The Quarterly Journal of Experimental Psychology

ISSN: 1747-0218 (Print) 1747-0226 (Online) Journal homepage: http://www.tandfonline.com/loi/pqje20

Good-enough linguistic representations and online cognitive equilibrium in language processing

Hossein Karimi & Fernanda Ferreira

To cite this article: Hossein Karimi & Fernanda Ferreira (2016) Good-enough linguistic representations and online cognitive equilibrium in language processing, The Quarterly Journal of Experimental Psychology, 69:5, 1013-1040, DOI: 10.1080/17470218.2015.1053951

To link to this article: http://dx.doi.org/10.1080/17470218.2015.1053951

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2016
Vol. 69, No. 5, 1013–1040, http://dx.doi.org/10.1080/17470218.2015.1053951

Routledge
Taylor & Francis Group

Good-enough linguistic representations and online cognitive equilibrium in language processing

Hossein Karimi and Fernanda Ferreira

Department of Psychology, University of California at Davis, Davis, CA, USA

(Received 3 September 2014; accepted 8 May 2015; first published online 23 June 2015)

We review previous research showing that representations formed during language processing are sometimes just "good enough" for the task at hand and propose the "online cognitive equilibrium" hypothesis as the driving force behind the formation of good-enough representations in language processing. Based on this view, we assume that the language comprehension system by default prefers to achieve as early as possible and remain as long as possible in a state of cognitive equilibrium where linguistic representations are successfully incorporated with existing knowledge structures (i.e., schemata) so that a meaningful and coherent overall representation is formed, and uncertainty is resolved or at least minimized. We also argue that the online equilibrium hypothesis is consistent with current theories of language processing, which maintain that linguistic representations are formed through a complex interplay between simple heuristics and deep syntactic algorithms and also theories that hold that linguistic representations are often incomplete and lacking in detail. We also propose a model of language processing that makes use of both heuristic and algorithmic processing, is sensitive to online cognitive equilibrium, and, we argue, is capable of explaining the formation of underspecified representations. We review previous findings providing evidence for underspecification in relation to this hypothesis and the associated language processing model and argue that most of these findings are compatible with them.

Keywords: Language processing; Heuristics; Underspecification; Online cognitive equilibrium.

When reading or hearing utterances in daily life, the language comprehension system should build not only intrasentential representations connecting the constituents within each utterance, but also intersentential (i.e., discourse) representations connecting the different utterances so that a seamless and coherent discourse representation is constructed. Based on classic theories of language processing, representations formed during language processing are accurate, precise, and detailed. However, recent evidence suggests that both intrasentential and intersentential representations can be sketchy and imprecise (e.g., Christianson, Hollingworth, Halliwell, & Ferreira, 2001; Christianson, Williams, Zacks, & Ferreira, 2006; F. Ferreira, 2003; F. Ferreira, Christianson, & Hollingworth, 2001; Greene, McKoon, & Ratcliff, 1992; Klin, Guzman, Weingartner, & Ralano, 2006; Levine, Guzman, & Klin, 2000; Stewart, Holler, & Kidd, 2007; Swets, Desmet, Clifton, & Ferreira, 2008).

For example, given the sentence *the dog was bitten by the man*, people often fail to compute the correct event representation: one in which the man (and not the dog) does the biting (F. Ferreira, 2003). Similarly, evidence from reference processing shows that the correct referents for

Correspondence should be addressed to Hossein Karimi, Department of Psychology, University of California at Davis, 135 Young Hall, One Shields Avenue, Davis, CA 95616, USA. E-mail: karimi@ucdavis.edu

referring expressions are not always found, resulting in shallow and imprecise intersentential representations (e.g., Greene et al., 1992; Klin et al., 2006; Levine et al., 2000; Stewart et al., 2007). To capture the sometimes superficial nature of the linguistic representations, Ferreira, Christianson, and colleagues (Christianson et al., 2001; F. Ferreira, 2003; F. Ferreira, Ferraro, & Bailey, 2002; F. Ferreira & Patson, 2007, also see Sanford & Sturt, 2002) proposed the theory of good-enough language processing (henceforth GE). At its core, the idea behind this theory is that linguistic representations built in the course of language processing are only good enough to tackle the task at hand and become elaborated only if mandated by the task at hand.

In this article, we first review evidence showing that intrasentential linguistic representations are indeed sketchy and incomplete, and we discuss them in relation to the GE approach. Then, we will explore how the GE approach can be applied to intersentential processing in general and reference processing in particular. Finally, we propose our own hypothesis about underlying mechanisms for shallow linguistic representations in relation to both intra- and intersentential reference processing. This hypothesis is consistent with and provides a possible set of mechanisms for the GE approach to language processing. We also propose a model of language processing that outlines a concrete and detailed architecture for the formation of linguistic representations, including those that are underspecified.

Good-enough intrasentential processing

Traditional theories of language processing assume that the representations created during the processing of linguistic information are complete and veridical. Under this view, the language processing system follows strict and clear syntactic algorithms to compute precise representations for the given linguistic input; words are accessed from the lexicon and then combined according to rules of syntax to arrive at the global sentence meaning. However, Ferreira and colleagues (F. Ferreira et al., 2002; F. Ferreira & Patson, 2007; Swets et al., 2008) have argued that algorithmic procedures for sentence processing are not only too costly but sometimes outright unnecessary (see also Boudewyn, Long, & Swaab, 2013; Kuperberg, 2007). In other words, the representations only need to be "good enough" to tackle the task at hand, and since oftentimes the task the listener needs to perform based on the linguistic input is quite minimal (nodding, confirming, executing simple motor actions, etc.), employing simple heuristic procedures is sometimes enough to fulfil the task of sustaining communication.

Heuristic processing is based on the application of simple rules that can output a quick overall representation of the information currently under processing and thus offer an economic advantage in terms of cognitive effort. As such, heuristic processing is consistent with the fundamental least effort principle, according to which human behaviour is driven by the general underlying tendency to minimize the average rate of work over time (Jaeger, 2010; Piantadosi, Tily, & Gibson, 2012; Zipft, 1935, 1949). Heuristic processing saves effortful processing by proceeding through "fast and frugal" heuristics rather than slow-going and computationally costly algorithmic processing (F. Ferreira et al., 2002).

A direct consequence of heuristic processing, however, is that the constructed representations may not be accurate reflections of the associated linguistic input. As counterintuitive as such a language processing architecture may sound, recent psycholinguistic evidence provides ample evidence that language processing sometimes leads to errors such that the correct meaning of some given linguistic input is not successfully computed. For example, it has been shown that when asked *How many of each type of animal did Moses take on the ark?* people often respond *two*, failing to realize that, according to biblical legend, it was Noah who took animals on the ark rather than Moses (Erickson & Matteson, 1981). Similarly, people also often fail to detect the anomaly in the question *Where should the authorities bury the survivors?* (i.e., they overlook the fact that *survivors* are not to be buried, Barton & Sanford, 1993).

Further evidence for underspecified representations comes from the processing of garden-path sentences such as *While Mary bathed the baby played in the crib*. As might be clear, the sentence is difficult to process because *the baby* is initially taken to be the object of the verb *bathe*. However, the remainder of the sentence rules out this analysis and forces a subject interpretation of *the baby* (i.e., the baby is doing the playing). Christianson et al. (2001) as well as F. Ferreira et al. (2001) provided evidence that the correct interpretation of this type of sentences might not always be computed. Specifically, they asked comprehension questions such as *Did Mary bathe the baby?* and found that people have the belief that Mary indeed bathed the baby. Later, Patson, Swensen, Moon, and Ferreira (2006) ruled out the possibility that the form of the question itself might have reinstated the initial wrong parse; instead of asking comprehension questions, they simply asked participants to recall the sentences. Their results showed that people recalled an event in which *Mary* bathes *the baby*, confirming the results of Christianson et al. (2001) and F. Ferreira et al. (2001).

In addition to garden-path sentences, people also seem to have difficulty computing the correct meaning of sentences that are both semantically and syntactically complex. For example, F. Ferreira (2003) manipulated the voice as well as the plausibility of experimental stimuli, producing active-plausible (e.g., "the dog bit the man"), active-implausible (e.g., "the man bit the dog"), passive-plausible (e.g., "the man was bitten by the dog"), and passive-implausible (e.g., "the dog was bitten by the man") sentences. The participants' task was to determine the thematic roles of the event participants—that is, who was the agent (doer) and who was the patient (doee). The results showed that passive-implausible sentences were the most difficult to process and led to the largest number of erroneous interpretations. Importantly, this effect was not attributable to the low frequency of passives because the meaning of rare cleft structures such as *It was the man who bit the dog* was virtually always identified correctly. This finding suggests that what is critical is not the frequency of the overall form but rather the extent to which thematic roles are assigned canonically—agent before patient. Based on these results, F. Ferreira (2003) argued that sentence processing proceeds through a heuristic process in which a noun–verb–noun syntactic template is quickly mapped into an agent–verb–patient thematic structure.

Obviously, this is not to say that algorithmic processing has no role in language processing, but that evidence seems to point to the conclusion that heuristic processing may occur in parallel with or even precede algorithmic processing. Recent evidence in favour of the priority of heuristic over algorithmic processing comes from the processing of quantifier scope ambiguity. Sentences displaying quantifier scope ambiguity such as *Every kid climbed a tree* have two main possible interpretations: Either there were several trees (one for each child) that the children climbed, or there was only one tree that the children all climbed. The former interpretation has been argued to be the default interpretation (e.g., Kurtzman & MacDonald, 1993). In order to investigate heuristic and algorithmic processing and their relative time courses, Dwivedi (2013) manipulated scope ambiguity (ambiguous vs. unambiguous) resulting in stimuli such as in (1).

1. a. Every kid climbed a tree. The tree was in the park.
   b. Every kid climbed that tree. The tree was in the park.

In a moving window paradigm, the participants were asked to perform a word-by-word, self-paced reading task (Experiment 1) and also to answer scope-related comprehension questions such as *How many trees were climbed?* (Experiments 2 and 3). Consistent with previous findings, they observed that the plural context sentence (i.e., the first sentence in 1a) read faster than the singular context sentence (i.e., the first sentence in 1b), but there was no reading time differences between continuation sentences (i.e., no difference between the second sentences in 1a and 1b). This suggests that the context and continuation sentences are processed in a shallow manner; were they processed deeply, the second sentence in singular ambiguous stimuli (1a) should have taken

longer to read than that with unambiguous stimuli (1b). However, question–response accuracy rates revealed lower accuracy for the singular ambiguous sentences than for unambiguous sentences, suggesting that algorithmic processing is invoked when task demands (comprehension questions) require it. Based on these results, Dwivedi (2013) argued that sentence processing primarily proceeds with heuristics and that deep, syntax-based, and algorithmic processing is invoked only afterwards and only if necessary.

However, there is also evidence showing that heuristic and algorithmic processes are simultaneously active during comprehension. For example, assuming that heuristic processing draws on semantic information and algorithmic processing on syntactic information, Lim and Christianson (2013a) had non-native speakers of English read for comprehension as well as verbally translate sentences from English to their own native language (Korean). Although the results showed that non-native speakers rely more heavily on semantic (plausibility) information than native speakers do during early stages of sentence processing, the results from later stages of processing from both tasks suggested that non-native speakers are able to access both the semantic and the syntactic information in real time. In a separate study, Lim and Christianson (2013b) found, in a translation task, that non-native speakers of English preserved the original syntactic structure of the to-be-translated sentences but sometimes switched the thematic roles of the event participants, suggesting that syntactic as well as semantic information are both available and used during sentence processing. Based on these results, Christianson and colleagues argued that language processing proceeds through two main routes operating in parallel—a semantic route and a syntactic route (Kuperberg, 2007)—and that good-enough representations arise because the outputs of these two routes are not successfully integrated.

The results from Dwivedi (2013) and those from Christianson and colleagues (Lim & Christianson, 2013a,b) run counter to each other in that the former supports a heuristics-before-algorithms architecture whereas the latter is consistent with the simultaneous operation of both heuristic and algorithmic routes. However, one possibility is that heuristic and algorithmic processing launch simultaneously but the former is faster, creating the impression that it precedes algorithmic processing. We return to this possibility when we propose our model below.

