Sentence Comprehension, Linguistic Complexity in

Introductory article

Edward Gibson, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

CONTENTS

The process of comprehending a sentence involves the moment-by-moment evaluation of a variety of informational and computational constraints. The combination of these constraints makes a sentence easy or difficult to process.

INTRODUCTION

Most sentences that we encounter in everyday life are easy to understand. But occasionally a sentence can be confusing or complex. For example, consider sentence (1):

The dog walked to the park had been chewing the bone. (1)

While reading (1), people generally get confused when the words 'had been chewing …' are encountered. People experience difficulty like this (a so-called 'garden-path effect') when there is an ambiguity earlier in the sentence for which they follow a likely interpretation, but which turns out to be incorrect. The temporary ambiguity in (1) is initiated at the word 'walked', which is ambiguous between a past tense verb and a passive participle verb. The past tense reading – the reading that is initially preferred – results in a main clause interpretation, in which 'the dog' is the actor (i.e. the agent carrying out the action) in a past-tense walking event. This interpretation is no longer viable when the words 'had been …' are input. The passive participle reading – the reading that is not usually noticed initially – involves 'the dog' as a patient (i.e. to whom the action is done) of a walking event, in which an unmentioned individual walked the dog.

Now consider sentence (2), which presents its readers with a second kind of processing difficulty:

The reporter that the senator that John met at the party attacked admitted the error. (2)

Sentence (2) consists of a main clause 'the reporter admitted the error', whose subject is modified by a relative clause (RC) 'that the senator … attacked'. The subject of the RC is then itself modified by another RC, 'that John met at the party'. This sentence is extremely difficult to understand, independent of any temporary ambiguities that it contains. In particular, sentence (2) is difficult to understand even when reading it for the second or third time. This is not so for (1): once it is known that the passive participle interpretation is the target structure, (1) becomes understandable.

Complicated sentences like (1) and (2) can be highly informative for telling us how the process of sentence comprehension occurs. In ambiguous sentences like (1), the combination of a variety of factors results in preferences for some structures over others, because not all structures can be pursued in parallel. In sentences like (2), without confusing temporary ambiguities, the combination of the same factors results in easier or harder processing, depending on the difficulty of the target interpretation. The analysis of a wide range of sentence comprehension studies involving many different structures across languages suggests that a variety of information sources are used in constructing an interpretation for a sentence, and that the process of building the target representation is constrained by the available computational resources. Some of the informational and resource factors are discussed below.

MEASURING LINGUISTIC COMPLEXITY EXPERIMENTALLY

Studies of sentence comprehension involve comparisons between target sentences and appropriate control sentences. For example, a control for sentence (1) above is a disambiguated version, as in (3):

The dog that was walked to the park had been chewing the bone. (3)

This sentence has the same structure as (1), but is disambiguated by the words 'that was' towards the passive participle interpretation of 'walked'. This sentence is correspondingly easier to understand than (1). A control sentence structure for sentence (2) is given in (4):

At the party, John met the senator that attacked the reporter that admitted the error. (4)

Sentence (4) contains all the same words as (2), with the same thematic relations among them, but (4) is much easier to understand. Thus, it is something about the structure of (2) that makes it difficult to understand.

A single individual's reactions to sentences like (1) and (2) on the one hand and their controls in (3) and (4) are not very informative by themselves. Because of individual variation, it is necessary to test hypotheses regarding sentence comprehension on a range of experimental participants and items. Perhaps the simplest method for gathering data on sentence comprehension is by means of acceptability (or grammaticality) judgments. This method consists of having experimental participants answer a questionnaire in which sentences are rated for their understandability, according to the participants' intuitions. A more objective method involves measuring reaction times to full sentences (presented either visually or auditorily) and accuracy to questions about the content of the presented sentences.

