *not* the case. Molinaro, Barber and Carreiras (2011) provide a comprehensive overview of studies examining various forms of syntactic violations in sentence medial position. There is considerable variability in these results, but the overall picture is that 28 of the 38 studies reported a negativity, which was more usually frontal, but sometimes appeared to have a more centro-parietal distribution, like the N400.

Violations at the ends of sentences were not examined by Molinaro et al. (2011), because of the sentence wrap-up confound. Nevertheless, their results strongly undermine the last remaining argument for "deformation" of the response to ungrammaticalities by sentence wrap-up effects, because they show that biphasic responses are the rule rather than the exception in medial positions.

In general, it appears that if a violation shows a negativity (whether LAN or N400) at the end of a sentence, it shows one in a sentence medial position as well. It is useful to have this verified by results from the same language and preferably very comparable materials, since there is clearly some variability as to whether the negativity is frontal or central-parietal or present at all (see Molinaro et al., 2011, for discussion). These are not impossible to find. A very clear illustration is provided by the results of Gunter, Friederici and Schriefers (2000), who tested gender agreement medially, and Koelsch et al. (2005), who used the same materials, except for the removal of a phrase following the violation. Examination of Figure 1A in the Gunter et al. study (p. 561) and Figure 8A in the Koelsch et al. study (p. 1572) shows that the results are practically identical; both showing a LAN followed by a P600. The only difference is the presence of a sentence-final positivity in the Koelsch et al. study. As already discussed in Section 4.2, the sentence-final positivity does not appear to be influenced by the presence of an ungrammaticality.

Below, we have repeated the studies cited in section 4.4.1 as reporting a biphasic negative-positive response for a violation in final position for convenience. In comparison, we provide the study with the most comparable design in sentence-medial position.

- phrase structure violation
    - final: Friederici et al., 1993, Hahne and Jescheniak, 2001
    - medial: Frisch, Hahne, and Friederici, 2004
- subject verb agreement violation
    - final: Hinojosa et al., 2003
    - medial: Silva-Pereyra and Carreiras, 2007
- non-finite verb mismatch
    - final: Hahne and Friederici, 1999; Gunter et al., 1997
    - medial: Osterhout and Nicol, 1999; Ditman et al., 2007
- reflexive number or gender agreement error
    - final: Osterhout, 1997
    - medial: Osterhout, 1997
- grammatical gender violation
    - final: Koelsch et al., 2005
    - medial: Gunter et al., 2000

The medial effects are much more abundantly available in the literature than this overview would suggest, of course; it is the matching final ones that are hard to find, since they have been difficult to publish. Interestingly, Hagoort, Wassenaar and Brown (2003) are among those who report a medial anterior negativity for subject verb agreement. In their discussion, they argue for an anterior effect separate from the N400, despite the earlier endorsement of the overlap hypothesis. The evidence throughout the literature demonstrates that the initial hypothesis that syntactic negativities are limited to final position is incorrect. This is reflected in the current state of the field, in which it is widely accepted that the LAN exists and not only in sentence-final position. Table 2 gives another overview of the general pattern:

syntactic violations elicit the same pattern medially and finally, while medial

presentations are followed by a later negativity.

Table 2
Distribution of ERP effects by violation type

Phenomenon	Examples	Negativity after medial	Biphasic final	Biphasic medial
Phrase structure order violation	*…the emotional *somewhat reaction of his <u>wife</u>.*[1] *… wurde im *<u>gepflückt</u>.* ("was in the *picked")[2] *…wurde im *<u>gepflückt</u> und...* ("was in the *picked and…")[2]	√	√	√
Subject-verb agreement violation	*… The men *hopes to <u>succeed</u>.*[3] *La prueba ocultada por el fiscal *<u>aparecí</u>.*[4] ("The proof hidden by the public prosecutor *appeared [1st person]")	√	√	√
Non-finite verb mismatch	*…won't eating the food on the <u>porch</u>.*[5] *The drowning child was *<u>rescue</u>.*[6] *The drowning child was *<u>rescue</u> from the river.*[6]	√	√	√
Subject reflexive agreement violation	*The man helped *themselves/herself to the <u>food</u>.*[3] *The man helped *<u>themselves/herself</u>.*[3] *The man helped *<u>themselves/herself</u> to the food*[8]	√	√	√
Grammatical gender agreement violation	*…*il/la scialle cammina lentamente per la <u>salita</u>.*[7] ("…an[wrong phonological form]/la[wrong gender] *shawl walks slowly up the slope") *…befährt den *<u>Land</u>.*[8] ("travels through the[masc] *country[neuter]) *…befährt den *<u>Land</u> in einem alter Wartburg.*[9] ("travels through the[masc] *country[fem] in an old Wartenburg car")	√	√	√
Semantic violation (Section 4.2)	*The cats won't *bake the food on the <u>porch</u>.*[5] *The cats won't *<u>bake</u>.*[10] *The cats won't *<u>bake</u> the food on the porch.*[5]	√	N400	N400