In sum, current psycholinguistic evidence shows that intrasentential representations may be shallow and imprecise presumably because heuristic processing is heavily involved in language processing along with deeper, algorithmic processing. In the following section, we outline evidence for shallow intersentential representations.

Good-enough intersentential (discourse) processing

Good-enough representations have also been shown to occur intersententially. One important case of underspecification in discourse is reference processing. Previous research has shown that the referents of ambiguous and even unambiguous pronominal referring expressions might remain unresolved during discourse processing (Greene et al., 1992; Levine et al., 2000). For example, using both probe recognition and self-paced reading tasks, Levine et al. (2000) demonstrated that if there is substantial backgrounding between the antecedent (*tart*) and the anaphor (*dessert*), and if a same-category and salient distractor (*cake*) intervenes between the antecedent and the anaphor, then reference resolution fails to take place. Interestingly, in their final experiment (Experiment 6), they showed that if participants were motivated to devote more attentional resources to the line containing the referring expression (by placing asterisks around the line containing the referring expression and explicitly informing them that comprehension questions will address information on that line), successful reference resolution did occur. From these results, Levine et al. (2000) argued that reference resolution is a function of two factors: first, the degree of accessibility of the antecedent, with reference resolution being more likely to take place if the antecedent is more accessible; and two, the extent to

which the resolution is necessary to establish a coherent discourse representation.

Following up on the factors that increase the probability of reference resolution, Klin, Weingartner, Guzman, and Levine (2004) investigated whether text factors (as opposed to instructional manipulations) would also lead to full reference resolution. To this end, they manipulated the perceived salience of the referring expression by means of syntactic focus (wh-clefts) and certain kinds of prenominal adjectives (*sumptuous*, *good*). They observed that reference resolution is influenced by perceived salience such that textually salient antecedents were more likely to be retrieved upon processing the associated referring expression, indicating that readers are efficient information processors and that they are sensitive to subtle linguistic cues in discourse. Similarly, in a text change paradigm, Sanford, Sanford, Molle, and Emmott (2006) showed that attentional capture devices such as italicization in written discourse and focus-driven stress in spoken discourse increase depth of processing.

Taken together and consistent with the GE approach to language processing, the results reviewed in this section show that discourse processing is sometimes superficial and fails to lead to successful reference resolution. However, when the task requires or at least encourages comprehension, the referent of the referring expression seems to be identified. In addition, when language users are implicitly motivated to allocate more attentional resources to reference processing (such as through linguistic attention capturing tools, text-based salience, or increased engagement), reference resolution is more likely to occur.

The online equilibrium hypothesis

So far, we have discussed evidence showing that both intra- and intersentential language processing is subject to good-enough processing such that the resulting representations may not be very detailed and accurate. We have also discussed work showing that task demands such as comprehension questions, textual perceived salience such as italicization, and enhanced engagement encourage greater depth of processing. However, the critical question is: What is the general underlying force that influences depth of processing? Borrowing insight and terminology from the theory of cognitive development proposed by Jean Piaget (Piaget, 1952, 1977, 1985), we suggest that whenever the language processing system is presented with a sentence or a piece of discourse, it is in fact presented with a cognitive challenge that disturbs "cognitive equilibrium" and that the language processing system is sensitive to *online* states of cognitive equilibrium and disequilibrium. Once the sentence or set of sentences is successfully processed or is believed to have been successfully processed, cognitive equilibrium is restored. Before going into the details of our online cognitive equilibrium hypothesis, let us briefly review Piaget's original notion of cognitive equilibrium.

Piaget's theory is primarily concerned with cognitive development and assumes that people learn to adapt to their environment. This adaption takes place through two different processes: assimilation and accommodation. Assimilation involves transforming information from the environment so that it fits preexisting cognitive structures called schemata, and accommodation refers to the alteration of existing cognitive structures so that new information can be incorporated within the current cognitive structure. Accommodation necessarily results in changes to existing schemata. According to Piaget, cognitive development reaches its ideal when a balance is struck between assimilation and accommodation—that is, between the current cognitive structure and the environment. He termed this state of balance "cognitive equilibrium". When equilibrium is achieved, the child has the schematic knowledge required to learn new information on her own. Disequilibrium occurs whenever children encounter new information that does not fit existing cognitive structures. For this new information to be processed, either it should be transformed or existing schemata should change to accommodate it so that equilibrium can be restored (Piaget, 1952, 1977, 1985).

We posit that language processing similarly reflects states of online cognitive equilibrium and disequilibrium and that equilibrium is the default

and desired cognitive state. As mentioned above, we propose that when the language processing system encounters the first words in a sentence, it enters a state of disequilibrium. This is because the sentence contains unprocessed information that poses uncertainty and must therefore be processed and integrated with the existing schemata to achieve a more certain and stable cognitive state (this idea is similar in many ways to Gibson's proposal that dependencies in language processing lead to processing costs; Gibson, 1998). In other words, a presented sentence creates online disequilibrium or cognitive challenge that needs to be resolved so that cognitive equilibrium can be restored. Importantly, achieving equilibrium is dependent on computing the general meaning of the sentence and successfully incorporating it in existing schemata. Since linguistic information unfolds over time, it necessarily takes some time for all relevant constituents to be delivered so that the general meaning becomes clear, and the disequilibrium can be resolved. As a result, processing a sentence(s) can be viewed as entering a state of disequilibrium and then moving towards a state of equilibrium.

It is well established that the language processing system processes input incrementally until its meaning is understood. Our proposal is related but different: We suggest that language processing is incrementally sensitive to changing states of equilibrium in real time and that many psycholinguistic findings can be explained by appeal to this sensitivity. We also propose two general principles for the online cognitive equilibrium hypothesis (henceforth OCE) that specify this sensitivity:

1. The cognitive system attempts to maximize equilibrium *at the earliest opportunity*.
2. Once at equilibrium, the language processing system prefers to stay in that state as long as possible and as long as there is no strong reason for abandoning that state.

It is important to highlight the relationship between these two principles of OCE and the GE approach to language processing, and to the distinction between heuristic versus algorithmic processing. We believe that the OCE approach provides a possible underlying mechanism for good-enough processing. Specifically, the reason why sometimes only fast and frugal heuristics rather than deep and time-consuming algorithms are applied during comprehension could be because heuristics offer a faster route to equilibrium (Principle 1). Similarly, the reason why the system is sometimes satisfied with a good-enough representation and does not exert the extra effort to engage in deeper processing could be because heuristics often provides enough equilibrium for the system, causing it to stay in that state for as long as possible, which prevents allocation of the full resources necessary for deeper processing (Principle 2).

It is also important to specify the relative timing of the heuristic and the algorithmic routes and how it relates to achieving equilibrium and the formation of good-enough representations. Figure 1 provides a graphical representation of a model of sentence processing that operates based on cognitive equilibrium. As can be seen in this figure, we believe that heuristics and algorithms launch at the same time. However, because heuristic processing is based on the implementation of simple rules (perhaps "rules of thumb"), it usually delivers an output before the algorithmic route. This way, the system reaches equilibrium faster by reaching an "interim" output, so to speak (Principle 1). Note that this output is perhaps not the best representation that the system could possibly build out of the linguistic input because it has been constructed using simple, error-prone heuristics. Crucially, when the output of heuristic processing becomes available, the algorithmic route is still not finished, and therefore the interim output of the heuristic route influences the ongoing algorithmic processing. We propose that it is this influence that results in the formation of good-enough linguistic representations: When the output of heuristic processing is available, the system reaches a state of equilibrium and therefore prefers to stay in that state, causing the system not to allocate further resources for algorithmic processing (Principle 2). In other words, the formation of an interim output by heuristic route causes the algorithmic route to become confirmatory in nature, leading

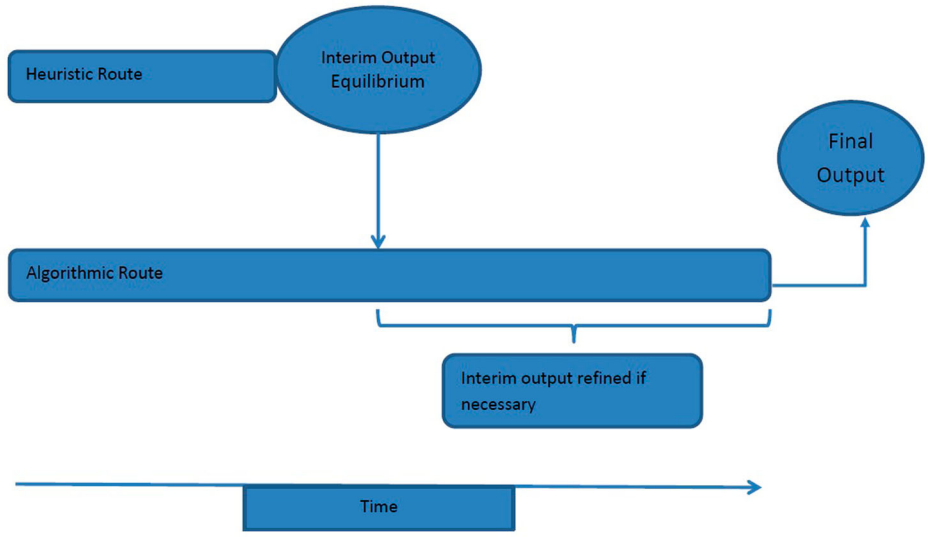

Figure 1. *The model of language processing according to the online equilibrium hypothesis.*

to good-enough final representations (for a similar approach, see also Townsend & Bever, 2001).

Although some studies have identified a few potential heuristics (e.g., F. Ferreira, 2003), the nature of the simple rules that guide heuristic processing is unclear. However, we believe that this processing relies more heavily on top-down information from semantic memory, whereas algorithmic processing seems to rely more heavily on linguistic knowledge to derive meaning in a bottom-up way, by organizing and combining the unfolding input using well-defined, successive linguistic rules. It is also not the case that the operation of the heuristic route and the formation of the interim output lead to a simple binary result whereby full equilibrium is either achieved or not. Instead, the interim output produces degrees of equilibrium. Crucially, the more certainty there is about the accuracy of the interim output, the greater the equilibrium and the more the system will rely on the interim output, which will, in turn, lead to more confirmatory/shallow algorithmic processing. The degree of certainty and therefore reliance on the interim output may depend on a number of factors such as how relevant and compelling the retrieved schemata and the memory-based representations are. Importantly, however, one influential factor could be task demands, such that when the task explicitly encourages deeper processing, the system will rely less on the interim output and more on the bottom-up, algorithmic route (e.g., F. Ferreira, Foucart, & Engelhardt, 2013; Swets et al., 2008).

At the same time, the interim output of the heuristic route is also influenced by ongoing algorithmic processing. That is, this output is refined by algorithmic processing if there is strong bottom-up evidence supporting a modification in the interim output. In other words, if there is enough evidence from the output of the algorithms to revise the heuristics-based interim output, the system will fall back to disequilibrium and will therefore allocate more processing resources to restore equilibrium. Thus, the interim output created from heuristics gets refined if necessary, and once the algorithmic processing is over, the final output is produced. Obviously, if the algorithmic-based evidence confirms the heuristics-based output, that output will not be modified.

