Constraints on sentence comprehension

Edward Gibson and Neal J. Pearlmutter

The process of comprehending sentences involves the integration of a variety of different information sources, constrained by the available computational resources. This paper surveys the evidence for four types of constraints on sentence comprehension: (1) lexical constraints, (2) contextual constraints, (3) locality-based computational resource constraints, and (4) phrase-level contingent frequency constraints. These four constraints, in combination with grammatical phrase-formation information and prosody, are sufficient to explain how interpretations are constructed for sentences as they are encountered, including complexity effects in unambiguous sentences and interpretation preferences in ambiguous sentences.

The process of understanding a sentence involves combining its meaning by combining the meanings of its individual words. Research on sentence comprehension examines the representations and processes that are needed to connect the identification of individual words in speech or reading1,2 with the mapping of sentence meanings onto mental models3 or discourse representations4. We will focus on a specific current approach, according to which sentence interpretations are formed during sentence comprehension are projected from the representations for individual words6,14–18, and that the later noun refers to properties of the speech signal beyond those related to the identity of words. These properties include variations in the pitch, amplitude and duration of individual speech sounds and cross-modal links in spatial attention46,17–18.

Recent results have suggested that constructing an interpretation for a sentence involves the moment-by-moment integration of a variety of different information sources, constrained by the available computational resources49,50. We assume that the representations constructed during sentence comprehension are projected from the representations for individual words19–23, and that the individual word representations are then combined to form phrases. (See Box 1 for a summary of the kinds of representations involved.) This combination process is constrained by phrase-formation information24–25, which delineates the range of possible combinations. For example, the dog is happy is a well-formed English sentence, but the dog the happy is not. In addition, phrase-formation information partially determines the interpretations for the combinations. For example, Eleanor loved Chris and Chris loved Eleanor mean different things in English (despite involving the same individual word representations), because phrase-formation constraints require that the initial noun in a simple noun–verb–noun sentence be interpreted as the performer of the action specified by the verb, and that the later noun be interpreted as the entity on which the action is performed. We can categorize the remaining constraints into four broad categories, which are relevant for both spoken and written sentence comprehension:

(1) Lexical, or word-level, constraints, which depend on knowledge associated with particular words in a language;

(2) Contextual constraints, which involve the communicative utility26 and the plausibility of different interpretations, given knowledge about the state of the world;

(3) Computational resource constraints, which depend on the availability of and access to working memory resources;

(4) Phrase-level contingent frequency constraints, defined as the probability of phrases occurring in particular phrase structure contexts.

We describe the four kinds of constraints in more detail below, assuming a framework in which all the constraints apply freely. One additional constraint is prosody, which refers to properties of the speech signal beyond those related to the identity of words. These properties include variations in the pitch, amplitude and duration of individual speech sounds and cross-modal links in spatial attention.
The meaning of a sentence is formed from a combination of the meanings of smaller parts of the sentence called constituents or phrases. The meaning of each phrase is inductively determined by the combination of the meanings of its constituents, down to the word level and below. A phrase structure for the sentence, The reporter who the senator attacked defended the report on the web, is provided in the Figure, with phrases shown down to the word level. Each phrase contains a particular word-level category, called the head of the phrase, which is obligatorily present in all instances of that phrase type. For example, the head of a verb phrase (VP) is a verb, the head of a noun phrase (NP) is a noun, etc. Thus, the head of the NP the reporter is the noun reporter, and the head of the VP defended the report on the web is the verb defended. All other categories or constituents whose heads of the phrase that immediately dominate them. For example, the article a is a dependent of the noun reporter in the NP the reporter. The NP the reporter and the prepositional phrase (PP) on the web are dependents of the verb defended in the phrase structure shown.

The sentence (S) in the Fig. is composed of two phrases at the top level of analysis: (1) an NP, the reporter who the senator attacked, and (2) a VP, defended the report on the web. The initial NP is referred to as the subject of the verb defended, and it is interpreted as the individual performing the action specified by the verb. The VP is further broken down into three constituents: (1) a verb, defended; (2) an NP immediately following the verb (the verb’s object), the report; and (3) the PP, on the web. The object NP is interpreted as the entity on which the action is performed, and the PP is interpreted as the location where the action takes place. This division represents one of two possible interpretations of the VP, one in which the act of defending the report took place on the web. The NP the reporter is divided into two further constituents: an article the, and a noun reporter. The PP is also divided into two constituents: a preposition on, and an NP the web.