Notes: * point of violation; Underscored: target word. The examples serve as examples; some liberties have been taken to make the relationship more obvious where it did not impact the sentence structure and translations were used where the structure is similar. For representative studies, see: [1] Hagoort et al. (1993); [2] Friederici et al. (1993) ; [3] Osterhout & Mobley (1995); [4] Hinojosa et al. (2003); [5] Osterhout and Nicol (1999); [6] Gunter et al. (1997); [7] Molinaro et al. (2008); [8] Koelsch et al. (2005); [9] Gunter et al. (2000); [10] Hagoort and Brown (2000).

4.5.2   Is the LAN a crypto-N400?

The issue of whether the LAN is a crypto-N400 is not essential to the argument about sentence wrap-up. As argued by e.g. Tanner and Van Hell (2014), and Tanner (2015), the LAN could result from overlapping components even at mid-sentence positions. The point of the overlap-argument is, as Osterhout pointed out in the quote above, that scalp distribution of ERPs is used to argue for a different neuronal source of the two effects. If this can be compromised by overlapping effects, then the same source can be accepted instead.

This argument from overlapping components requires certain assumptions.

- The negativity would have to have a broader distribution than the P600, that is, it would have to show up on frontal electrodes where there is no P600 or the P600 is too small to overwrite it.

- The amplitude would have to be similar for both effects in the time window of the N400 in the region in which they overlap so that they cancel each other out.

There seems to be quite a few reasons to doubt these assumptions.

First, as has already been noted, Molinaro et al.'s (2011) overview shows that an N400 is sometimes seen instead of a LAN before a P600. If the hypothesis is correct, one would expect the negativity always to be frontal. If it sometimes is not, there should be some account of 1) why the P600 is sometimes able to overwrite the N400 and sometimes does not, as well as 2) why when it is not overwritten, it does not generally show any evidence of a frontal extension (the part that is *not* overwritten when it is *smaller*). Second, the model suggests that the P600 ought to overwrite *semantic* N400s as well, but the P600 and the lexical semantic N400 effects are additive according to several studies (Gunter, et al., 1997; Osterhout & Nicol, 1999); an account for this raises the same issues as above with regard to Molinaro et al.'s overview. Probably most convincing is recent evidence (Pakulak and Neville, 2010) that the distribution of the anterior negativity does not become more N400-like when the size of the P600 decreases. Under the assumptions sketched above, the crypto-N400 should slowly emerge as the P600 decreases.

4.5.3   Why maintain the sentence wrap-up dogma?

The sentence wrap-up argument began as a potential explanation for an undesired "syntactic negativity". This goal has not been achieved, and given the recent discussion of the real distinction between the N400 and P600 (Bornkessel-Schlesewsky & Schlesewsky, 2008; Brouwer & Crocker, 2017; Brouwer, Fitz & Hoeks, 2012; Kuperberg, 2007), it would fail in any case. The consensus is that the syntax/semantics distinction is out of the running as an explanation for the distinction between the N400 and P600. The next twist was to regard the sentence-final confound as an overlapping negativity and positivity. The rationale was to reduce the number of components related to language processing, with the caveat that the N400 to syntactic violations should be limited to final position. The acceptance of a syntactic negativity, the LAN, seems to be widespread in the syntactic ERP world, even in sentence medial position (but see, e.g. Tanner & Van Hell, 2014). A substantial argument for the sentence wrap-up confound is gone with that acceptance, somehow the sentence wrap-up confound remains uncontested.

4.6   (Lack of) evidence for an interaction of language ERPs with position

In the prior sections evidence has been presented that the responses for violations are similar at medial and final positions, arguing against the sentence-final confound. However, it is "common knowledge" that violation effects are different at the end of a sentence. That common knowledge is based on four experiments[7] (two in Osterhout, 1997; two in Hagoort and Brown, 1999, and Hagoort, 2003, with essentially the same design). There are some clear limitations of the conclusions

---

[7] There are other citations in the literature, but many of those simply refer back to formulations of the hypothesis or to these experiments. This is one of the mechanisms by which a dogma comes into existence. Since the claim is copied and recopied it appears to be very well-documented, unless the references are actually checked.

which can be drawn from these studies, despite a very high citation index as proof of position effects.