This model of sentence processing is similar to recent multistream models of language processing (e.g., Hagoort, Baggio, & Willems, 2009; Jackendoff, 2007; Kuperberg, 2007; van Herten, Chwilla, & Kolk, 2006; van Herten, Kolk, &

Chwilla, 2005) and runs counter to models that assume a single stream to meaning construction (e.g., Boland & Tanenhaus, 1991; F. Ferreira & Clifton, 1986; Frazier & Rayner, 1982; Trueswell & Tanenhaus, 1994; see Brouwer, Fitz, & Hoeks, 2012, for a detailed discussion). A well-known example of a multistream model is the dual-route model of Kuperberg (2007), according to which comprehension is performed through two independent but highly interacting routes: a semantically driven heuristic route and a syntactically driven algorithmic route. Importantly, in Kuperberg's (2007) model, the syntactic route can be influenced by the semantic information that becomes available via the heuristic route. Our model depicted above also makes use of two independent but highly interacting routes. However, there are two main differences between our model and that of Kuperberg (2007). First, unlike Kuperberg's model in which the heuristic route is responsible only for creating semantic relations between individual content words but not for creating global representations, the heuristic route in our model is capable of generating a global sentence meaning. In fact, one of the fundamental assumptions of our model is that an early global meaning (i.e., interim output) is formed by the heuristic route before algorithmic processing is completed (see also Townsend & Bever, 2001). Second, unlike Kuperberg's model, ours specifies that heuristic and algorithmic routes launch simultaneously, but heuristic processing reaches completion earlier because it proceeds through simpler rules and makes use of less information.

Let us now apply the equilibrium hypothesis and its associated model to intrasentence and intersentence processing. We focus somewhat more heavily on reference processing (intersentence processing) as this topic has not received enough attention in the literature researching into good-enough language processing.

Online cognitive equilibrium and intrasentential underspecification

Under the online cognitive equilibrium hypothesis, intrasentential good-enough representations might be formed because the amount of equilibrium in the interim, heuristics-based output is sometimes so high that it preempts algorithmic processing. More specifically, if the amount of disequilibrium that a word or structure causes through the heuristic route is small, the algorithmic route will not allocate enough resources to processing the bottom-up input, resulting in the formation of less-than-veridical representations.

For example, based on OCE, the Moses illusion might be assumed to occur because the high degree of semantic overlap between the right and the wrong agents (*Moses* and *Noah*) creates little disequilibrium, resulting in heavy reliance on the interim output and causing the ongoing algorithmic processing to become rather confirmatory, which, in turn, may result in failure to detect the anomaly. Similarly, the burying-the-survivors example could be caused by the high amount of semantic overlap between *survivors* and other possible words that could replace it (e.g., *casualties*, *victims*, etc.), which may have led to small disequilibrium at heuristic processing, resulting in a heavier influence of its output on algorithmic processing, which in turn might have led to failure to detect the anomaly. Finally, for the sentence *the dog was bitten by the man*, our proposal is that when the concepts associated with *dog* and *biting* are activated, a compelling schema containing the information that dogs bite people leads to the formation of an interim event representation in which such is the case, creating strong equilibrium. This output then influences the ongoing algorithmic processing so heavily that the accurate syntactic parse of the sentence is contaminated, as it were, by this output, creating an erroneous sentence representation.

This possibility is strengthened by the fact that, in the correct event representation, *the dog* is the agent of the action of biting. Because syntactic subjects and thematic agent roles often coincide (Christianson & Ferreira, 2005; F. Ferreira, 1994; Osgood & Bock, 1977; Reinhart, 1982; Tomlin, 1995), the mapping of a noun–verb–noun syntactic template to an agent–verb–patient template is valid up until the verb (*dog* plus *bite*). As such, little initial disequilibrium is produced, leading to an

interim output with a high degree of equilibrium and shallow subsequent algorithmic processing.

It is important to note that although influenced by the output of the heuristic route, the algorithmic route still independently processes the unfolding information in a bottom-up manner and often successfully detects any anomalies, as evidenced by the fact that implausible passive sentences are usually interpreted correctly (74% of the time in F. Ferreira's, 2003, study). Moreover, it has been shown that manipulations of the linguistic input through cleft constructions result in increased detection rates in the Moses illusion (Bredart & Modolo, 1988).

Our model of sentence processing is also compatible with studies that have found evidence for the simultaneous availability of syntactic as well as semantic information even in underspecified representations. For example, in a follow-up study to F. Ferreira (2003), Christianson, Luke, and Ferreira (2010) examined the potential influence of syntactic and semantic complexity on language production using a picture description task. Interestingly, they observed that the syntactic structures of the productions were influenced not only by the syntactic form of the primes (e.g., Pickering & Ferreira, 2008), but also by their plausibility. Specifically, both plausible passives and implausible actives primed passive productions. These results were taken to suggest that both semantic and syntactic information are available in real time and that neither source of information in abandoned in favour of the other (see also, van Herten et al., 2006). Good-enough representations, according to the authors, are the result of a compromise designed to reconcile the outputs of syntactic and semantic routes.

Similar to the explanation of Christianson et al. (2010), under our model, these results can be explained by the interaction between the heuristic and the algorithmic routes. Specifically, it is possible that when processing an anomalous active sentence ("man biting dog"), the plausible representation (i.e., "dog biting man") has also been activated through the heuristic (top-down) route. Since *the man* is the syntactic subject in the given sentence, but is the patient in the correct event representation, the combination of the syntactic role from the actual output (i.e., subject) and the thematic role from the activated correct event representation (i.e., patient), might have led to the production of the passive structure in the picture description task. Similarly, the reason why passive-implausible sentences ("the dog was bitten by the man") did not prime passive voice productions in that study could be because the activation of the correct event representation might have resulted in the activation of an agent concept for *dog*. However, in the actual input, *dog* is the syntactic subject of the sentence. The combination of the correct thematic role from the correct event representation (i.e., agent) and the syntactic role from the actual linguistic input (i.e., subject) might have led to more active than passive productions. However, in the case of plausible passives ("the man was bitten by the dog"), the plausible (correct) event representation has probably not been activated through the heuristic, top-down route, resulting in more reliance on surface structure and leading to more passive productions.

Online cognitive equilibrium and intersentential processing and underspecification

How would the online equilibrium hypothesis apply to intersentential (discourse) in general and reference processing in particular? It seems reasonable to assume that in the case of reference processing, when a sentence containing potential antecedents is processed, a new semantic representation is established. The following sentence containing the referring expressions now has to be incorporated with this newly built schema (i.e., the initial sentence) for equilibrium to be achieved. Therefore, when the referring expression is encountered, it needs to be resolved for equilibrium to take place.

To relate this back to our model, we argue that when the referring expression is encountered, the fast-going heuristic route sets out to quickly establish the correct referent for the referring expression to achieve equilibrium as soon as possible (Principle 1). Once the referential problem is solved, the system reaches and prefers to stay in

interim equilibrium, and therefore slow-going algorithmic processing is influenced by the output of referential processing (Principle 2). Again, the amount of influence from the heuristic-based interim output is contingent on the amount of equilibrium it produces: The more equilibrium is produced by the interim output, the more it will influence algorithmic processing and vice versa. As such, the depth of processing of a referring expression is a function of the amount of disequilibrium induced by the referring expression: The more disequilibrium is created by the referring expression, the smaller the effect of the interim output will be on the algorithmic processing, resulting in deeper processing of that referring expression. The flip side of this is that the more equilibrium there is at the time of processing the referring expression, the more influential the output of the heuristic processing will be on the ongoing algorithmic processing, resulting in shallower processing of the referring expression.

Recent research on anaphora lends support to an OCE approach towards reference processing. Specifically, recent evidence shows that the initial stages of reference processing are based on quick semantic fit between the referring expression and the potential referents but that later stages are more bottom-up in nature. For instance, Cook (2014) investigated the effect of goodness-of-fit between an anaphor and the referent. Unambiguous anaphors were correct, incorrect but semantically highly related, or incorrect but unrelated to preceding referents. The results showed that incorrect but semantically related anaphors were read faster than incorrect and semantically less related anaphors, suggesting that goodness-of-fit is a driving force underlying reference resolution. However, less related anaphors continued to cause processing difficulty in the spillover region whereas highly-related anaphors did not produce such an effect, suggesting, according to the authors, that referring expressions might be checked against the context for referential (as opposed to semantic) analysis in a later stage.

Cook's (2014) findings are consistent with OCE, because based on this model, it could be argued that upon encountering the referring expression, the heuristic route attempts to quickly establish reference. When there is enough semantic overlap between the referring expression and a potential referent, the heuristic route simply links the two (although it might be wrong), producing early equilibrium. In the case of semantically less related referring expressions, disequilibrium would be greater, which reduces the amount of influence from the interim output on the algorithmic processing, leading to more processing difficulty for semantically less related referring expressions.

A question now arises as to what determines the degree of disequilibrium that a referring expression induces? We propose that the magnitude of disequilibrium incurred by a referring expression is inversely related to its efficiency in successfully distinguishing the referent from other competing antecedents. In other words, the more efficient the referring expression is at distinguishing the referent from competing antecedents, the less disequilibrium it will induce, and vice versa. This is because the system needs to resolve the referring expression before and in order to arrive at equilibrium. As such, the efficiency of the referring expression in establishing reference will influence cognitive equilibrium. Importantly, the distinguishing efficiency of the referring expression is dependent not only on the linguistic features of the referring expression itself such as gender or number, but also on how conceptually confusable the potential antecedents (including the referent) are as well. If the antecedents are highly confusable (for instance, if they are two human noun phrases of the same gender), then the referring expression should be more discriminating towards the correct referent for equilibrium to be achieved as early as possible, but when the antecedents are not confusable, the discriminating efficiency of the referring expression need not be very high.

As an example, given the sequence *The cowboy talked to the sheriff. He was very jokey*, reference to either of the antecedents with just the pronoun *he* would not be efficient because it prevents a quick identification of the correct referent; a (repeated) noun phrase referring expression (e.g., *the cowboy*) would be better. However, if *sheriff* is replaced

with *cowgirl*, then a pronoun would suffice to single out the correct referent. As such, an ambiguous referring expression would produce more disequilibrium than an unambiguous referring expression because the former would make it difficult to pick out the correct referent and thereby establish coherence (and therefore equilibrium) by incorporating the sentence containing the referring expression with the preceding sentence containing the antecedents. However, an unambiguous pronoun would allow quick and efficient resolution of the referring expression, which would lead to early equilibrium. This view of reference resolution is consistent with theories maintaining that language has primarily evolved as a means of communication, and therefore efficiency in conveying messages is a central principle guiding language processing (e.g., Fine & Jaeger, 2013; Jaeger & Snider, 2013; Piantadosi, Tily, & Gibson, 2011, 2012; Qian, & Jaeger, 2012; Smith & Levy, 2008).

The online cognitive equilibrium view of reference resolution can explain a wide range of established findings in the reference processing literature. For example, it has been found that referential ambiguity results in processing difficulty, whereas unambiguous referring expressions produce no such difficulty. In other words, previous findings show that the resolution of unambiguous pronouns is rapid and automatic, whereas ambiguous pronouns lead to difficulty. For example, Rigalleau and Caplan (2000) measured naming latencies for pronouns following an antecedent. The results showed a clear effect of gender congruence between the pronoun and the antecedent, with congruent pronoun-antecedent pairs resulting in shorter naming latencies and incongruent pronoun–antecedent pairs producing longer naming tendencies. Based on these results, the authors argued for an immediate and automatic process of gender matching between a pronoun and the referent. Similarly, employing a reading-time paradigm, Rigalleau, Caplan, and Baudiffier (2004) confirmed that a gender mismatch between the antecedent and the pronoun resulted in increased reading times, suggesting that gender coindexation between a pronoun and its associated antecedent is nonstrategic and automatic.

Moreover, in an eye-tracking visual world study, Arnold, Eisenband, Brown-Schmidt, and Trueswell (2000) found evidence for rapid use of gender information in pronominal anaphora processing, with quick looks to the image associated with the gender-matching pronoun. Similarly, Sanford and Filik (2006) showed that there is a processing cost associated with the gender-unspecified, singular *they* due to number incompatibility in sentences such as *I was looking for the railway station when I saw someone on the other side of the street. I asked them if they knew where it was.*

These results can be explained by arguing that when the referring expression contains enough information to efficiently distinguish between the referent and the competing antecedents, the heuristic route quickly resolves the referential problem through a simple coindexation process and produces early equilibrium. In other words, the processing of unambiguous referring expressions is facilitated because the comprehension system quickly reaches equilibrium by establishing the referential link between the referring expression and the antecedent through a simple, quick, and heuristics-based coindexation process, leading to little if any processing difficulty. In fact, because there is enough information in the linguistic input itself for referential resolution, it is highly likely that in the case of unambiguous pronoun resolution, the algorithmic route reaches completion almost as early as the heuristic route does, leading to no apparent asynchrony in their respective outputs.