Under a second interpretation of the VP, the report was on the web, and the reporter defended this report using an unstated medium (speaking, perhaps).

References

Fig. A phrase structure for the sentence, The reporter who the senator attacked defended the report on the web. (See text for explanation of abbreviations.)

Box 1. The hierarchical structure of sentences

The meaning of a sentence is formed from a combination of the meanings of smaller parts of the sentence called constituents or phrases. The meaning of each phrase is inductively determined by the combination of the meanings of its constituents, down to the word level and below. A phrase structure for the sentence, The reporter who the senator attacked defended the report on the web, is provided in the Figure, with phrases shown down to the word level. Each phrase contains a particular word-level category, called the head of the phrase, which is obligatorily present in all instances of that phrase type. For example, the head of a verb phrase (VP) is a verb, the head of a noun phrase (NP) is a noun, etc. Thus, the head of the NP the reporter is the noun reporter, and the head of the VP defended the report on the web is the verb defended. All other categories or constituents whose heads of the phrase that immediately dominate them. For example, the article a is a dependent of the noun reporter in the NP the reporter. The NP the reporter and the prepositional phrase (PP) on the web are dependents of the verb defended in the phrase structure shown.

The sentence (S) in the Fig. is composed of two phrases at the top level of analysis: (1) an NP, the reporter who the senator attacked, and (2) a VP, defended the report on the web. The initial NP is referred to as the subject of the verb defended, and it is interpreted as the individual performing the action specified by the verb. The VP is further broken down into three constituents: (1) a verb, defended; (2) an NP immediately following the verb (the verb’s object), the report; and (3) the PP, on the web. The object NP is interpreted as the entity on which the action is performed, and the PP is interpreted as the location where the action takes place. This division represents one of two possible interpretations of the VP, one in which the act of defending the report took place on the web. The NP the reporter is divided into two further constituents: an article the, and a noun reporter. The PP is also divided into two constituents: a preposition on, and an NP the web.

Under a second interpretation of the VP, the report was on the web, and the reporter defended this report using an unstated medium (speaking, perhaps).

The meanin...
the effect of additional word-level properties has been considered. However, the variety and importance of such constraints in immediate moment-to-moment interpretation has been demonstrated only recently, in studies examining temporarily ambiguous sentences. One such type of ambiguous sentence is shown in (1).

(1) Amanda believed the senator…
(a) …during the speech.
(b) …was lying to the committee.

In (1), the relationship between the phrase the senator and the rest of the sentence is temporarily ambiguous; if the sentence continues as in (1a), then it is the senator who believes. If instead the sentence continues as in (1b), then the senator turns out to be the subject of the embedded sentence the senator was lying to the committee. The ambiguity arises because verbs like believed can license an object noun phrase (NP) [the interpretation forced in (1a)] or an embedded sentence [forced in (1b)], and either possibility is acceptable prior to encountering one of the endings. Furthermore, the ambiguous region can be lengthened considerably (e.g. Amanda believed the senator from the glorious Commonwealth of Massachusetts was lying…), and thus the sentence processor must make some decision about a preferred interpretation during the ambiguous region. One way of making a decision about how to interpret the ambiguous is to make use of some detailed information associated with the verb believed: in addition to knowing that it permits either an object NP or an embedded sentence, comprehenders might keep track of the relative frequency of these alternatives. The relative frequency can be determined by having subjects write whole-sentence completions for fragments like Amanda believed…Believed is used more than three times as often with an embedded sentence (e.g. (1b)) as with an object NP (e.g. (1a)), and thus if the sentence comprehension mechanism relies on this kind of information, it will tend to prefer the embedded sentence alternative for the verb believed. Verbs with relative frequencies favoring the object NP alternative (e.g. understood, which licenses an object NP roughly eight times more often than an embedded sentence) would create the opposite preference—in favor of the object NP alternative. Comprehenders are sensitive to these relative frequencies, showing little or no difficulty in reading continuations like (1b) when the verb licensing the ambiguity more often takes an embedded sentence (e.g. believed), but showing substantial difficulty when the verb licensing the ambiguity more often occurs with an object NP (e.g. understood)6,7.