Osterhout (1997) set out to determine whether sentence-final position has a clear effect on the processing of syntactic violations. In the first study, violations of gender and number agreement on reflexive pronouns were examined (e.g. *The salesman congratulated himself/*herself/*themselves*…). Whether the target pronoun was sentence-final or not, an anterior negativity and a subsequent positivity were found. Osterhout reports a main effect of violation with no interaction with position in the main analysis. That is to say, there was no evidence of a different response at the sentence-final position. Examining the electrode where the effect was strongest showed an interaction, suggesting that the effect was *bigger* at the end of the sentence, but there was no evidence that the scalp distribution was altered or that the negativity does not occur in medial position. Osterhout's conclusion from this experiment was that morphosyntactic violations on closed class words do indeed show an anterior negativity, although the scalp distribution might be due to overlap (Section 4.5). In the subsequent discussion within this article, Osterhout is careful to limit further concerns about individual differences in the presence of P600 and N400 effects to open class words.

The second experiment reported in this article followed up the results from Osterhout and colleagues' earlier studies with temporarily ambiguous sentences. The sentences were presented with the disambiguating word in final position or with a following phrase (e.g., *The boat sailed down the river [and] sank [during the storm]*). Without *and*, the sentence is ungrammatical unless it is reanalyzed to the non-preferred reduced relative clause structure, meaning *The boat which was floated down the river sank*, so garden path effects are expected at *sank,* similar to those reported by Osterhout, Holcomb and Swinney (1994). This is visible as a P600 time-locked to *sank*. As predicted by the sentence-final confound hypothesis, *sank* also shows evidence of an N400 effect when it is in final position.

What is not consistent with the original hypothesis is that the response is biphasic (as seen in other studies: 4.4.1) and what is not consistent with the extended variant is that virtually the same waveform is visible when *sank* is not sentence-final (as seen in other studies: 4.5.1). Position did not interact with sentence type in either the N400/LAN or the P600 time window. The ERP effects for garden path *sank* and a semantically anomalous control (*barked*) were virtually identical. The results of this experiment thus provide no evidence that the N400 can be overwritten by the P600 to produce a frontal negativity. Nor do they provide evidence that the N400 is limited to sentence-final position. There was a small difference in the lateralization of the scalp distribution, but no bigger than one might expect, given that the experiment was carried out as a between-subject design; different subjects saw the sentence-final version than saw the sentence medial versions. Because there is some variability between individual subjects, it is important to use a within-subject design in order to establish that a difference in scalp distribution is meaningful.

The claim that the study nevertheless ***does*** provide evidence for a special status for the sentence-final position is based primarily on a post-hoc analysis. Osterhout examined the waveforms of individuals and concludes that the biphasic pattern results from different strategies across individuals, some using a semantic analysis to resolve the problem thus producing an N400, and some showing the typical response to a garden path (P600). More subjects used the semantic strategy in sentence-final position (12 out of 15) than in sentence medial position (2 out of 15). This conclusion is weak, since the differentiation did not show up in the statistics for the amplitude measures, and as noted above, the results involved different subjects, who may simply have differed from each other. The individual patterns obviously could not be put to a statistical test after creating groups in this way. The effect has never (to the best of our knowledge) been replicated. Although this article is one of those that are frequently cited to support the notion of sentence wrap-up as a serious confounding factor in ERP research, it does not strongly support the

hypothesis. The effects are certainly not strong enough to invalidate the evidence that medial and final positions are in general quite similar with regard to syntactic violations presented in 4.5.

The other set of experiments which is cited as evidence for negativity in the final position was reported by Hagoort and Brown (1999) and Hagoort (2003). In these experiments they presented Dutch sentences containing nouns that agreed or failed to agree with the preceding determiner in grammatical gender. The noun phrases included an intervening adjective and either began the sentence (*De/*het kapotte paraplu*...: lit. *The[common gender]/the[neuter gender] broken umbrella[common gender]*) or completed it (*Cindy sliep slecht vanwege de/*het griezelige droom.;* lit: *Cindy slept badly due to the[common gender]/the[neuter gender] gruesome dream [common gender]*). Only when the critical word was in sentence-final position an N400 was observed. However, it is well-established that words in general become more predictable in the course of a sentence (Van Petten and Kutas, 1990), and in some cases become highly predictable, resulting in a smaller N400. In these materials, *dream* at the end of the sentence is much more predictable than *umbrella* is at the start —but only when preceded by the appropriate article. With the incorrect article, *dream* is semantically congruent, but if a gender marked determiner modifies the predictability of the upcoming word, *dream* is less congruent than it is with the correct article. The extent to which the incongruent article can be expected to have an effect depends on how quickly that sort of constraint is expected to affect the activation of the target noun. It has been demonstrated that listeners use grammatical gender to predict upcoming nouns (e.g. Dahan, Swingley, Tanenhaus & Magnusen, 2000). Gender may not be as effective at deactivating words as it is in increasing their activation, but that has not been clearly established.