In contrast to unambiguous pronouns, ambiguous pronouns have been shown to lead to processing difficulty. For example, in a probe recognition task, MacDonald and MacWhinney (1990) presented participants with sentences containing two antecedents and found that reaction times to referents are shorter, and reaction times to nonreferents are longer, following a referring pronoun, suggesting that pronouns facilitate retrieval of the referent but inhibit that of the nonreferent. Moreover, and more importantly, they also found that this facilitation effect is substantially delayed when the pronoun is ambiguous than when it is unambiguous. Research using the visual world paradigm has also demonstrated that people distribute their gaze almost

evenly on all antecedents when processing referentially ambiguous anaphoric expressions (Chambers, Tanenhaus, Eberhard, Filip, & Carlson, 2002; Sedivy, Tanenhaus, Chambers, & Carlson, 1999; Spivey, Tanenhaus, Eberhard, & Sedivy, 2002).

The processing difficulty induced by ambiguous referring expressions has also been demonstrated using methods from cognitive neuroscience such as the recording of event-related potentials (ERPs). For instance, building on the discovery that referential ambiguity has its own unique ERP signature (van Berkum, Brown, & Hagoort, 1999; van Berkum, Brown, Hagoort, & Zwitserlood, 2003), which is the same for both noun phrase and pronominal anaphoric expressions (van Berkum, Zwitserlood, Bastiaansen, Brown, & Hagoort, 2004), Nieuwland and van Berkum (2006) showed that ambiguous pronouns elicit a sustained frontal negative shift reflecting their processing difficulty, but unambiguous pronouns produced no such shift in the ERP signal. Interestingly, the magnitude of this referential ambiguity effect was significantly influenced by contextual bias such that when the two antecedents were more balanced in terms of probability of being the referent of the ambiguous pronoun, the ERP effect was stronger, but when the context biased the interpretation of the pronoun towards one of the antecedents, the effect was weaker. This pattern of results further suggests that the more ambiguous the referent of the pronoun, the more processing difficulty the system experiences.

Under the online equilibrium hypothesis, these findings can be explained by arguing that the amount of disequilibrium caused by ambiguous referring expressions (such as ambiguous pronouns) is higher than that produced by unambiguous referring expressions. In fact, the heuristic route cannot easily establish reference in the case of ambiguous referring expressions, and as a result, no early equilibrium can be achieved. Thus, finding the correct referent is automatically pursued by algorithmic route. Moreover, since heuristics cannot really operate, algorithmic processing of reference is pursued with little influence from the heuristic route. Since algorithmic processing is more dependent on bottom-up input, pronominal ambiguity takes longer to complete (because enough contextual information for reference resolution should be found in the linguistic output before the ambiguity could be resolved). As such, the system remains in disequilibrium for a longer time, creating processing difficulty.

Consistent with this explanation, there is evidence that the resolution of ambiguous pronouns is delayed (e.g., MacDonald & MacWhinney, 1990; Stewart et al., 2007). For example, employing self-paced reading-time paradigm, Stewart et al. (2007) showed that the resolution of ambiguous pronouns is delayed whereas the processing of unambiguous pronouns is not. Specifically, they provided evidence that the processing system makes an initial but shallow commitment with regard to the referent of the ambiguous pronoun and waits for later additional disambiguating information before fully resolving the referential ambiguity. As such, there appears to be no initial and automatic commitment to a certain interpretation of a pronoun upon encountering it.

However, one of the most interesting and challenging paradoxes about referential ambiguity is that it has been shown to inhibit referential processing (as discussed above) as well as to facilitate it. Evidence for the facilitative effect of referential ambiguity comes from studies conducted on relative clause attachment resolved by reflexive pronouns (Traxler, Pickering, & Clifton, 1998; van Gompel, Pickering, Pearson, & Liversedge, 2005; van Gompel, Pickering, & Traxler, 2001). For instance, using stimuli such as (2), Traxler et al. (1998) showed that the reflexive pronoun and immediately following regions of the sentence were read faster when the pronoun was ambiguous (2c) and more slowly when it was unambiguous (2a and 2b).

2. a. The daughter$_i$ of the colonel$_j$ who shot herself$_{i/*j}$ on the balcony had been very depressed.
   b. The daughter$_i$ of the colonel$_j$ who shot himself$_{*i/j}$ on the balcony had been very depressed.
   c. The son$_i$ of the colonel$_j$ who shot himself$_{i/j}$ on the balcony had been very depressed.

The paradoxical effect of referential ambiguity could be explained by the online equilibrium hypothesis. Specifically, a closer scrutiny of Traxler et al.'s (1998) results reveals that the two potential antecedents are ontologically one single entity, not two. That is, *the daughter of the colonel* refers to just one person in the outside world, not two, despite the fact that there are two different nouns in this complex noun phrase. As such, it is plausible that this noun phrase is stored and represented in working memory as a single complex entity encompassing both antecedents, resulting in a single, merged referent. In other words, the representation of *the daughter of the colonel* is one entity that merges the singular representations of *daughter* as well as *colonel*. If this is the case, then an unambiguous reflexive referring to *the daughter of the colonel* would disrupt equilibrium, because it would force the processing system to distinguish the antecedent representations in the merged representation—in this case, the feminine entity from the one that is masculine. This requirement to distinguish the antecedents that have been stored as a merged representation would produce disequilibrium because a referential link is forced between only one of the antecedents and the referring expression. That is, when the unambiguous reflexive pronoun is encountered, the heuristic route cannot establish reference and produce early equilibrium, which leads to algorithmic processing assuming a more prominent role in establishing reference, which manifests in longer reading times.

Note that as shown by previous research, the complex representation is eventually distinguished, resulting in the resolution of attachment ambiguity (see Swets, Desmet, Hambrick, & Ferreira, 2007), but the resolution is more difficult than the resolution of ambiguous antecedents. When the referring expression is ambiguous, there is no requirement for the referring expression to distinguish between the individual antecedents in the merged representation of the complex noun phrase referent. In other words, the referring expression is allowed to refer to the merged entity as a whole. As such, little, if any, disequilibrium is induced by the ambiguous reflexive pronoun because the heuristic route can operate quickly, producing enough equilibrium to reduce the depth of processing of the algorithmic route and to eventually result in faster reading of ambiguous reflexive pronouns than of unambiguous reflexive pronouns.

Recent evidence supports the idea that comprehenders sometimes form merged representations. Specifically, Poesio, Sturt, Artstein, and Filik (2006) showed that underspecified anaphoric expressions are special representations that cover all the potential preceding referents. Their work is based on Frazier and Rayner's (1990) finding that homonymous words (i.e., ambiguous words that have two unrelated meanings, such as *pitcher*) result in garden-pathing when an unexpected meaning must be instantiated, but polysemous words (i.e., ambiguous words with two or more closely related meanings, such as *newspaper*) do not produce that same garden-path or surprisal effect. From this, Frazier and Rayner (1990) argued that because the different meanings of polysemous words are closely related, an initial preliminary/partial interpretation would be activated covering all the meanings of the polysemous word, and weakening the requirement that words need to be immediately and fully interpreted (e.g., F. Ferreira & Clifton, 1986; F. Ferreira & Henderson, 1991; Fodor & Inoue, 1998, 2000; Frazier & Fodor, 1978; Pickering & Frisson, 2001).

Extending Frazier and Rayner's (1990) work to anaphora, Poesio et al. (2006), focused on what is known as mereological pronouns: Pronouns that have no specific antecedent. For example, in (3), it is unclear whether the pronoun *it* refers to the engine, to the boxcar, or to both.

3. Hook up engine 2 to the boxcar and send it to Corning.

Mereological pronouns are similar to the case of polysemous words because they too can receive an initial preliminary interpretation that covers all the possible interpretations (in this case, antecedents). As such, they also provide a useful situation for investigating the nature of good-enough representations in relation to referring expressions. This is because once the two separate entities (e.g., the engine and the boxcar) are merged, they become one new entity, and therefore from the point of

view of the listener, all that matters is that the pronoun interpretation restricts the potential referents to the entity obtained through the merger. After confirming that mereological pronouns are fairly common in corpora, Poesio et al. (2006) provided offline (acceptability judgement) and online (eye-tracking) evidence supporting the view that mereological pronouns invoke a good-enough representation. Specifically, sentences such as (4a), which contain a verb that gives rise to a mereological pronoun (*hook up*), were judged to be more acceptable than sentences with a preceding context that did not permit a mereological pronoun (4b).

4. a. The engineer hooked up the engine to the boxcar and sent it to London.
   b. The engineer separated the engine to the boxcar and sent it to London.

In addition, in an eye-tracking experiment, they showed that sentence regions containing mereological pronouns are read faster than control sentences containing a normal plural pronoun, suggesting that the referent associated with the singular pronoun might indeed have received an underspecified interpretation.

Thus, similar to mereological pronouns that result in formation of a representation that represents all the preceding antecedents, syntactically complex noun phrases such as *the daughter of the colonel* might also result in a representation that merges both individual entities in the complex noun phrase. The reason for forming such a merged representation could be that the entire noun phrase refers to just a single entity in the world, or the merger might occur simply because the noun phrase is syntactically complex, and therefore it is easier and more efficient to underspecify the representation by forming a merged and sketchy representation. Whatever the reason, the bottom line is that such a merged representation would be easier to process when referred to by an ambiguous reflexive pronoun than with an unambiguous one because an ambiguous referring expression would facilitate establishing reference between the pronoun and the merged representation, whereas an unambiguous referring expression would require conceptually breaking apart the merged representation, which would in turn result in disequilibrium and therefore processing difficulty.

Online cognitive equilibrium and the production of referring expressions

The online equilibrium approach to reference processing can also explain several established findings regarding the production (as opposed to the comprehension) of referring expressions. One relevant phenomenon is that the presence of a competing antecedent as well as the semantic similarity between the antecedents results in greater use of noun phrase anaphora overall (i.e., regardless of which antecedent is referred to), whereas the absence of a competitor or lack of semantic similarity between antecedents results in greater use of pronouns.

For example, Arnold and Griffin (2007) presented their participants with two-panel cartoons. Participants saw the first panel and simultaneously heard the first sentence of a story, and then they were asked to generate the next sentence based on the second panel of the cartoon. The manipulations were whether a second character was present in the cartoons and also whether the second character, if present, was of the same or a different gender to the main character [*Mickey went for a walk (with Donald/Daisy) in the hills one day*]. Since in all critical items the target character was the subject of the first sentence, no difference in pronominalization rates in the one-character versus two-character conditions was expected. Nonetheless, the results showed that speakers tended to use fewer pronouns to realize an entity when a second entity was also present in the scene than when there was no second entity. Additionally, they found a robust significant difference between same-gender and different-gender conditions, with significantly more repeated noun referring expressions in the same-gender condition. They attributed this effect to the semantic interference resulting from having two characters of the same gender available as potential referents. Note that since there was a significant difference between the one- and two-character conditions, the gender effect could not be attributed to ambiguity avoidance.

More recently, Fukumura, van Gompel, and Pickering (2010) replicated the effect of the presence of competing antecedents, and Fukumura and van Gompel (2011) also found the same semantic interference effect with animacy rather than gender: They observed that participants used more repeated nouns when two antecedents were both animate than when one was animate and the other inanimate.

According to OCE, in order for the referring expression to result in equilibrium, it should efficiently distinguish between the referent and other competing antecedents so that incorporation of the two sentences could be achieved as quickly as possible. As such, when there are two antecedents instead of just one, there will necessarily be some degree of confusability between them, which would then require that the referring expression be more discriminating towards its referent. Thus, the effect of having a competing antecedent in the preceding discourse could be because repeated nouns are more distinguishing than pronouns (because they pick out their referents with more certainty).