Similar lexically-based frequency constraints have been shown to influence preferences in a variety of other ambiguities as well, examples of which are shown in (2) and (3). In (2), the verb examined can be the head of the main verb phrase (VP) of the sentence, as is required by (2a), but instead might initiate a relative clause modifying the defendant, as in (2b), where the main VP is had always declared… Preferences in this ambiguity and comprehension difficulty in resolving it are predicted by a set of interrelated lexical properties of the triggering verb (examined), including how often it is used with an object NP and how often it is used in the past tense [required in (2a) as opposed to as a past participle [required in (2b)]6,7,10,11.समानान्तर गठन, जिसे अंग्रेजी में sentence ambiguity in (1) (Ref. 23), the PP attachment ambiguity in (5) (Ref. 33), and the gerund/present-participle ambiguity in (5) below (Ref. 32), and filler–gap ambiguities (Ref. 33).

(2) The defendant examined…
(a) …the confession and then declared his innocence.
(b) …by the lawyer had always declared his innocence.

In (3), the prepositional phrase (PP) with the binoculars can modify the seeing event (so that the spy is using them), but it might instead modify the cop (so that the cop has the binoculars rather than the spy). Preferences in this ambiguity have been argued to depend on the verb’s preference to take PPs headed by particular prepositions as arguments9, as well as the noun’s preference to take the PP NP (e.g. (5a) as opposed to as a past participle [required in (5b)]17,27,28.

(3) The spy saw the cop with the binoculars…

Contextual constraints In addition to constraints associated with specific lexical representations, information that can be computed from the sentence or from the discourse context can also immediately influence sentence comprehension. This is particularly apparent in ambiguity resolution, where a variety of studies have examined the effects of contextual influences on the recognition of ambiguity and referential complexity.

In the case of plausibility, most studies have examined the effect of making one alternative for an ambiguity implausible, typically by manipulating the semantic relationships between verbs and their arguments within a clause or sentence. For example, (4) contains the same temporary ambiguity as (2). The implausibility of evidence examining something leads to a preference for the relative clause interpretation. The evidence examined by the lawyer had always been considered circumstantial.

(4) The evidence examined by the lawyer had always been considered circumstantial.

Plausibility influences in comprehension have also been shown with respect to the object NP versus embedded sentence ambiguity in (1) (Ref. 25), the PP attachment ambiguity in (5) (Ref. 33), the gerund/present-participle ambiguity in (5) below (Ref. 32), and filler–gap ambiguities (Ref. 33).

(5) Visiting relative…
(a) …are fun.
(b) …a fun.

Referential contexts have also been shown to influence many of the same ambiguities, in situations in which the alternatives require reference to single versus multiple referents. For example, modification of a noun by a relative clause (RC) or PP typically involves a presupposition that the noun has multiple referents in the discourse, from which the RC or PP selects. Thus in the ambiguity in (2), the RC alternative (2b) requires a presupposition that the matrix VP alternative (2a) does not, making the RC alternative illicit. However, if the prior discourse context contains the presupposed multiple referents, the RC alternative may be preferred17,27,28.
Referential context effects have also been demonstrated for PP attachment using preceding linguistic context and using concurrent visual context. In the latter case, for example, comprehenders heard the instruction in (6), which contains a temporary ambiguity involving the attachment of the PP on the towel.

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

The participants in the experiment showed no difficulty in interpreting the ambiguity correctly when the instruction was presented concurrently with a display containing two referents for apple, one on a towel and one on a napkin. When only a single referent was visible, however, they had significantly more difficulty interpreting the instruction.