Often we want to know about the time course of processing load in a sentence: where in the sentence does the difficulty begin and end? End-of-sentence (offline) measures do not address such questions. Online measurements are necessary. In reading, measuring participants' eye movements to visually presented sentences is one online technique. At points of high complexity, participants slow down and/or regress to previous regions. A commonly used online alternative to tracking eye movements is self-paced reading with a moving window display. In this method, a sentence is initially presented on a computer screen as a series of dashes marking the length and position of the words in the sentence. Participants press a key (usually the spacebar) to reveal each region of the sentence, often one word at a time. As each new region appears, the preceding region disappears. The amount of time the participant spends reading each region is recorded as the time between key-presses. As in eye-tracking, participants tend to slow down at points of high complexity. Unlike eye-tracking, however, there is no way in self-paced reading for a participant to back up and re-read a region that was confusing.

It is more difficult in auditory language presentation to obtain an online dependent measure such as reading time. One way to obtain an online measure of difficulty in auditory presentation of language is via cross-modal priming. In a cross-modal priming paradigm, participants listen to sentences over headphones, and then perform a different task – such as deciding whether a string of letters is a word – at predetermined locations in the target sentences. People are faster at the interrupting task at points of lower complexity, such as when a word in the linguistic context is semantically related to the target stimulus (e.g. 'doctor' is related to 'nurse'). A second online auditory method involves tracking participants' eye movements with respect to a visually presented scene while the participants listen to instructions about the scene. This method is particularly informative with respect to the question of how the context of the presented scene influences sentence understanding. A third method involves monitoring the scalp with electrodes in order to measure event-related potentials (ERPs): minute voltage changes due to differences in neural activity while participants listen to sentences. This method, which is also used in reading research, requires an understanding of what the voltage changes mean. Some plausible interpretations are currently being proposed in this area.

RESOURCE CONSTRAINTS THAT AFFECT PROCESSING LOAD

In the process of understanding a sentence, it is necessary to integrate structures for incoming words into the structure(s) that have been built thus far, such that the potential integrations for an incoming word are determined by the syntactic rules for the language. According to one current theory – the dependency locality theory (DLT) – the processing cost of integrating a new word w is proportional to the distance between w and the syntactic item with which w is being integrated. Structural integration cost has been shown to be an important factor in accounting for online processing load. For example, consider the RC structures in (5) and (6):

The reporter that attacked the senator admitted the error. (5)

The reporter that the senator attacked admitted the error. (6)

In (5), the RC pronoun 'that' is interpreted as the subject of the verb 'attacked', whereas in (6), the same pronoun is interpreted as the object of the verb 'attacked'. People read the verb 'attacked' more slowly in a sentence like (6) than in a sentence like (5). This difference can be explained by integration distances. In (5), there is one local integration when processing 'attacked': this verb is integrated with the preceding RC pronoun as its subject. In contrast, there are two integrations at the point of processing 'attacked' in (6): this verb must be integrated as the verb for the subject 'the senator' (a local integration) and the object position of 'attacked' must be integrated with the RC pronoun 'that', a nonlocal integration. As a result of the extra nonlocal integration, the processing load at 'attacked' is larger in (6) than in (5), resulting in longer reading times at this word. Furthermore, reading times are slow in both sentence types for the verb 'admitted', a point of long-distance integration with the subject 'the reporter' in both sentence types. Reading times are relatively faster for the other words in the sentences, because integrations at all other positions are local.

Interestingly, aphasic (speech loss) stroke patients understand the subject-RCs in sentences like (5), but they do not understand the object-RCs in sentences like (6), as evidenced by their inability to reliably answer questions about the object-RCs. This suggests that the brain damage suffered by these aphasic patients has reduced the available resources for processing sentences.

An interesting question raised by a distance-based theory of integration cost is how distance is quantified. It appears that complexity may depend not only on the number of words or syllables between two integration points but also on the complexity of the intervening discourse structures. In particular, the ease or difficulty of constructing and/or accessing the referents in the intervening material affects the complexity of integrations across these referents. For example, people read the verb 'attacked' in the object-RC in (6) more quickly when the subject of the RC 'the senator' is replaced with a pronoun such as 'I' or 'you'. This decrease in complexity is arguably due to the fact that the pronouns 'I' or 'you' indicate highly accessible referents in the discourse – the speaker/writer and the hearer/reader – whereas 'the senator' refers to an individual who is not part of the current discourse, in a single sentence paradigm. Reading times for the embedded verb 'attacked' also decrease substantially when there is a referent for the subject noun-phrase (NP) 'the senator' in the current context. This result provides more support for the discourse-based integration cost metric, assuming that it is easier to access a structure for a referent that has just been built than it is to build a structure for a new referent.