Similarly, the semantic similarity effect could be explained by arguing that similarity necessarily produces confusability, requiring that the referring expression be more discriminating towards its referent, which in turn results in the use of more noun phrase referring expressions. Notice too that our explanation is agnostic between speaker-based (e.g., Arnold & Griffin, 2007; V. S. Ferreira, Slevc, & Rogers, 2005; Fukumura & van Gompel, 2012) and audience-design (e.g., Brennan & Clark, 1996; Brown-Schmidt & Tanenhaus, 2006; Horton & Gerrig, 2002, 2005) accounts of referent formulation; that is, speakers might try to reduce confusion in their own representations to facilitate their production of utterances, or they might attempt to reduce listener confusion (or both). Our account is compatible with any of these views.

Challenges to referential processing under OCE

Earlier we discussed research showing that sometimes referring expressions are not given specific referents. For example, Levine et al. (2000) suggest a model of anaphor resolution in which referring expressions may be left unresolved (see above). Further evidence for this model comes from Greene et al. (1992), who presented participants with stories such as (5) via rapid serial visual presentation (RSVP) and asked them to respond "Yes" to a probe if it had appeared in the text and "No" if it had not.

5. Mary and John were doing the dishes after dinner.
   One of them was washing while the other dried.
   Mary accidentally scratched John with a knife and then she dropped it on the counter.

The probe words were the two human characters introduced in the story (*Mary* or *John*), or a nonreferent control (*dishes*). The probes appeared immediately before the pronoun in the last line (*she*), after the word following the pronoun (*dropped*), or at the end of the last line. They found no difference in reaction times to the referent and the nonreferent probes (i.e., *Mary* vs. *John*), suggesting that the correct referent of the pronoun was not accessed following the processing of the pronoun.

Their results were later replicated by Love and McKoon (2011). However, Love and McKoon showed that when extra text is added between the referents and the pronoun, automatic resolution of the pronoun does take place. Specifically, Love and McKoon lengthened Greene et al.'s (1992) stories from four to eight lines, creating stories such as (6), and observed a reaction time advantage for the referent compared with the nonreferent. Interestingly, this effect of lengthening the stories did not depend on whether the extra lines rementioned the referents. Based on these results, Love and McKoon argued that automatic pronoun resolution in longer stories is due to the greater engagement of participants in the stories, which presumably led to greater attentional resources being allocated to processing and consequently to the resolution of the pronoun. In other words, mental effort is required to bind a referent to a pronoun, and readers do not expend that effort unless the task is interesting or significant to them.

6. Rita and Walter were writing an article for a magazine.
   They had to get it done before next Tuesday.
   Rita didn't trust Walter to get the facts right.
   Once, he'd written a piece about aliens landing in Chicago.
   "I'm going to get dragged down with you," Rita said at the time.
   However, neither of them had been fired.
   Rita edited the section Walter had written and then she smoked a cigarette to relax.

Obviously, nonresolution of referring expressions is inconsistent with the online equilibrium hypothesis, according to which referring expressions should be resolved so that equilibrium is achieved. However, more recent evidence suggests that if the processing measures are sensitive enough and/or if the manipulations are strong enough, evidence for anaphora resolution can be detected (e.g., Cook, 2014; Klin et al., 2006). Specifically, note that both Levine et al. (2000) and Greene et al. (1992) employed the probe recognition task for investigating the encoding of the antecedents, which has been argued to be sensitive mainly to a surface encoding of words and not to the formation of abstract conceptual representations (Fincher-Kiefer, 1993; Gordon, Hendrick, & Foster, 2000). As such, if the antecedent had been partially encoded in Levine et al.'s (2000) and Greene et al.'s (1992) studies such that only an abstract level of representation associated with the antecedent was reinstated upon processing of the anaphor, the probe recognition task would not be sensitive to that effect.

To increase the probability of detecting processes related to the resolution of referring expressions (assuming such resolution is indeed taking place), Klin et al. (2006) employed a lexical decision task, which has been argued to be more sensitive to subtle and abstract representations (Fincher-Kiefer, 1993; Lucas, Tanenhaus, & Carlson, 1990). They also manipulated the presence of antecedents and distractors to examine processing of referring expressions. Their results showed a null effect of the presence of a distractor in the preceding text when a lexical decision task was used (Experiment 2). However, when they employed an even more sensitive measure (i.e., reading time) and manipulated the very presence of the antecedent by comparing conditions in which there was an antecedent in the preceding text with conditions in which there was no antecedent, they found clear evidence for partial encoding: Reading times for the line including the referring expression were faster in the antecedent-present condition than in the antecedent-absent condition, suggesting that at least a subset of information associated with the antecedent was retrieved upon processing of the pronoun (Experiment 3).

Thus, at best, referring expressions were only underspecified and not left completely unresolved in Levine et al.'s (2000) and Greene et al.'s (1992) studies. But why did these two studies obtain the results they did? That is, why did they find evidence for nonresolution? In addition to the use of a task thought to be an insensitive measure of sentence integration, some other factors might have contributed to their results. Specifically, Levine et al.'s (2000) results could have been obtained because, unlike pronouns that are void of semantic meaning and therefore only pose a referential problem, noun phrase referring expressions do contain semantic meaning of their own. Since previous research has shown that the computation of semantic meaning precedes referential processing (Nieuwland & van Berkum, 2008), it could have been the case that the high semantic overlap between the referring expression (*dessert*) and the antecedents (*tart* and *cake*) has created a temporary semantic coindexation based on a simple and heuristics-based semantic fit assessment between the anaphor and both of the antecedents, resulting in high interim equilibrium and shallow subsequent bottom-up/algorithmic processing. In fact, there is recent evidence showing that the initial stages of resolving noun phrase anaphors could be based on quick semantic matching (Cook, 2014; see above).

Another possible reason for the undespecification observed in Levine et al. (2000) could be that the antecedents were inanimate entities, and inanimate entities have been shown not to result in semantic confusability compared with animate entities. Specifically, Fukumura and van Gompel

(2011) showed that congruence in terms of animacy reduces pronoun use, suggesting that semantic confusability exists between the two animate entities. However, interestingly, the animacy congruence effect was not observed for inanimate–inanimate pairs. Fukumura and van Gompel (2011) argued that this could be because inanimate entities receive overall less activation in memory due to less enhanced processing (a levels of processing effect; Craik & Tulving, 1975), leading to less semantic confusability. Similarly, in Levine et al.'s (2000) study, because the two antecedents were both inanimate entities (*cake* and *tart*), the magnitude of semantic confusability produced by the two antecedents might have been so low that a nondiscriminating anaphor (*dessert*) could not possibly cause much disequilibrium, which could have in turn led to shallow processing. Interestingly, Klin et al. (2006) found evidence for reference resolution only when they compared reading times on referring expressions with and without an antecedent, where semantic congruence was obviously not an issue. This analysis also offers another explanation of the Greene et al. (1992) results suggesting that pronouns are not resolved during normal comprehension: It might be that their materials encouraged formation of merged representations for the two antecedents (Poesio et al., 2006). Specifically, recall that in their stimuli, the two human characters were conjoined in the first sentence and that a collective pronoun was used in the second sentence: See (5). As such, the two antecedents might have been merged together in single representations, and, as a result of this, unambiguous pronouns might have rendered pronoun resolution difficult, leading to nonresolution.

We also argue that the observation by Love and McKoon (2011) that adding text to the preceding discourse increased the likelihood of anaphoric resolution is compatible with the online equilibrium hypothesis (recall that Love & McKoon, 2011, used the same stimuli as those used in Greene et al., 1992, but lengthened the stories). This is because the additional text might have allowed the formation of a richer representation for the sentence containing the antecedents, which might in turn have increased the probability of reference resolution. More specifically, a richer preceding representation might have allowed for richer schemata, facilitating the integration of the referring expression with the representation of the preceding sentence and therefore achieving equilibrium. In fact, our own previous studies show that length plays a role in pronoun resolution by increasing the "semantic richness" of the associated antecedent. Specifically, we showed that longer antecedents are more likely to be later realized with pronouns rather than with repeated noun phrases (Karimi, Fukumura, Ferreira, & Pickering, 2014) and also that longer antecedents are more likely to interpreted as referents of ambiguous pronouns than shorter antecedents (Karimi & Ferreira, 2015). Similarly, a richer representation might facilitate retrieval of any antecedents contained in that representation for establishing reference and therefore achieving equilibrium. Moreover, such rich representation might have aided the separation of individual antecedents contained in any merged representation.

Another important psycholinguistic finding that also may seem to pose a serious challenge for the OCE approach to reference processing is the so-called "repeated-name penalty" (Gordon, Grosz, & Gilliom, 1993). The repeated-name penalty is the phenomenon wherein the subsequent reference to a subject antecedent produces processing difficulty if reference is made through a repeated noun as opposed to a pronoun. For example, given *Bruno was the bully of the neighbourhood. Bruno chased Tommy all the way home from school one day*, it takes people significantly longer to read *Bruno watched Tommy hide behind a big tree and start to cry* than to read *He watched Tommy hide behind a big tree and start to cry*. Based on the online equilibrium hypothesis, the most effective way of distinguishing between two preceding antecedents would be to use a repeated noun to eliminate ambiguity altogether. Thus, the fact that a repeated name causes difficulty rather than ease of processing runs counter to the predictions of the online equilibrium hypothesis.

Interestingly, however, there is evidence showing that the repeated name realization of a syntactic subject produces no processing difficulty

at the *initial* stages of processing. In one study, Almor and Nair (2007) found that the repeated name penalty was not observed when the text was presented word by word, but only when it was presented as larger chunks or whole sentences, indicating, according to the authors, that initial stages of referential processing (where the representation of the antecedents and that of the referring expression is merely activated, but no binding occurs) are not affected by the presence of a repeated name, but only later, integrative stages are (when the referring expression is integrated with the preceding discourse, and binding takes place). Almor and Eimas (2008) also reported similar findings in spoken language where they found the repeated name penalty effect only in a delayed recall task but not in lexical decision latencies. Similarly, employing a visual world paradigm, Almor and Phillips (2006) showed that semantic overlap between the referring expression and the antecedents facilitates initial stages of referential processing but also impedes later stages.

Under OCE, it could be argued that in the case of repeated-noun reference, the initial stage of establishing reference that is driven by heuristics resolves the ambiguity and brings about equilibrium, thereby producing no processing difficulty at the early stages of reference processing. Note that since the magnitude of the equilibrium in the interim output is high (because the referring expression is maximally unambiguous), the algorithmic processing should be heavily influenced by output of the heuristic route. However, there is strong bottom-up evidence from the linguistic input that opposes the output of the heuristic route. Specifically, and as argued by Almor (1999, 2004), there is no pragmatic justification for reactivating an antecedent that is already the most highly activated antecedent in discourse. In fact, overactivating a concept may cause disequilibrium because it might open up the possibility that there is a third, unmentioned antecedent that the referring expression is referring to (who happens to have the same name as the intended referent), reducing the discriminatory power of the referring expression. This argument is supported by the fact that the repeated name penalty only occurs for highly salient antecedents such as the syntactic subject and not for less prominent antecedents (e.g., Swaab, Camblin, & Gordon, 2004). As such, the algorithmic route might fail to establish reference for the syntactic subject when it is subsequently realized with a repeated noun, which might surface as difficulty in later stages of referential processing.

Online cognitive equilibrium and some established psycholinguistic findings

In this section, we focus on the compatibility of OCE with some important and established properties of the language processing system.

One of the most significant findings in the field of psycholinguistics in the past decade is that language processing is a predictive operation: Upcoming information is preactivated before they even appear in the unfolding linguistic input (Altmann & Kamide, 1999; Altmann & Mirkovic, 2009; Kamide, Altmann, & Haywood, 2003; Levy, 2008; van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005). Many models of language processing attribute prediction to the probabilistic nature of language processing (e.g., Jaeger, 2010; Levy, 2008). For example, under Bayesian approaches to language processing, the prior probability of a word (i.e., its base frequency) and the conditional probability of that word (i.e., the probability of that word occurring given a preceding string of words) conspire to lead to prediction of upcoming material (see, Courville, Daw, & Touretzky, 2006; McClelland, 1998; Qian, Jaeger, & Aslin, 2012).