Computational resource constraints: locality

An important constraint affecting sentence comprehension, which is not reducible to lexical or constituent constraints, is locality9,14,17. In ambiguous structures, the locality constraint causes a preference for an interpretation associated with a local attachment over an interpretation associated with a less local attachment. For example, locality explains the existence of the strongly preferred interpretation of (7), in which the adverbial yesterday is associated to the clause the senator left the country, rather than to the clause headed by sold.

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

In addition to its application in ambiguity resolution, locality is also an important factor in determining the processing difficulty of unambiguous structures. For example, in English, a relative clause (RC) whose relative pronoun is coindexed with a gap in object position, such as in (8a), is more complex on a variety of measures than a relative clause whose relative pronoun is coindexed with a gap in subject position, as in (8b) (Refs 41–43). Furthermore, aphasic stroke patients cannot reliably answer comprehension questions about object-GCs, although they perform well on subject-GCs (Ref. 44).

(8a) The reporter who the senator attacked e admitted the error.

(8b) The reporter who e attacked the senator admitted the error.

This difference can be explained by locality considerations: the distance between the relative pronoun who and the gap e is longer with an object-gap, where the NP the senator intervenes, than with a subject-gap, where the two are adjacent. According to one recent theory, locality constrains two central components of sentence comprehension: first, structural integration, consisting of integrating new input words into the currently existing syntactic and discourse structures; and secondly, storage of the syntactic categories that are necessary to complete the current input string as a grammatical sentence. Locality constrains these components as follows: (1) the greater the distance between an incoming word and the head or dependent to which it attaches, the greater the integration cost; and (2) the longer a predicted category must be kept in memory before being encountered, the greater the cost for maintaining that prediction. Under this theory, locality is compared in terms of new discourse referent phrase representing objects and events in the world that can later be referred to using a referential expression, such as a pronoun. Locality accounts for preferences in ambiguous attachment structures such as (7) under the assumption that the sentence comprehension mechanism attempts to minimize the integration cost at each step: integrating yesterday to the VP headed by told involves crossing structures for the intervening material that the suspect left the country, whereas integrating yesterday to the VP headed by sold involves crossing only the NP the country.

The severe complexity of nested structures9,14,17,19 is also explained by locality9. A syntactic category A is said to be nested within another category B if B contains A, a constituent to the left of A, and a constituent to the right of A. The difficulty of processing nested clauses is illustrated in (9):

(9) (9a) [The scientist collaborated with the professor [who, e had advised the student [who, e copied the article]].]

(9b) [The student [who, the professor [who, the scientist collaborated with e] had advised e] copied the article].

Sentence (9a) contains no nested clauses and is relatively easy to understand. On the other hand, in (9b), the relative clause who the scientist collaborated with is nested within the embedded clause the professor had advised, and this clause is nested within the outer clause the student copied the article. The doubly nested structure of (9b) makes this sentence very difficult to understand.

According to the locality-based proposal, the processing difficulty associated with nested structures has to do with the distance between heads and their dependents. In non-nested structures, heads are close to their dependents, often adjacent to one another, whereas some heads and dependents are widely separated in nested structures. For example, in the non-nested structure in (9a), the relative pronouns and their associated gaps are adjacent. In contrast, the relative pronouns and their associated gaps are adjacent.

Sentence (9a) contains no nested clauses and is relatively easy to understand. On the other hand, in (9b), the relative clause who the scientist collaborated with is nested within the embedded clause the professor had advised, and this clause is nested within the outer clause the student copied the article. The doubly nested structure of (9b) makes this sentence very difficult to understand.

Interestingly, doubly-nested structures like (9b) are easier to understand when a first- or second-person pronoun is in the subject position of the outer clause, as in (9c), [e.g. I or you replacing the student in (9b)] (Refs 8,47). The discourse-bias for the computation of locality explains this result. Because all discourses implicitly include a speaker/writer and a hearer/reader, first- and second-person pronouns are old referents in every discourse. Thus the distance between the relative pronouns and their associated gaps is less when I replaces the scientist in (9b), and the structure is correspondingly less complex. Complexity rating experiments across a range of nested constructions in different languages support the predictions of the locality-based theory.