Integration distances provide a partial explanation for the extreme complexity of (2). The integrations at the words 'attacked' and 'admitted' are all substantially longer in (2) than in (6), with the inclusion of the RC 'that John met at the party' modifying the NP 'the senator'. Such long-distance integrations, in combination with storage costs at these processing states, may be too complex for the limited capacity of sentence-processing resources. In contrast, all integrations are local in the control for (2) in (4).

The processing of ambiguous structures provides further evidence that longer-distance integrations are more costly. Consider (7):

The bartender told the detective that the suspect left the country yesterday. (7)

The adverbial 'yesterday' can be associated with either the most local verb 'left' or the earlier verb 'told'. The more local integration is strongly preferred. This preference is consistent with the hypothesis that people attempt to minimize integration costs when faced with ambiguity.

At the same time that the process of integration is going on in understanding a sentence, it is also necessary to store the partially processed structures. Both structural integration and storage consume computational resources. According to the DLT, the resources required for storing a partially processed structure are proportional to the number of incomplete syntactic dependencies at that point in processing the structure. As a result, processing load increases when the number of incomplete syntactic dependencies increases. Results from processing ambiguous structures further support the hypothesis that the sentence comprehension mechanism is sensitive to the number of incomplete dependencies in a structure: at choice points, structures with fewer incomplete dependencies are preferred over structures with more incomplete dependencies.

It should be noted that the DLT is just one theory of computational resource use in sentence comprehension. According to an earlier theory, ambiguity resolution is guided by two principles: a locality principle (like the integration component of the DLT), and a principle known as Minimal

Attachment, which prefers phrase structures involving fewer phrase structure rule applications. This theory and the DLT make largely the same predictions with respect to processing ambiguous inputs. It remains an open question precisely how resources constrain sentence interpretation.

INFORMATIONAL CONSTRAINTS THAT AFFECT PROCESSING LOAD

The difficulty of an integration depends not only upon its resource use, but perhaps even more importantly upon the informational complexity of the resulting structure. Recent work has demonstrated that people's preferred interpretations of (temporary) ambiguities in sentences are affected by factors such as (1) the frequency of the different lexical entries for the word being integrated; (2) the plausibility of the meaning of the resultant structure in the world; and (3) the context that the sentence is uttered in. Consider sentence (1) once again:

The dog walked to the park had been chewing the bone. (1)

The preference for initially analyzing the word 'walked' as a past tense verb rather than a passive participle verb is driven by a number of factors. First, the past tense lexical entry for 'walked' is much more frequently used than the passive participle lexical entry for 'walked', which is used in the passive interpretation. Second, although it is plausible for a dog to be walked, it is even more plausible for a dog to be walking, because there is a general bias to treat animate beings as agents or experiencers of events. Third, there is more syntactic storage cost associated with the passive participle structure at the point of processing 'walked': the passive participle structure requires at least a verb to make a complete sentence (and possibly also a modifier for the verb 'walked'), whereas the past tense structure requires no further words to make a grammatical sentence. Note that integration costs do not have a bearing on the preference, because both potential integrations of the word 'walked' are local.

Now consider (8), a sentence with the same structural ambiguity as in (1), but which does not cause processing difficulty:

The evidence examined by the lawyer turned out to be unreliable. (8)

Like the verb 'walked', the verb 'examined' is ambiguous between a past tense and a passive participle. There are three differences between this ambiguity and the ambiguity in (1) that make the ambiguity in (8) much easier to resolve as a passive participle. First, the relative frequencies of past tense and passive participle lexical entries for the verb 'examined' are less biased towards the past tense reading. Second, and most importantly, plausibility information is highly biased towards the passive participle structure: it is plausible for evidence to be examined, as in the passive participle interpretation, but it is not plausible for evidence to examine something. Third, syntactic storage costs are less biased in favor of the past tense reading, because unlike the past tense entry of 'walked', which is optionally intransitive, the verb 'examined' obligatorily requires a noun phrase object. Thus there is a smaller difference in storage costs between the past tense and passive participle structures than for 'walked'.