This confound with predictability of the noun is a very plausible source for a difference between the effect seen in final position and the N400-less effect in medial position. When Osterhout and Holcomb (1992) found an unexpected N400,

48

they reran the experiment with the same materials in both conditions to make sure that cloze probability differences did not explain their results, but a replication containing this check was never published for these materials. Again in light of the results discussed above, the results would have to be considerably stronger to be convincing that sentence-final position causes a qualitative change in the processing of a syntactic violation.

To conclude, even though it is "common knowledge" that responses to syntactic violations are very different at medial and sentence-final position, the actual evidence for clear interactions between position and syntactic effects is minimal. To have built a dogma leading to such a strongly enforced ban on designs with sentence-final targets on this basis is not reasonable.

5   Summary and conclusion

5.1 No real evidence for sentence wrap-up

Below is a summary of the various stones that have formed the foundation of the sentence wrap-up dogma discussed in the preceding sections. Following each is a summary of the evidence, and an explanation of why this evidence is not strong enough to support the claim that sentence-final ERP effects reflect sentence wrap-up. This summary demonstrates that the sentence wrap-up dogma is endorsed only by convention, and represents an unnecessary and severe limitation on ERP research.

● Sentence-final lengthening/regression and other behavioral evidence provide clear evidence for a process of sentence wrap-up.

As Section 3 demonstrates at some length, although there are certainly sentence-final effects, the evidence does not suggest that it reflects processes of sentence wrap-up that can also be expected to show up as an ERP effect.

- There is a sentence-final positivity which reflects sentence wrap-up.

There is indeed a sentence-final positivity (4.1), but it does not interact with semantic factors as measured by the N400 (section 4.2), nor is it modified by syntactic factors as represented by the P600 (section 4.4) when either of these is measured at the end of a sentence. If it is affected by factors that are relevant to any existing concept of sentence wrap-up, this has not been demonstrated to date.

- There is a sentence-final slow negativity following ungrammaticalities earlier in the sentence which may also be elicited by ungrammaticalities at the end of the sentence, causing a negative response *instead* of a P600 to ungrammaticalities at that position.

There is indeed a slow negativity following many if not all sentence-medial ungrammaticalities (section 4.3). It certainly does not block a P600 response, which is generally present for ungrammaticalities measured at the end of the sentence (section 4.4.1). Stronger yet, the slow negativity is not elicited by a violation presented at the end of a sentence (section 4.4.2), so it cannot cause a confound for those ungrammaticalities. In fact it is not even obvious that it is limited to the end of the sentence, since in studies which present multi-word ERPs, it appears earlier in the sentence as well. This observation undermines its status as a sign of sentence wrap-up (section 4.4.3).

This brings up a problem at a more theoretical level. An effect elicited only by ill-formed sentences does not constitute evidence for the theoretical constructs of sentence wrap-up that have been discussed in psycholinguistics beyond the bounds of ERP research. Wrap-up is considered to be a process that occurs as a standard process in sentence comprehension, not just when dealing with ungrammaticality. That is, evidence for sentence-final effects limited to ungrammatical sentences was never valid as an argument against measuring at the end of normal *grammatical* sentences.

- When ungrammaticalities are measured at the end of the sentence, there is an N400 effect due to message level processes occurring specifically at the end of the sentence in response to a syntactic violation. However, due to overlap with the P600, this effect has a frontal distribution (LAN). This negativity cannot be generalized to other positions in the sentence.

For most of the last fifteen years or so, this has been the core assumption underlying the ERP wrap-up dogma. However, it is abundantly evident that frontal negativities are not limited to the end of sentences (section 4.5), therefore invalidating any argument of this nature. This conclusion holds regardless of whether the frontal distribution of the negativity in the biphasic response can or should be explained as being due to overlapping N400 and P600 effects, although the thorough review of sentence medial negativities provided by Molinaro et al. (2011) shows that the issue of the scalp distribution of the negativity is much more complex than can be explained by the overlap hypothesis. The review of the assumptions of the overlapping argument in 4.5.2 also suggests that the claim that the LAN is a masked N400 is incorrect.

- Interactions have been reported in the literature between grammaticality and position, which strongly support different processes at the end of the sentence.