The memory cost component of the theory also provides an account of comprehension effects involving locally ambiguous filler-gap dependencies. It has been observed as follows: (1) the greater the distance between an incoming word and the head or dependent to which it attaches, the greater the integration cost; and (2) the longer a predicted category must be kept in memory before being encountered, the greater the cost for maintaining that prediction. Under
Review

in a variety of languages that, given a filler, people prefer to posit a gap for the filler as soon as possible\(^\text{49–51}\). This preference is found in the memory cost component of the theory: by positing a gap as soon as possible, the processor minimizes the number of syntactic predictions that it needs to store. In addition, the position in which a gap is posited is modulated by plausibility as well as the lexical properties of the licensing head (e.g. a verb)\(^\text{33}\).

Locality in ambiguity resolution interacts with a second factor. This can be seen in the processing of Spanish sentences like (10) below.

(10) El astronauto predijo la órbita del planeta (que se observó desde el satélite).

(The astronaut predicted the orbit of the planet (that was observed from the satellite)).

Outstanding questions

• What is the source of individual differences in sentence comprehension?

• What is the relative strength of the constraints? Some evidence exists that lexical constraints are stronger than contextual constraints\(^\text{52,55–57}\), and that lexical constraints are stronger than the locality constraint in certain circumstances\(^\text{65}\). However, it remains unclear whether these are general properties of the different constraints or whether they are specific to the cases examined. In addition, it remains to be seen how phrase-formation constraints are weighted with respect to the other constraints discussed here\(^\text{19}\).

• What is the source of individual differences in sentence comprehension? Differences in working memory capacity\(^\text{58–60}\), in processing efficiency\(^\text{61}\), and in amount of exposure to language\(^\text{62–64}\) have all been considered. Similarly, what is the relationship between individual differences in language comprehension and in other cognitive systems?\(^\text{65–69}\)

• What is the nature of the representations underlying the constraints? For example, in language production, the existence of phrase-level priming phenomena\(^\text{70}\) has been used to argue for the existence of phrase-level representations. There is more limited evidence for the existence of phrase-level priming in comprehension\(^\text{19}\).

• Are phrase-level contingent frequency constraints necessary to explain comprehension performance, or are the remaining types of constraints sufficient? If phrase-level contingent frequency constraints are necessary, can they subsume the effects of other constraints (e.g. locality)?

• How is distance determined with respect to locality? New discourse refers to an important contributing factor, but whether there are additional components (e.g. other discourse factors, intervening words) is still unclear. In addition, the specific function relating these factors to difficulty remains to be determined.

• What is the source of cross-linguistic attachment preference differences? Differences in corpus frequencies,\(^\text{71}\) in word order\(^\text{72}\), and in lexical pronoun frequencies\(^\text{73}\) have each been proposed as possible explanations.

• What is the source of cross-linguistic attachment preference differences? Differences in corpus frequencies, in word order, and in lexical pronoun frequencies have each been proposed as possible explanations.

• What is the source of individual differences in sentence comprehension?

Because attachments to non-local sites become more difficult with increasing distance, locality weighs more strongly against the attachment of the RC to the first site. The higher locality cost outweighs the competing non-local attachment factor, so that attachment (to planet) is preferred. Attachment to the first site is the next easiest, because the non-local attachment factor favors this attachment.

Attachment to the intermediate site (orbita) is not favored by any constraints, and it is correspondingly the most difficult attachment to make (Ref. 40 and E. Gibson, N. Pearlmutter and V. Torrens, unpublished data). Furthermore, locality and the second factor cannot be lexical-based, because both the two- and three-site ambiguities involve the same attaching phrase (a RC) as well as two of the same potential attachment sites immediately preceding it, yet one site is favored in the two-site ambiguity (orbits in (10)), while the other site is favored in the three-site ambiguity (planet in (11)). Interestingly, the non-local attachment preference observed in the Spanish two-site ambiguity in (10) is not universally present across languages. In particular, English displays a local attachment bias in corresponding English items\(^\text{74}\). To account for the cross-linguistic variability, it has been proposed that the cost associated with violating the non-local attachment factor varies across languages\(^\text{75–77}\).