Referential context also strongly affects people's initial interpretations of ambiguous structures, as evidenced by monitoring people's eye movements while they listen to commands spoken to them over headphones. Consider the command in (9):

Put the apple on the towel in the box. (9)

In a context with only one apple, there is a strong preference to interpret 'on the towel' as the goal for the verb 'put', even if the apple is already on a towel. But in a context with two apples, one of which is already on a towel, the phrase 'on the towel' is initially analyzed as a modifier of the apple, specifying the apple's source location.

Another informational factor which affects people's initial interpretations of an ambiguous input is the intonation of the speech signal. Intonational (or prosodic) properties include variations in the pitch, amplitude, and duration of individual speech sounds and larger segments, as well as the place of pauses. Intonational phrasing which is consistent with the target syntactic structure helps sentence understanding, and intonational phrasing which conflicts with the target syntactic structure makes sentence understanding more difficult.

IMPLICATIONS AND OPEN QUESTIONS

The details of the relative timing and strengths of the resource and informational constraints are currently not known. An influential early hypothesis in the sentence processing literature was the modularity hypothesis: that syntactic preference constraints (constraints related to the resource constraints described here) apply first, followed by informational constraints such as context and plausibility. This hypothesis predicts that there should be reanalysis effects (small garden-path

effects) in instances where syntactic preferences favor one interpretation, and plausibility and/or contextual constraints favor another interpretation. Recent work using eye-tracking methods has failed to observe such a reanalysis effect. For example, there is no measurable initial preference for the goal interpretation of the phrase 'on the towel' in sentence (9) when the visual context favors the source location, despite the fact that syntactic preferences favor the goal interpretation. As a result, an interactive position has gained in popularity, in which resource and information constraints apply immediately to the available alternatives.

The interacting constraints have been described here as if they are all independent of one another. However, this may not be the case. Many researchers argue that language is implemented in a highly interactive architecture, such as a connectionist neural network of some kind. What may superficially look like independent constraints may be emergent properties of such an architecture. Much further research is needed to investigate this issue.

Further Reading

Crain S and Steedman M (1985) On not being led up the garden path: the use of context by the psychological parser. In: Dowty D, Karttunen L and Zwicky A (eds) *Natural Language Processing: Psychological, Computational and Theoretical Perspectives*, pp. 320–358. Cambridge, UK: Cambridge University Press.

Elman JL (1991) Distributed representations, simple recurrent networks and grammatical structure. *Machine Learning* **7**: 195–225.

Frazier L and Clifton C Jr (1996) *Construal*. Cambridge, MA: MIT Press.

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

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

Gibson E and Pearlmutter N (1998) Constraints on sentence comprehension. *Trends in Cognitive Science* **2**: 262–268.

Grodner D, Gibson E and Tunstall S (2002) Syntactic complexity in ambiguity resolution. *Journal of Memory and Language* **46**: 267–295.

Kjelgaard MM and Speer SR (1999) Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguities. *Journal of Memory and Language* **40**: 153–194.

MacDonald M, Pearlmutter N and Seidenberg M (1994) The lexical nature of syntactic ambiguity resolution. *Psychological Review* **101**: 676–703.

McElree B and Griffith T (1998) Structural and lexical constraints on filling gaps during sentence comprehension: a time-course analysis. *Journal of Experimental Psychology: Learning, Memory and Cognition* **24**: 432–460.

Tabor W, Juliano C and Tanenhaus MK (1997) Parsing in a dynamical system: an attractor-based account of the interaction of lexical and structural constraints in sentence processing. *Language and Cognitive Processes* **12**: 211–272.

Tanenhaus MK, Spivey MJ, Eberhard KM and Sedivy JC (1995) Integration of visual and linguistic information in spoken language comprehension. *Science* **268**: 1632–1634.

Tanenhaus MK and Trueswell JC (1995) Sentence comprehension. In: Miller J and Eimas P (eds) *Speech, Language, and Communication*. San Diego, CA: Academic Press.

Trueswell JC, Tanenhaus MK and Garnsey SM (1994) Semantic influences on parsing: use of thematic role information in syntactic disambiguation. *Journal of Memory and Language* **33**: 285–318.