There are two sources which report interactions (section 4.6). In one case the "syntactic" negativity was possibly slightly stronger at the end than in the sentence medial position, although the statistical evidence for this claim is not strong, but negativities occurred in both positions (Osterhout, 1997; Exp 1 and Exp 2). In the other set of experiments, there is a clear confound involving differences in the cloze probability of the target nouns with no negativity and those containing a negativity (Hagoort and Brown, 1999; Hagoort, 2003). For this reason they cannot be accepted as strong evidence for an interaction. When examined in the context of the overwhelming evidence that negative-positive biphasic responses are regularly

reported in medial position as well as in final position, these results do not provide a counter-argument.

To sum up, despite the general intuitive appeal of the idea of sentence wrap-up, there is very little evidence that it exists at all, much less that it should determine how ERP research is carried out. For the LAN and P600, there is essentially no evidence that the end of the sentence incurs additional sentence wrap-up which will influence the results of any measurement made at that point. At most slight differences have been demonstrated.

5.2 The way forward

Not only is the "sentence-final wrap-up" dogma unsupported by evidence, the embargo on sentence-final positions has likely blocked research on important effects. As one can glean from this paper, hardly any research has been carried out as to what actually happens at the end of a sentence, clause, or prosodic boundary, what the differences are between the boundary types, and which processes are and are not affected by the point at which they occur in a sentence. As demonstrated by Kuperberg et al. (2011), there may indeed be some processes that are affected by sentence position, including discourse processes as reflected in the N400. These effects should be a point of interest and research, rather than taken as further evidence for an embargo, because it will actually tell us something about how comprehension proceeds. Along the same lines, the sentence-final positivity as well as the sentence-final effects in behavioral research deserve further investigation.

Current sentence processing research has been focused on early effects of on-line incremental processes that are performed as each word comes in, or even in anticipation of the input. This obviously does not give a complete picture of human sentence processing. It is therefore time that attention be devoted also to slower, later processes, and to what happens at the end of the clause, sentence, or other boundary.

Further research should also bear task effects in mind. Task demands and response preparation can heavily affect how readers and listeners are processing sentences, at sentence medial as well as final positions (Sabourin & Stowe, 2004; Swets et al., 2008). Some tasks create position effects, for instance, when the sentence final words contains information that is critical for a following comprehension question or judgment task. When designing experiments in which task effects are not of interest, it therefore is advisable to avoid tasks involving a decision that is confounded with position.

An important implication of this review is that if ERP measurements have to occur at the end of a clause or sentence, for example due to the structure of a language or a language's most common word order, they should be carried out there. If the design involves processes other than the LAN and P600 which were specifically reviewed here, replicating at other points in the sentence would be useful; if a component exhibits position-specific effects, this is an important feature of that component.

A final point is that although this article is directed at a particular dogma from a particular literature, this review shows that it is important to reassess those ideas that are considered "general knowledge" in the literature from time to time. The basic assumptions may rest on a foundation of non-existent or misinterpreted stones which in hindsight do not provide as sturdy a support as they seemed to when first proposed.

References

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*(3), 247–264.

Bertram, R., Hyönä, J., & Laine, M. (2000). The role of context in morphological processing: Evidence from Finnish. *Language and Cognitive Processes*, *15*(4–5), 367–388.

Bever, T. G., Fodor, J. A., & Garrett, M. (1966). *Speech perception: Some experimental results for sentences*. Paper presented at the International Congress of Psychology, Moscow, Russia.

Bever, T. G., Lackner, J. R., & Kirk, R. (1969). The underlying structures of sentences: All the primary limits of immediate speech processing. *Perception & Psychophysics*, 5, 225–234.

Bohan, J., Leuthold, H., Hijikata, Y., & Sanford, A. J. (2012). The processing of good-fit semantic anomalies: An ERP investigation. *Neuropsychologia, 50*, 3174–3184.

Boland, J. (2004). Linking eye movements to sentence comprehension in reading and listening. In M. Carreiras & C. Clifton Jr. (Eds.), *The on-line study of sentence comprehension* (pp. 51–76). New York: Psychology Press.

Bornkessel, I., McElree, B., Schlesewsky, M., & Friederici, A. D. (2004). Multi-dimensional contributions to garden path strength: Dissociating phrase structure from case marking. *Journal of Memory and Language, 51*(4), 495–522.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on "semantic P600" effects in language comprehension. *Brain Research Reviews*, *59*(1), 55–73.

Braze, D., Shankweiler, D., Ni, W., & Palumbo, L. C. (2002). Readers' eye movements distinguish anomalies of form and content. *Journal of Psycholinguistic Research*, *31*(1), 25–44.

Brouwer, H., & Crocker, M. W. (2017). On the proper treatment of the N400 and P600 in language comprehension. *Frontiers in Psychology, 8*(1327).