However, the nature of the factor favoring non-local attachment in these constructions is uncertain, as is the source of these cross-linguistic differences.

Phrase-level contingent frequency constraints

An alternative possible explanation of cross-linguistic and cross-structure preference differences relies on keeping track of the frequencies of ambiguity resolutions contingent on different lexical and phrase structure environments\(^\text{54–58}\). This contingent frequency approach differs from the use of purely lexical constraints in that the frequencies involved are assumed to be tabulated over syntactic constructions rather than individual lexical items (e.g. the NP-Prep-NP-Prep-NP-RC construction in (10), and the NP-Prep-NP-Prep-NP-RC in (11)). However, there is evidence that ambiguity resolution frequencies in naturally produced written text do not always match comprehension preferences\(^\text{76}\). If the
texts that have been analysed thus far are representative of
those that people are normally exposed to, then the contin-
gent frequency approach might have difficulty in accounting
for such a discrepancy.

The strongest evidence for the need to keep track of
contingent frequencies of phrase structures in sentence
comprehension is provided by a set of experiments involv-
ing the grammatical category ambiguity of the word
*that*. For example, in (12a) and (13a) the word *that* is a
demonstrative article, modifying the noun *hotel*. In (12b)
and (13b) on the other hand, the word *that* is a compo-

cent, introducing an embedded sentence *(cheap hotels
w/ clean and comfortable...).*

(12a) That cheap hotel was clean and comfortable.
(12b) That cheap hotels were clean and comfortable.
(13a) The lawyer insisted that cheap hotel was clean and
comfortable.
(13b) The lawyer insisted that cheap hotels were clean and
comfortable.

In sentence-initial contexts like (12), there is a preference
to resolve the ambiguity in favor of the article interpretation
(12a), whereas in post-verbal contexts like (13), there is a
preference to resolve the ambiguity in favor of the compo-

cent (13b) *(Reh 56;57).* Keeping track of contingent frequencies is one way to account for these re-

sults, but there are others. In particular, memory resource
factors might favor the article interpretation in the sent-

ence-initial phrase context, but not differentiate the two
in the post-verbal context, so that the higher frequency of
the complementizer interpretation *(lexical frequency inde-

pendent of phrase-level context)* could determine the pref-

erence in this environment.

Conclusion

We have discussed four types of constraints that are opera-
tive during sentence comprehension: lexical constraints,
contextual constraints, a locality-based computational resource
constraint, and phrase-level contingent frequency constraints.

We have seen a variety of evidence for the first three of these
but noted that the status of contingent frequency constraints is
less clear. These sources of information, in combination with
phrase-formation constraints and, in speech, prosodic
constraints, apply rapidly during normal sentence compre-

hension to determine the interpretations *(for an incoming
string of words). When the incoming string is consistent with
only a single interpretation, the constraints determine the
difficulty of maintaining that interpretation. When the
incoming string is consistent with multiple interpretations
*(in cases of ambiguity), the constraints also determine the

salience preferences for the different interpretations. We pro-

pose that this set of constraints will be sufficient to explain
results from the sentence comprehension literature.

References

Sci. 2, 157–202


Allbritton, D., McKoon, G. and Ratcliff, R. (1996) The reliability of


Garnsey, S.M. et al. (1997) The contributions of verb bias and

plausibility to the comprehension of temporarily ambiguous sentences. J. Mem. Lang. 35, 546–585


Albott, B., McLean, G. and Ratcliff, R. (1994) The reliability of


Garnsey, S.M. et al. (1997) The contributions of verb bias and

plausibility to the comprehension of temporarily ambiguous sentences. J. Mem. Lang. 35, 546–585

MacDonald, M.C. (1994) Probabilistic constraints and syntactic


Trueswell, J.C. (1994) The role of lexical frequency in syntactic


Schütze, C. and Gibson, E. (1996) Argument preference in English PP

attachment, Paper presented at the Linguistic Society of America, San

Diego, CA


Boland, J.E.

Tanenhaus, M.K.


The functional anatomy of language processing, by C.J. Price.