Under OCE, prediction is the natural corollary of the interim output influencing ongoing algorithmic processing. As described above, the output of the heuristic route is produced earlier to arrive at equilibrium quickly. However, precisely because the heuristic route strives to achieve equilibrium as soon as possible, it may use whatever information is currently available and relate it to existing schemata to come up with a complete representation before the unfolding linguistic input is fully available. Once such an output is formed, it starts to project its contents on the algorithmic route, resulting in predictions about the identity of

upcoming words. In such a processing architecture, prediction is the result of activating an entire event representation rather than word-level cumulative priming (Otten & Van Berkum, 2009; cf. Kukona, Fang, Aicher, Chen, & Magnuson, 2011).

A comparison between OCE and the current prediction-based approaches merits some discussion at this point. A key claim of the current prediction-based approaches is that processing effort is the underlying force driving prediction, such that predictions are constantly made during language processing to reduce the processing difficulty of the upcoming material (e.g., Fine, Jaeger, Farmer, & Qian, 2013; Jaeger, 2010; Jaeger & Tily, 2011; Levy, 2008; see also Drieghe, Brysbaert, Desmet, & De Baecke, 2004; Frisson, Rayner, & Pickering, 2005; McDonald & Shillcock, 2003; Rayner, Li, Juhasz, & Yan, 2005; Staub & Clifton, 2006). However, these prediction-based approaches do not specify why the system should expend effort to predict in the first place. In other words, since prediction necessarily involves cognitive effort, why would the system not simply wait for upcoming information? In fact, being predictive runs counter to the fundamental least effort principle in cognitive processing (Jaeger, 2010; Piantadosi et al., 2012; Zipft, 1935, 1949). The system could save effort by not predicting in the first place and then use those saved resources when unexpected words appear. However, under OCE, prediction is the natural consequence of the formation of the interim output through top-down activation of prestored information and the projection of its contents onto the ongoing algorithmic processing, as outlined above. Prediction does not require resources under OCE; indeed, an interesting prediction of the OCE is that resources would be required to *inhibit* the generation of a prediction if the output of the heuristic route is sufficiently compelling.

Moreover, and with regard to underspecification in language processing, there seems to be no clear explanation in current prediction-based views for the shallow processing illustrated in phenomena such as the Moses illusion. One possibility, however, is that since predicted information is activated before it is explicitly encountered, little effort should be allocated to processing it, which may lead to underspecification. For instance, failure to detect the anomaly in the Moses illusion could occur because the preceding information has given rise to preactivation of enough semantic features of the agent so as to lead to diminished processing effort when the agent is actually encountered. However, it could be argued that since predicted information is more expected, the presence of other information in place of the predicted information should actually be more salient and therefore easier to detect. In fact, some previous research has shown that such is indeed the case (e.g., Luke & Christianson, 2012). Therefore, one might expect that comprehenders would be more likely to detect the wrong agent (i.e., Moses) in the Moses illusion if they predicted *Noah,* and they should therefore be jarred when they encounter *Moses* instead.

Another established psycholinguistic finding are syntactic garden-paths. In a sentence such as *While Mary bathed the baby played in the crib*, people have been shown to initially interpret *the baby* as the object of *bathe* (e.g., Christianson et al., 2001, 2006). It is only after encountering *played* that it becomes apparent that the initial parse is incorrect, and a revision has to be made. Under the garden-path model of sentence processing (Frazier & Fodor, 1978; Frazier & Rayner, 1982), misanalysis occurs because the language comprehension system obeys the late closure principle, according to which incoming information should be attached to the clause or phrase that is currently being processed, as long as the attachment is licensed by syntax. As such, *the baby* is initially attached to the clause in which *Mary* is the subject and *bathe* is the verb. Other models of language processing attribute the difficulty of processing garden path sentences to prediction error. Specifically, the argument under these models is that the syntactic structure of garden path sentences is infrequent, whereas simple subject–verb–object structure is highly frequent. Thus, the garden path structure is less predicted and less expected, which leads to

processing difficulty (Fine et al., 2013; Jaeger & Tily, 2011; Levy, 2008).

Under OCE, the garden path phenomenon could be explained by arguing that when the word immediately following the verb (*the baby* in the above example) is encountered, the system will reach equilibrium fastest if it interprets it as the object of the verb, rather than leave it unattached or interpret it as the subject of the following clause. This is because the object interpretation of *the baby* permits the formation of an interim output (including all three arguments—the agent, the action, and the patient), whereas a subject interpretation will postpone the formation of an interim output (see also Gorrell, 1987, for a similar argument). As such, these words are immediately interpreted as the object of the verb by the heuristic route, resulting in an erroneous parse. Obviously, later, when the interim output interacts with the ongoing algorithmic processing, the nonviability of the initial parse will become apparent.

The OCE approach towards the garden-path phenomenon can explain some important findings with regard to this phenomenon. One important observation has been that remaining committed to a certain analysis makes it more difficult to later disengage from that analysis. For example, F. Ferreira and Henderson (1991) lengthened the ambiguous region (underlined in the following example) in garden path sentences (*When the men hunt the birds typically scatter* vs. *When the men hunt the birds with bright plumage typically scatter*) and showed that the longer the ambiguous region, the harder it is for participants to recover from the incorrect analysis. In other words, the longer the processor remains committed to a certain syntactic analysis, the harder it becomes for the processor to disengage from that analysis in the event that it turns out to be wrong. Based on OCE, this length of commitment effect, as it were, could be explained by arguing that commitment results in equilibrium in the system (because an interim general representation has been formed; Principle 1), and the longer the system stays in equilibrium, the greater the equilibrium will be (Principle 2). Consequently, in the face of solid evidence from the algorithmic, bottom-up route disconfirming the interim representation, greater disequilibrium will ensue.

Another intriguing finding with regard to processing garden-path sentences is that the erroneous parse of the sentence and the associated representation are not completely overwritten by the revised analysis. Instead, such representations linger in memory. For example, recall that Christianson et al. (2001) presented participants with sentences such as *While Bill hunted the deer ran into the wood*, and asked comprehension questions such as *Did Bill hunt the deer?* Surprisingly, the found that the participants were significantly more likely to give incorrect "yes" responses to the garden-path than to non-garden-path sentences, indicating that the lingering of erroneous representations depends on the formation of an initial incorrect representation. According to OCE, it could be assumed that the lingering representations are in fact the interim heuristics-based output that is formed at the initial stages of sentence processing. Recall that OCE assumes that once equilibrium is achieved, the system prefers to stay in the concomitant equilibrium state as long as possible (Principle 2). We argue that this preference to stay in equilibrium is responsible for the lingering erroneous representations observed by Christianson et al. (2001). Obviously, once the heuristic route forms the erroneous representation and attains equilibrium, the interim output gets modified when the strong, bottom-up evidence is encountered. However, the intrinsic tendency in the system to stay in equilibrium might cause the system not to completely abandon the wrong, initial representation, producing the lingering erroneous representation effect.

This explanation of the results of Christianson et al. (2001) is in line with their other findings: That the tendency to give an incorrect "yes" response to the comprehension questions is contingent on the plausibility of the disambiguating word as well as on the length of the ambiguous region, such that plausible ambiguous words and longer ambiguous regions led to more incorrect "yes" responses than implausible ambiguous words and short ambiguous regions, respectively. We can

assume, based on OCE, that perhaps both plausibility and length confer a greater degree of equilibrium to the interim output and therefore to the system, causing the preference to stay in equilibrium and to retain the interim representation to grow even stronger.

The predictions of OCE

A strong test of the OCE involves the effects of individual differences. For example, with regard to the attachment ambiguity research by Traxler et al. (1998) and others (e.g., Swets et al., 2007; van Gompel, Pickering, Pearson, & Liversedge, 2005; van Gompel, Pickering, & Traxler, 2001), it has been shown that individuals with low working memory spans attach an ambiguous relative clause high (i.e., to the first-mentioned noun phrase) but individuals with high working memory spans attach it low (i.e., to the second-mentioned noun phrase).

One explanation of this tendency is that it is due to different chunking strategies on the part of people with different working memory capacities (see Swets et al., 2007). In addition, it is possible that low-span individuals need to arrive at equilibrium faster than high-span individuals. It is well established in the working memory literature that maintaining structured information in memory is easier than maintaining unstructured information (e.g., Bor, Duncan, Wiseman, & Owen, 2003; Imbo, Szmalec, & Vandierendonck, 2009), in line with theories of language processing that assume that partially processed but incomplete syntactic dependencies overtax working memory and thereby increase parsing complexity (Chomsky & Miller, 1963; Gibson, 1998, 2000; Pickering & Barry, 1991). If that is the case, attaching the ambiguous relative clause to one of the preceding noun phrases will provide structure (of a syntactic type) to otherwise unstructured information and increase equilibrium. Since people with shorter memory spans are assumed not to be able to keep much (unstructured) information in memory, the need to arrive at equilibrium might be more pressing for them, causing them to prefer to attach the relative clause faster than people with higher working memory spans. An early attachment then would increase the probability of high attachment because NP1 is presumably syntactically more prominent early on, and this prominence might fade with the passage of time. However, individuals with higher spans can hold more (unstructured) information in their memory and, thus, may not need to attach the relative clause immediately. This late attachment then could be the reason for their preference to attach low, presumably because the prominence of NP1 declines, and the recency of NP2 assumes more force.

Put differently, people with smaller working memory spans may try to come up with an interpretation as soon as the relative clause is encountered to reduce disequilibrium (uncertainly) and to be able to more efficiently keep the information under processing in memory. Since the syntactic prominence of NP1 over NP2 is greater early on during processing, the probability of attaching high increases greatly if an attachment is made early on. Conversely, high working memory span individuals can cope with the uncertainty/disequilibrium created by the relative clause for a longer time because they can keep unstructured information in memory more efficiently. Therefore, they do not immediately attach the relative clause to one of the NPs, resulting in more NP2 attachments. In fact, using online (and offline) measures, a recent study by Payne et al. (2014) showed that working memory span affects attachment preferences and that this effect is strongest for older adults, such that low-span older adults have the strongest tendency to attach high (NP1). More experiments that can record attachment preferences in real time would allow this prediction to be further tested.

If people with high working memory spans are more capable of tolerating disequilibrium or ambiguity, then there should be significant differences between high- and low-span individuals in how long they delay the resolution of ambiguous pronouns as well (MacDonald & MacWhinney, 1990; Stewart et al., 2007), with high-span individuals delaying resolution longer than low-span individuals. Again, testing the ambiguous pronoun resolution in real time could put this prediction to

test. Partial support for this prediction comes from ERP research carried out by Nieuwland and van Berkum (2006), showing that pronoun ambiguity results in processing difficulty for high- but not for low-span individuals. As the authors discuss, this pattern might occur because low-span individuals have not even noticed the ambiguity [they "immediately took on the—to them, and at that moment—most plausible referential interpretation" (Nieuwland & van Berkum, 2006, p. 163)]. In OCE terminology, this is to say that low-span individuals immediately interpret the ambiguous pronoun (because they need to arrive at equilibrium faster), failing to entertain other potential referents for the ambiguous pronoun. Related to these findings, Lee and Federmeier (2012) found that older adults (with lower working memory spans) have a harder time resolving lexical ambiguity because they cannot suppress the irrelevant meaning of homographs and, more importantly, that this negatively influences subsequent processing, suggesting that when equilibrium cannot be immediately achieved, successful comprehension is more difficult to attain for low-span individuals.

Also, if low-span individuals cannot keep unstructured information in memory for long, they should predict faster than high-span individuals because the need to arrive at equilibrium is more pressing for them. Otten and van Berkum (2009) investigated this possibility but found no difference between high- and low-span readers in predicting words. However, their conclusion was based on null results, and we think this possibility might be checked again with different materials and more sensitive methodology to pursue the question further and more efficiently. Despite finding no difference between high- and low-span individuals, however, Otten and van Berkum (2009) observed a late negativity when low-span readers encountered determiners that were not consistent with the gender of the predicted words. It is possible that individuals with smaller memory spans might have predicted earlier and consequently might have remained committed to the output of the heuristic route for a longer time, or they might have relied on the interim output more heavily. Thus, the observed negativity might reflect the greater disturbance of equilibrium in the face of conflicting bottom-up evidence.