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about semantic illusions: rethinking the functional role of the P600 in language comprehension. *Brain Research*, *1446*, 127–143.

Camblin, C. C., Gordon, P. C., & Swaab, T. Y. (2007). The interplay of discourse congruence and lexical association during sentence processing: Evidence from ERPs and eye tracking. *Journal of Memory and Language*, *56*(1), 103–128.

Carpenter, P. A., & Daneman, M. (1981). Lexical retrieval and error recovery in reading: A model based on eye fixations. *Journal of Verbal Learning and Verbal Behavior*, *20*(2), 137–160.

Dahan, D., Swingley, D., Tanenhaus, M. K., & Magnuson, J. S. (2000). Linguistic gender and spoken-word recognition in French. *Journal of Memory and Language, 42*(4), 465–480.

Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9(4), 561–584.

De Vincenzi, M., Job, R., Di Matteo, R., Angrilli, A., Penolazzi, B., Ciccarelli, L., & Vespignani, F. (2003). Differences in the perception and time course of syntactic and semantic violations. *Brain and Language*, *85*(2), 280–296.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*(8), 1117–1121.

Ditman, T., Holcomb, P. J., & Kuperberg, G. R. (2007). An investigation of concurrent ERP and self-paced reading methodologies. *Psychophysiology*, *44*(6), 927–935.

Ehrlich, K., & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior*, *22*(1), 75–87.

Ferreira, F., & Patson, N. D. (2007). The 'Good Enough' approach to language comprehension. *Language and Linguistics Compass, 1*(1-2), 71–83.

Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2001). Syntactic working memory and the establishment of filler-gap dependencies: Insights from ERPs and fMRI. *Journal of Psycholinguistic Research, 30*(3), 321–338.

Fodor, J. A., & Bever, T. G. (1965). The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior*, *4*(5), 414–420.

Fodor, J. A., Garrett, M., & Bever, T. G. (1968). Some syntactic determinants of sentential complexity, II: Verb structure. *Perception & Psychophysics, 3*(6), 453–461.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *14*(2), 178–210.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6*(2), 78–84.

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, *1*(3), 183–192.

Friedman, D., Simson, R., Ritter, W., & Rapin, I. (1975). The late positive component (P300) and information processing in sentences. *Electroencephalography and Clinical Neurophysiology, 38*(3), 255–262.

Frisch, S., Hahne, A., & Friederici, A. D. (2004). Word category and verb–argument structure information in the dynamics of parsing. *Cognition, 91*(3), 191–219.

Garnsey, S. M., Tanenhaus, M. K., & Chapman, R. M. (1989). Preferred verb argument structure in sentence comprehension – an ERP study. *Bulletin of the Psychonomic Society, 27*(6), 522–522.

Garrett, M., Bever, T., & Fodor, J. (1966). The active use of grammar in speech perception. *Perception & Psychophysics*, *1*(1), 30–32.

Garrod, S., & Terras, M. (2000). The contribution of lexical and situational knowledge to resolving discourse roles: Bonding and resolution. *Journal of Memory and Language*, *42*(4), 526–544.

Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience*, *12*(4), 556–568.

Gunter, T. C., Stowe, L. A., & Mulder, G. (1997). When syntax meets semantics. *Psychophysiology*, *34*(6), 660–676.

Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience, 15*(6), 883–899.

Hagoort, P., & Brown, C. M. (1994). Brain responses to lexical ambiguity resolution and parsing. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing*. (pp. 45–81). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hagoort, P., & Brown, C. M. (1999). Gender electrified: ERP evidence on the syntactic nature of gender processing. *Journal of Psycholinguistic Research*, *28*(6), 715–728.

Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech: semantic ERP effects. *Neuropsychologia, 38*(11), 1518-1530.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, *8*(4), 439–483.

Hagoort, P., Wassenaar, M., & Brown, C. M. (2003). Syntax-related ERP-effects in Dutch. *Cognitive Brain Research*, *16*(1), 38–50.

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience*, *11*(2), 194–205.

Hahne, A., & Jescheniak, J. D. (2001). What's left if the Jabberwock gets the
semantics? An ERP investigation into semantic and syntactic processes during
auditory sentence comprehension. *Cognitive Brain Research*, *11*(2), 199–212.

Hill, R., & Murray; W. (2000). Commas and spaces: Effects of punctuation on
eye movements and sentence processing. A. Kennedy, D. Heller, J. Pynte (Eds.),
Reading as a perceptual process (pp. 565–589), Amsterdam: Elsevier.