Another related prediction of OCE has to do with lingering garden-path representations and individual differences. Specifically, and as mentioned above, it could be the case that low-span individuals need to achieve equilibrium faster than high-span individuals because they are less capable of retaining unstructured information in memory, and therefore the need for formation of an interim output is greater for them. On the other hand, we know from the results of F. Ferreira and Henderson (1991) and Christianson et al. (2001) that committing to certain syntactic analysis results in stronger garden-path effects. As such, low-span individuals might stay with the erroneous garden-path representation for a longer time (because they come up with it earlier) and might rely on the interim representation more strongly. Consequently, garden-path representations should linger more for low-span than for high-span individuals. This prediction stands in contrast to the prediction stemming from some previous research that has demonstrated that all interpretations of an ambiguous structure (or word) are initially activated and that high-span individuals maintain these representations in an active state for a longer time than low-span individuals (see, Farmer, Cargill, Hindy, Dale, & Spivey, 2007; MacDonald, Just, & Carpenter, 1992). Partial support for this prediction comes from Chirstianson et al. (2006) who showed that older adults are more accepting of the incorrect interpretations of garden-path sentences, presumably due to age-related deficits in working memory span. In any case, further investigation of the potential interaction between working memory capacity and lingering representations is another way of putting OCE to test.

Moving away from variations in individual differences, another way to test OCE is through sentence complexity. If a sentence is hard to process, there should be more reliance on the heuristic route (at least early on during processing) to create an interim output and attain equilibrium earlier than when the sentence is easier to process. This is because in the case of a simple sentence, heuristic and algorithmic processing might not

greatly differ in how early they reach completion, but for complex sentences, this time difference is greater (see F. Ferreira, 2003; F. Ferreira et al, 2002). This prediction is partially supported by research showing that more complex anaphoric relations cause early resolution. For example, Sorace and Filiaci (2006) presented participants with the Italian equivalent of sentences such as (7) and asked them to perform a picture verification task in which they had to choose pictures corresponding to the subordinate clause and thus identify the referent of null and/or overt subject pronouns.

7. a. The mother kisses her daughter, while she is wearing her coat. (forward anaphora)
   b. While she is wearing her coat, the mother kisses her daughter. (backward anaphora)

They observed that the tendency to choose the subject of the main clause (i.e., mother) as the referent of the subject pronoun in the subordinate clause (i.e., she) was stronger in the backward anaphora condition (7b), where the processing demands were presumably higher than in the forward anaphora condition (7a). From an OCE standpoint, these results might have occurred because in backward anaphora there is a greater need for earlier resolution (equilibrium); a dangling referring expression produces more ambiguity (disequilibrium) than a referring expression with a clear antecedent. In other words, the first clause is more ambiguous in the case of backward anaphora (7b) than in the case of forward anaphora (7a). Thus, referential ambiguity is resolved as soon as the first candidate (i.e., *the mother*) is encountered.

In a recent study of our own (Karimi & Ferreira, 2015), we presented participants with sentences containing two antecedents. We made either or none of the antecedents long by attaching a relative clause to it. The results showed that when the given sentence was complex (i.e., when one of the antecedents was made long by a relative clause), a following ambiguous pronoun was resolved more quickly than in the baseline condition where none of the antecedents was long. Future research can investigate this further. For example, it would be interesting to examine whether low-span individuals are more likely to make an initial commitment to one of the antecedents in complex sentences than high-span individuals.

A final prediction of OCE is related to the formation of merged representations: If merged representations are indeed created, and if separating the individual concepts contained in the merged representation disturbs equilibrium, then encouraging the formation of merged representations should produce processing difficulty when unambiguous pronouns are used and facilitated processing when ambiguous pronouns are employed. To test this prediction, stimuli such as (8) could be made in which pronoun ambiguity as well as motivation for forming merged representations are manipulated.

8. a. Mary went shopping with John. They bought a lot of grocery. She/He was very excited.
   b. Mary went shopping with John. The store was huge. She/He was very excited.

As is clear, the formation of a merged representation of the antecedents is encouraged by the use of a collective pronoun in the second sentence in (8a). As such, the prediction from the OCE is that the underlined pronoun in (8a) should take longer to read than that in (8b), because it may force separating the two individual antecedents in the merged representation (if any).

Conclusions

All in all, the online equilibrium hypothesis provides a new framework for viewing language processing in general and reference processing in particular. This hypothesis provides an underlying mechanism for the formation of good-enough linguistic representations. We proposed a language processing model that brings together the concepts of online cognitive equilibrium as well as heuristic and algorithmic processing and proposes a relative timing for their operation. Findings from both intra- and intersentential processing seem to be compatible with the online equilibrium approach to language processing and the associated model. Specifically, intrasentential processing seems to

result from the low level of disequilibrium caused by certain words or structures that may confer high degrees of equilibrium and therefore trigger superficial bottom-up processing. Similarly, intersentential (discourse) processing also seems to be consistent with the attempt to maintain cognitive equilibrium in the sense that processing difficulty for a referring expression seems to be sensitive to the magnitude of cognitive disequilibrium incurred by that referring expression. Underspecified reference processing appears to be the result of the system arriving at equilibrium as early as possible and/or not allocating substantial resources to reference processing once equilibrium has been achieved.

Obviously, the online equilibrium hypothesis is not meant to explain all the complex and sometimes opposing results that have been found in the psycholinguistics literature. However, we hope that it provides a new perspective for building and testing theories of language processing.

REFERENCES

Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review*, *106*(4), 748–765.

Almor, A. (2004). A computational investigation of reference in production and comprehension. In J. Trueswell & M. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language -as-action traditions* (pp. 285–301). Cambridge, MA: MIT Press.

Almor, A., & Eimas, P. D. (2008). Focus and noun phrase anaphors in spoken language comprehension. *Language and Cognitive Processes*, *23*(2), 201–225.

Almor, A., & Nair, V. A. (2007). The form of referential expressions in discourse. *Language and Linguistics Compass*, *1*(1–2), 84–99.

Almor, A., & Phillips, M. (2006, November). *Category NP anaphors in spoken language comprehension*. Paper presented at the 47th Annual Meeting of the Psychonomics Society, Houston, TX.

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

Altmann, G. T. M., & Mirkovic, H. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science*, *33*, 583–609.

Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. *Cognition*, *76*, B13–B26.

Arnold, J. E., & Griffin, Z. (2007). The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language*, *56*, 521–536.

Barton, S. B., & Sanford, A. J. (1993). A case study of anomaly detection: Shallow semantic processing and cohesion establishment. *Memory & Cognition*, *21*, 477–87.

van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language*, *41*, 147–182.

van Berkum, J. J. A., Brown, C. M., Hagoort, P., & Zwitserlood, P. (2003). Event-related brain potentials reflect discourse-referential ambiguity in spoken language comprehension. *Psychophysiology*, *40*(2), 235–248.

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., Zwitserlood, P., Bastiaansen, M. C. M., Brown, C. M., & Hagoort, P. (2004). *So who's 'he' anyway? Differential ERP and ERSP effects of referential success, ambiguity and failure during spoken language comprehension*. Annual meeting of the Cognitive Neuroscience Society (CNS-2004), San Francisco, April 18–20.

Boland, J. E., & Tanenhaus, M. K. (1991). The role of lexical representations in sentence processing. *Advances in Psychology*, *77*, 331–366.

Bor, D., Duncan, J., Wiseman, R. J., & Owen, A. M. (2003). Encoding strategies dissociate prefrontal activity from working memory demand. *Neuron*, *37* (2), 361–367.

Boudewyn, M. A., Long, D. L., & Swaab, T. Y. (2013). Effects of working memory span on processing of lexical associations and congruence in spoken discourse. *Frontiers in Psychology*, *4*. doi:10.3389/fpsyg.2013.00060

Bredart, S., & Modolo, K. (1988). Moses strikes again: Focalization effect on a semantic illusion. *Acta Psychologica*, *67*(2), 135–144.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, *22*, 1482–1493.

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.

Brown-Schmidt, S., & Tanenhaus, M. K. (2006). Watching the eyes when talking about size: An investigation of message formulation and utterance planning. *Journal of Memory and Language*, *54*, 592–609.

Chambers, C. G., Tanenhaus, M. K., Eberhard, K. M., Filip, H., & Carlson, G. N. (2002). Circumscribing referential domains during real time language comprehension. *Journal of Memory and Language*, *47*, 30–49.

Chomsky, N., & Miller, G. A. (1963). Introduction to the formal analysis of natural languages. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 2, pp. 269–321). New York: Wiley.

Christianson, K., & Ferreira, F. (2005). Conceptual accessibility and sentence production in a free word order language (Odawa). *Cognition*, *98*(2), 105–135.

Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology*, *42*, 368–407.

Christianson, K., Luke, S. G., & Ferreira, F. (2010). Effects of plausibility on structural priming. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *36*, 538–544.

Christianson, K., Williams, C. C., Zacks, R. T., & Ferreira, F. (2006). Younger and older adults' "good enough" interpretations of garden path sentences. *Discourse Processes*, *42*, 205–238.

Cook, A. E. (2014). Processing anomalous anaphors. *Memory & Cognition*, *42*(7), 1171–1185.

Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world. *Trends in Cognitive Sciences*, *10*(7), 294–300.

Craik, E. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*, 268–294.

Drieghe, D., Brysbaert, M., Desmet, T., & De Baecke, C. (2004). Word skipping in reading: On the interplay of linguistic and visual factors. *European Journal of Cognitive Psychology*, *16*(1–2), 79–103.

Dwivedi, V. D. (2013). Interpreting quantifier scope ambiguity: Evidence of heuristic first, algorithmic second processing. *PLoS ONE*, *8*(11): e81461.

Erickson, T. A., & Matteson, M. E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learn and Verbal Behavior*, *20*, 540–551.

Farmer, T. A., Cargill, S. A., Hindy, N. C., Dale, R., & Spivey, M. J. (2007). Tracking the continuity of language comprehension: Computer mouse trajectories suggest parallel syntactic processing. *Cognitive Science: A Multidisciplinary Journal*, *31*, 889–909.

Ferreira, F. (1994). Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language*, *33*, 715–736.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*, 164–203.

Ferreira, F., Christianson, K., & Hollingworth, A. (2001). Misinterpretations of garden-path sentences: Implications for models of reanalysis. *Journal of Psycholinguistic Research*, *30*, 3–20.

Ferreira, F., & Clifton, C., Jr. (1986). The independence of syntactic processing. *Journal of Memory and Language*, *25*, 348–368.

Ferreira, F., Ferraro, V., & Bailey, K. G. D. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*, 11–15.

Ferreira, F., Foucart, A., & Engelhardt, P. E. (2013). Language processing in the visual world: Effects of preview, visual complexity, and prediction. *Journal of Memory and Language*, *69*, 165–182.

Ferreira, F., & Henderson, J. M. (1991). Recovery from misanalyses of garden path sentences. *Journal of Memory and Language*, *30*, 725–745.

Ferreira, F., & Patson, N. (2007). The good enough approach to language comprehension. *Language and Linguistics Compass*, *1*, 71–83.

Ferreira, V. S., Slevc, L. R., & Rogers, E. S. (2005). How do speakers avoid ambiguous linguistic expressions? *Cognition*, *96*, 263–284.

Fincher-Kiefer, R. (1993). The role of predictive inferences in situation model construction. *Discourse Processes*, *16*, 99–124.

Fine, A. B., & Jaeger, T. F. (2013). Evidence for implicit learning in syntactic comprehension. *Cognitive Science*, *37*(3), 578–91.

Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS ONE*, *8*(10), e77661.

Fodor, J. D., & Inoue, A. (1998). Attach anyway. In J. D. Fodor & F. Ferreira (Eds.), *Reanalysis in sentence processing* (pp. 101–141). Dordrecht: Kluwer.