Hinojosa, J.A., Martín-Loeches, M., Casado, P., Munoz, F., and Rubia, F.J. (2003).
Similarities and differences between phrase structure and morphosyntactic
violations in Spanish: An event-related potentials study. *Language and Cognitive
Processes*, *18(2)*, 113–142.

Hirotani, M., Frazier, L., & Rayner, K. (2006). Punctuation and intonation effects on
clause and sentence wrap-up: Evidence from eye movements. *Journal of Memory
and Language*, *54*(3), 425–443.

Ito, K., & Speer, S. (2008). Anticipatory effects of intonation: Eye movements during
instructed visual search. *Journal of Memory & Language*, *58*, 541–573.

Jarvella, R. J. (1971). Syntactic processing of connected speech. *Journal of Verbal
Learning and Verbal Behavior*, *10*(4), 409–416.

Just, M. A., & Carpenter, P. A. (1978). Inference processes during reading: Reflections
from eye fixations. *Eye movements and the higher psychological functions.
Hillsdale, NJ: Erlbaum*, 157–174.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to
comprehension. *Psychological Review*, *87*, 329–354.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in
reading comprehension. *Journal of Experimental Psychology: General*, *111*(2),
228–238.

Kaan, E. (2002). Investigating the effects of distance and number interference in
processing subject-verb dependencies: An ERP study. *Journal of Psycholinguistic
Research*, *31*(2), 165–193.

Kamide, Y., Scheepers, C., & Altmann, G. T. M. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research*, 32(1), 37–55.

King, J., & Kutas, M. (1995). Who did what and when? Using word-and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, *7*(3), 376–395.

Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory and Language*, *29*(2), 133–159.

Koelsch, S., Gunter, T. C., Wittfoth, M., & Sammler, D. (2005). Interaction between syntax processing in language and in music: An ERP study. *Journal of Cognitive Neuroscience*, *17*(10), 1565–1577.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*, 23–49.

Kuperberg, G. R., Kreher, D. A., Goff, D., McGuire, P. K., & David, A. S. (2006). Building up linguistic context in schizophrenia: Evidence from self-paced reading. *Neuropsychology*, *20*(4), 442–452.

Kuperberg, G. R., Paczynski, M., & Ditman, T. (2011). Establishing causal coherence across sentences: An ERP study. *Journal of Cognitive Neuroscience*, *23*(5), 1230–1246.

Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, *17*(1), 117–129.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205

Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition*, *11*(5), 539–550.

Kutas, M., Lindamood, T. E., & Hillyard, S. A. (1984). Word expectancy and event-related brain potentials during sentence processing. In S. Kornblum & J. Requin (Eds.), *Preparatory States and Processes* (pp. 217–237), Hillsdale: Erlbaum.

McDonald, J. L., & MacWhinney, B. (1995). The time course of anaphor resolution: Effects of implicit verb causality and gender. *Journal of Memory and Language*, *34*(4), 543–566.

Miller, G. A., & Chomsky, N. (1963). Finitary models of language users. In R. D. Luce, R. R. Bush & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 2, pp. 419–491). New York: Wiley.

Mitchell, D. C., & Green, D. W. (1978). The effects of context and content on immediate processing in reading. *Quarterly Journal of Experimental Psychology, 30,* 609–636.

Molinaro, N., Barber, H. A., & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex*, *47*(8), 908–930.

Molinaro, N., Vespignani, F., & Job, R. (2008). A deeper reanalysis of a superficial feature: An ERP study on agreement violations. *Brain Research*, *1228*, 161–176.

Münte, T. F., Heinze, H. J., & Mangun, G. R. (1993). Dissociation of brain activity related to syntactic and semantic aspects of language. *Journal of Cognitive Neuroscience*, *5*(3), 335–344.

Nakamura, C., Arai, M., & Mazuka, R. (2012). Immediate use of prosody and context in predicting a syntactic structure. *Cognition, 125*(2), 317–323.

Neville, H., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, *3*(2), 151–165.

Osterhout, L. (1997). On the brain response to syntactic anomalies: Manipulations of word position and word class reveal individual differences. *Brain and Language,* 59(3), 494–522.

Osterhout, L., Allen, M. D., McLaughlin, J., & Inoue, K. (2002). Brain potentials elicited by prose-embedded linguistic anomalies. *Memory & Cognition*, *30*(8), 1304–1312.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*(6), 785–806.

Osterhout, L., & Holcomb, P. J. (1993). Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes*, *8*(4), 413–437.

Osterhout, L., & Holcomb, P. J. (1995). Event related potentials and language comprehension. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind: Event-related brain potentials and cognition* (pp. 171–215). New York, NY: Oxford University Press.