Fodor, J. D., & Inoue, A. (2000). Garden path reanalysis: Attach (anyway) and revision as last resort? In M. de Vincenzi & V. Lombardo (Eds.), *Cross-linguistic perspective on language processing* (pp. 21–61). Dordrecht: Kluwer.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, *6*, 291–325.

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.

Frazier, L., & Rayner, K. (1990). Taking on semantic commitments: Processing multiple meanings vs. multiple senses. *Journal of Memory and Language*, *29*, 181–200.

Frisson, S., Rayner, K., & Pickering, M. J. (2005). Effects of contextual predictability and transitional probability on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(5), 862–877.

Fukumura, K., & van Gompel, R. P. (2011). The effect of animacy on the choice of referring expression. *Language and Cognitive Processes*, *26*(10), 1472–1504.

Fukumura, K., & van Gompel, R. P. G. (2012). Producing pronouns and definite noun phrases: Do speakers use the addressee's discourse model? *Cognitive Science*, *36*, 1289–1311.

Fukumura, K., van Gompel, R. P. G., & Pickering, M. J. (2010). The use of visual context during the production of referring expressions. *Quarterly Journal of Experimental Psychology*, *63*, 1700–1715.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*(1), 1–76.

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In Y. Miyashita, A. Marantz, & W. O'Neil (Eds.), *Image, language, brain* (pp. 95–126). Cambridge, MA: MIT Press.

van Gompel, R. P. G., Pickering, M. J., Pearson, J., & Liversedge, S. P. (2005). Evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language*, *52*, 284–307.

van Gompel, R. P. G., Pickering, M. J., & Traxler, M. J. (2001). Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. *Journal of Memory and Language*, *45*, 225–258.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, *17*, 311–347.

Gordon, P. C., Hendrick, R., & Foster, K. (2000). Language comprehension and probe-list memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 766–775.

Gorrell, P. G. (1987). *Studies of human syntactic processing: Ranked-parallel versus serial models* (Unpublished doctoral dissertation). University of Connecticut, Stores.

Greene, S. B., McKoon, G., & Ratcliff, R. (1992). Pronoun resolution and discourse models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 266–283.

Hagoort, P., Baggio, G., & Willems, R. M. (2009). Semantic unification. In *The cognitive neurosciences* (4th ed., pp. 819–836). Cambridge, MA: MIT Press.

van Herten, M., Chwilla, D. J., & Kolk, H. H. (2006). When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception. *Journal of Cognitive Neuroscience*, *18*(7), 1181–1197.

van Herten, M., Kolk, H. H., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research*, *22*(2), 241–255.

Horton, W. S., & Gerrig, R. J. (2002). Speakers' experiences and audience design: Knowing when and knowing how to adjust utterances to addressees. *Journal of Memory and Language*, *47*(4), 589–606.

Horton, W. S., & Gerrig, R. J. (2005). Conversational common ground and memory processes in language production. *Discourse Processes*, *40*(1), 1–35.

Imbo, I., Szmalec, A., & Vandierendonck, A. (2009). The role of structure in age-related increases in visuo-spatial working memory span. *Psychologica Belgica*, *49*(4), 275–291.

Jackendoff, R. (2007). A parallel architecture perspective on language processing. *Brain research*, *1146*, 2–22.

Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology*, *61*, 23–62.

Jaeger, T. F., & Snider, N. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, *127*(1), 57–83.

Jaeger, T. F., & Tily, H. (2011). On language 'utility': Processing complexity and communicative efficiency. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2* (3), 323–335.

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*, 133–156.

Karimi, H. & Ferreira, F. (2015). *Informativity renders a referent more accessible: Evidence from eye-tracking*. Manuscript submitted for publication.

Karimi, H., Fukumura, K., Ferreira, F., & Pickering, M. J. (2014). The effect of noun phrase length on the form of referring expressions. *Memory & Cognition*, *42*, 993–1009.

Klin, C. M., Guzman, A. E., Weingartner, K. M., & Ralano, A. S. (2006). When anaphor resolution fails: Partial encoding of anaphoric inferences. *Journal of Memory and Language*, *54*, 131–143.

Klin, C. M., Weingartner, K. M., Guzman, A. E., & Levine, W. H. (2004). Readers_sensitivity to linguistic cues in narratives: How salience influences anaphor resolution. *Memory & Cognition*, *32*, 511–522.

Kukona, A., Fang, S. Y., Aicher, K. A., Chen, H., & Magnuson, J. S. (2011). The time course of anticipatory constraint integration. *Cognition*, *119*(1), 23–42.

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

Kurtzman, H. S., & MacDonald, M. C. (1993). Resolution of quantifier scope ambiguities. *Cognition*, *48*, 243–279.

Lee, C. L., & Federmeier, K. D. (2012). Ambiguity's aftermath: How age differences in resolving lexical ambiguity affect subsequent comprehension. *Neuropsychologia*, *50*(5), 869–879.

Levine, W. H., Guzman, A. E., & Klin, C. M. (2000). When anaphor resolution fails. *Journal of Memory and Language*, *43*, 594–617.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126–1177.

Lim, J. H., & Christianson, K. (2013a). Second language sentence processing in reading for comprehension and translation. *Bilingualism: Language and Cognition*, *16*(3), 518–537.

Lim, J. H., & Christianson, K. (2013b). Integrating meaning and structure in L1–L2 and L2–L1 translations. *Second Language Research*, *29*(3), 233–256.

Love, J., & McKoon, G. (2011). Rules of engagement: Incomplete and complete pronoun resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(4), 874–887.

Lucas, M. M., Tanenhaus, M. K., & Carlson, G. N. (1990). Levels of representation in the interpretation of anaphoric reference and instrument inference. *Memory & Cognition*, *18*, 611–631.

Luke, S. G., & Christianson, K. (2012). Semantic predictability eliminates the transposed-letter effect. *Memory & Cognition*, *40*(4), 628–641.

MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, *24* (1), 56–98.

MacDonald, M. C., & MacWhinney, B. (1990). Measuring inhibition and facilitation in pronouns. *Journal of Memory and Language*, *29*, 469–92.

McClelland, J. L. (1998). Connectionist models and Bayesian inference. In M. Oaksford & N. Chater (Eds.), *Rational models of cognition* (pp. 21–52). New York: Oxford University Press.

McDonald, S. A., & Shillcock, R. C. (2003). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science*, *14*(6), 648–652.

Nieuwland, M. S., & Van Berkum, J. J. A. (2006). Individual differences and contextual bias in pronoun resolution: Evidence from ERPs. *Brain Research*, *1118*, 155–167.

Nieuwland, M. S., & Van Berkum, J. J. A. (2008). The interplay between semantic and referential aspects of anaphoric noun phrase resolution: Evidence from ERPs. *Brain and Language*, *106*, 119–131.

Osgood, C. E., & Bock, J. K. (1977). Salience and sentencing: Some production principles. In S. Rosenberg (Ed.), *Sentence production: Developments in research and theory* (pp. 89–140). Hillsdale, NJ: Lawrence Erlbaum.

Otten, M., & Van Berkum, J. J. (2009). Does working memory capacity affect the ability to predict upcoming words in discourse? *Brain Research*, *1291*, 92–101.

Patson, N. D., Swensen, E., Moon, N., & Ferreira, F. (2006). *Individual differences in syntactic reanalysis*. Poster presented at the Annual CUNY Conference on Human Sentence Processing, March 2006, New York, NY.

Payne, B. R., Grison, S., Gao, X., Christianson, K., Morrow, D. G., & Stine-Morrow, E. A. (2014). Aging and individual differences in binding during sentence understanding: Evidence from temporary and global syntactic attachment ambiguities. *Cognition*, *130*(2), 157–173.

Piaget, J. (1952). *The origins of intelligence*. New York: International University Press.

Piaget, J. (1977). *The development of thought: Equilibration of cognitive structures*. New York: Viking.

Piaget, J. (1985). *The equilibrium of cognitive structures*. Chicago: University of Chicago Press.

Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*, *108* (9), 3526–3529.

Piantadosi, S. T., Tily, H., & Gibson, E. (2012). The communicative function of ambiguity in language. *Cognition*, *122*, 280–291.

Pickering, M. J., & Barry, G. (1991). Sentence processing without empty categories. *Language and Cognitive Processes*, *6*, 229–259.

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, *134*, 427–459.

Pickering, M. J., & Frisson, S. (2001). Processing ambiguous verbs: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(2), 556–573.

Poesio, M., Sturt, P., Artstein, R., & Filik, R. (2006). Underspecification and anaphora: Theoretical issues and preliminary evidence. *Discourse Processes*, *42*, 157–175.

Qian, T., & Jaeger, T. F. (2012). Cue effectiveness in communicatively efficient discourse production. *Cognitive Science*, *36*(7), 1312–1336.

Qian, T., Jaeger, T. F., & Aslin, R. (2012). Learning to represent a multicontext environment: More than detecting changes. *Frontiers in Psychology*, *3*, 1–9.

Rayner, K., Li, X., Juhasz, B. J., & Yan, G. (2005). The effect of word predictability on the eye movements of Chinese readers. *Psychonomic Bulletin & Review*, *12* (6), 1089–1093.

Reinhart, T. (1982). Pragmatics and linguistics: An analysis of sentence topics. *Philosophica*, *27*, 53–94.

Rigalleau, F., & Caplan, D. (2000). Effects of gender marking in pronominal coindexation. *Quarterly Journal of Experimental Psychology*, *53A*, 23–52.

Rigalleau, F., Caplan, D., & Baudiffier, V. (2004). New arguments in favour of an automatic gender pronominal process. *Quarterly Journal of Experimental Psychology*, *57A*, 893–933.

Sanford, A. J., & Filik, R. (2006). *They* as a gender-unspecified singular pronoun: Eye tracking reveals processing cost. *The Quarterly Journal of Experimental Psychology*, *60*(2), 171–178.

Sanford, A. J., Sanford, A., Molle, J., & Emmott, C. (2006). Shallow processing and attention capture in written and spoken discourse. *Discourse Processes*, *42* (2), 109–130.

Sanford, A. J., & Sturt, P. (2002). Depth of processing in language comprehension: Not noticing the evidence. *Trends in Cognitive Sciences*, *6*, 382–386.

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, *71*, 109–148.

Smith, N., & Levy, R. (2008). Optimal processing times in reading: A formal model and empirical investigation. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *The 30th annual meeting of the Cognitive Science Society (CogSci08)* (pp. 595–600). Austin, TX: Cognitive Science Society.

Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research*, *22*, 339–368.

Spivey, M. J., Tanenhaus, M. K., Eberhard, K. M., & Sedivy, J. C. (2002). Eye-movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology*, *45*, 447–481.

Staub, A., & Clifton, C., Jr. (2006). Syntactic prediction in language comprehension: Evidence from either … or. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(2), 425–436.

Stewart, A. J., Holler, J., & Kidd, E. (2007). Shallow processing of ambiguous pronouns: Evidence for delay. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *60*, 1680–1696.

Swaab, T. Y., Camblin, C. C., & Gordon, P. C. (2004). Electrophysiological evidence for reversed lexical repetition effects in language processing. *Journal of Cognitive Neuroscience*, *16*(5), 715–726.

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

Swets, B., Desmet, T., Hambrick, D. Z., & Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: A psychometric approach. *Journal of Experimental Psychology: General*, *136*(1), 64–81.

Tomlin, R. S. (1995). Focal attention, voice, and word order: An experimental, cross-linguistic study. In P. Downing & M. Noonan (Eds.), *Word order in discourse* (pp. 517–554). Amsterdam: John Benjamins Publishing.

Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules* (Vol. *1950*). Cambridge, MA: MIT Press.

Traxler, M. J., Pickering, M. J., & Clifton, C. (1998). Adjunct attachment is not form of lexical ambiguity resolution. *Journal of Memory and Language*, *39*, 558–592.

Trueswell, J. C., & Tanenhaus, M. K. (1994). Toward a lexicalist framework of syntactic ambiguity resolution. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 155–180). Hillsdale, NJ: Erlbaum.

Zipf, G. K. (1935). *The psychobiology of language*. London: Routledge.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. New York: Addison-Wesley.