Osterhout, L., Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 786.

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, *34*(6), 739–773.

Osterhout, L., & Nicol, J. (1999). On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Language and Cognitive Processes*, *14*(3), 283–317.

Pakulak, E., & Neville, H. J. (2010). Proficiency differences in syntactic processing of monolingual native speakers indexed by event-related potentials. *Journal of Cognitive Neuroscience*, *22*(12), 2728–2744.

Perfetti, C., & Adlof, S. M. (2012). Reading comprehension: A conceptual framework from word meaning to text meaning. *Measuring up: Advances in how we assess reading ability*, 3–20.

Pexman, P. M., Ferretti, T. R., & Katz, A. N. (2000). Discourse factors that influence online reading of metaphor and irony. *Discourse Processes*, *29*(3), 201–222.

Phillips, C., Kazanina, N., & Abada, S. H. (2005). ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research*, *22*(3), 407–428.

Rayner, K. (1975). Parafoveal identification during a fixation in reading. *Acta Psychologica, 39*(4), 271–281.

Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *Quarterly Journal of Experimental Psychology*, 53A, 1061–1080.

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*, 4(special issue), 21–49.

Rösler, F., Pütz, P., Friederici, A., & Hahne, A. (1993). Event-related brain potentials while encountering semantic and syntactic constraint violations. *Journal of Cognitive Neuroscience*, *5*(3), 345–362.

Ruchkin, D. S., Johnson Jr, R., Canoune, H., & Ritter, W. (1990). Short-term memory storage and retention: An event-related brain potential study. *Electroencephalography and Clinical Neurophysiology*, *76*(5), 419–439.

Sabourin, L., & Stowe, L. (2004). Memory effects in syntactic ERP tasks. *Brain and Cognition, 55*(2), 392–395.

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology*, *14*(4), 489–537.

Silva-Pereyra, J. F., & Carreiras, M. (2007). An ERP study of agreement features in Spanish. *Brain Research, 1185*, 201–211.

Steinhauer, K. (2003). Electrophysiological correlates of prosody and punctuation. *Brain and Language*, *86*(1), 142–164.

Steinhauer, K., & Friederici, A. D. (2001). Prosodic boundaries, comma rules, and brain responses: The closure positive shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. *Journal of Psycholinguistic Research*, *30*(3), 267–295.

Stowe, L. A. (1986). Parsing WH-constructions: Evidence for on-line gap location. *Language and Cognitive Processes*, 1(3), 227–245.

Stowe, L. A. (1991). Thematic structure and verb preferences. In Y. No & M. Libucha (Eds.), *Escol '90. Proceedings of the Eastern States Conference on Linguistics* (pp. 272-285). The Ohio State University, Columbus, OH.

Swets, B., Desmet, T., Clifton, C., & Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory & Cognition, 36*(1), 201–216.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re) consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, *18*(6), 645–659.

Tanner, D. (2015). On the left anterior negativity (LAN) in electrophysiological studies of morphosyntactic agreement: A Commentary on "Grammatical agreement processing in reading: ERP findings and future directions" by Molinaro et al., 2014. *Cortex, 66*(Supplement C), 149–155.

Tanner, D., & Van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia, 56*(0), 289–301.

Ueno, M., & Kluender, R. (2003). Event-related brain indices of Japanese scrambling. *Brain and Language*, *86*(2), 243–271.

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 443–467.

Van Berkum, J. J. A., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience*, *11*(6), 657–671.

Van Berkum, J. J. A., Koornneef, A. W., Otten, M., & Nieuwland, M. S. (2007). Establishing reference in language comprehension: An electrophysiological perspective. *Brain Research*, *1146*, 158–171.

Van Berkum, J. J. A., Zwitserlood, P., Hagoort, P., & Brown, C. M. (2003). When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research*, *17*(3), 701–718.

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition*, *18*(4), 380–393.

Van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open-and closed-class words. *Memory & Cognition*, *19*(1), 95–112.

Warren, T., White, S. J., & Reichle, E. D. (2009). Investigating the causes of wrap-up effects: Evidence from eye movements and E-Z Reader. *Cognition, 111*(1), 132–137.

Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience, 16*(7), 1272–1288.

Figure Caption

Figure 1

Reproduced with permission from Kutas and Hillyard (1980). In this study it was demonstrated that a word which did not fit the context semantically elicited a negativity relative to a normal sentence. The negativity differed from the surprise effect elicited by a word in larger font, which elicited a positivity. In terms of the current article, it demonstrates the general form of the sentence-final positivity, which is much larger than that for the preceding words, indicated by the double-headed arrows.

Figure 1