

Parsing as Prediction: Evidence from SVO and OVS constructions in German

Matthew W. Crocker
Psycholinguistics Group
Department of Computational Linguistics
Saarland University
Saarbrücken, Germany
crocker@coli.uni-sb.de

Abstract

In this extended abstract, I briefly outline a number of recent experimental results suggesting that people not only incrementally interpret the utterances they hear, but also make predictions about what they expect to come next. In particular, I will discuss what anticipatory eye movements in visual scenes can reveal about both the mechanisms of incremental comprehension and associated predictive processes.

Introduction

The experimental psycholinguistic literature abounds with evidence that human language comprehension takes place incrementally: people construct rich, compositional interpretations, word-by-word, as utterances are heard or read. One important consequence of incrementality is that people are forced to resolve local ambiguities – occurring during an utterance – in the face of incomplete information. In addition, Tanenhaus *et al.* (1995) have shown that when sentences are heard simultaneously with presentation of a related visual scene, eye movements to objects in the scene can reveal much about which interpretation people adopt and how visual and linguistic information interact in sentence comprehension.

Recent evidence further suggests that the human sentence processor is *anticipatory*: Upon hearing the initial words of an utterance, the sentence processor appears to construct hypotheses about

what is likely to come next. Altmann & Kamide (1999), for example, followed subjects' eye movements in a scene containing a boy, a cake, and a number of toys, accompanied by unambiguous spoken sentences such as "The boy will eat the cake". Upon hearing the verb "eat", subjects launch anticipatory eye movements to the appropriate objects (i.e., the cake) before those objects are uttered.

Here I will briefly summarize recent results from my group regarding *ambiguous* sentence constructions. These studies are novel in that they exploit anticipatory eye movements to determine *preferred* interpretations when hearers encounter a temporary ambiguity. More specifically, the studies focus on ascertaining the influence of morpho-syntactic (Case marking) and probabilistic (word order) constraints, lexical (verb) information, and the use of information available in a visual scene.

Compositional Prediction

One criticism of Altmann & Kamide's findings is that the evidence for prediction (i.e., increased looks to the edible object in the scene), could simply be due to a direct association between "eat" and "cake". To establish that prediction is indeed based on incremental parsing and interpretation, Kamide *et al.* (2002) examined the processing of unambiguous Subject Verb Object (SVO) and Object Verb Subject (OVS) sentences in German. Ambiguity is eliminated in such sentences through the use of Case marking on the sentence initial masculine noun phrase (NP1). Subjects were presented with a scene containing a hare (Hase), a fox (Fuchs), a cabbage (Kohl), and a distractor object. They

simultaneously heard sentences of the following sort:

- (1) Der Hase frißt gleich den Kohl (**SVO**)
The hare-**nom** eats shortly the cabbage-**acc**
- (2) Den Hasen frißt gleich der Fuchs (**OVS**)
The hare-**acc** eats shortly the fox-**nom**

An examination of eye movements in the scene during the post-verbal adverb region (*gleich*) revealed an increase in looks to the appropriate object (the cabbage) in condition (1), and to the appropriate subject (the fox) in condition (2). Crucially, these findings suggest the rapid, compositional use of various information sources in anticipating forthcoming arguments based on entire left context:

- Morphosyntax (Case marking and grammatical function)
- Lexical knowledge (verb selectional restrictions and role assignment)
- World knowledge (plausibility)
- Visual context (scene)

Prediction in Ambiguous Sentences

Both studies outlined above (Altmann & Kamide, 1999; Kamide *et al*, 2002) demonstrate that subjects presented with unambiguously Case marked NP + verb sequences are able to draw upon a range of compositional linguistic constraints in making predictions about what will come next.

The central idea behind the experiments I discuss here is to see what predictive eye movements reveal about the human comprehension of locally ambiguous utterances. The two experiments I will focus on are concerned with the SVO vs. OVS ambiguity in German. While Subjects and Objects can be unambiguously Case marked (masculine), as in the study outlined above, the feminine and neuter forms are ambiguous in German. Our materials all begin with an ambiguous (feminine) NP, followed by the main verb and the disambiguating second NP that is unambiguously marked as either the Subject (Nom) or Object (Acc) of the sentence:

- (3) Die Krankenschwester schubst den Sportler
The nurse-? is pushing the sportsman-**acc**.
- (4) Die Krankenschwester fönt der Priester.
The nurse-? is blow-drying the priest-**nom**.

As sentences are heard, eye movements are monitored in scenes containing the three characters performing the two actions described (i.e. the nurse pushing the sportsman, and the priest blow-drying the nurse). Thus, the nurse is not only ambiguous as Nom/Acc in the sentence, she is also both an agent and patient in the visual scene. Several experiments were conducted using this basic design, with the aim of determining what predictive eye movements can reveal about comprehension during the ambiguous region of the sentence. In particular, we found evidence for the following behaviours:

SVO bias: After fixating the character in the scene corresponding to the initial NP, subjects shift attention to the potential patient/Object (the Sportsman) *before* the verb is processed. This suggests that people have a structural preference to interpret a sentence initial Case ambiguous NP as Subject/Nom (rather than Object/Acc). It further suggests that people assign an Agent role to this noun phrase and launch anticipatory eye movements towards the likely Patient.

Verb+scene effect: Once the verb is processed, subjects are able to combine the verb's role information with the events depicted in the scene to correctly resolve the SVO vs. OVS ambiguity when the initial NP is ambiguous.

Syntactic priming: The SVO bias in comprehension is also subject to syntactic priming. The likelihood of interpreting an ambiguous sentence initial NP as Object is increased if listeners have just read an unambiguous OVS priming sentence, and *vice versa*.

Implications for Computational Models

While numerous models of incremental sentence processing have been developed (see Crocker, 1999, for an overview), such models do not typically address the issue of prediction. In the case of serial processing models, it seems

natural to assume that anticipatory hypotheses are derived from the unique interpretation that has been built so far. However, it is also the case that many predictions are typically possible even if a single interpretation of the preceding context has been selected. It is therefore an important issue whether people construct a single hypothesis about what comes next or do so in some non-deterministic, probabilistic fashion.

In the case of ranked parallel models, the situation is even more complex: Does only the highest ranked interpretation lead to predictions or do all interpretations lead to some non-deterministic set of anticipatory hypotheses? While the data do not at this time allow us to decide among these alternatives, I will outline a general solution based on recent probabilistic, or *experience-based*, models which have been proposed (e.g. Jurafsky, 1996; Crocker & Brants, 2000), arguing that anticipations are best modeled as expectations based on our prior linguistic experience.

At present, however, probabilistic models of human syntactic processing are primarily concerned with ranking the possible interpretations of what has been heard/read so far in terms of their likelihood. Models must therefore be extended to make predictions, characterized loosely as:

(5) $P(\text{Expectations}|\text{Experience, Scene, Context})$

Derived expectations can then be evaluated against human data such as that presented here. Current models must be substantially revised, however, to account for the complex array of linguistic and non-linguistic information sources that contribute to the construction of anticipatory hypotheses during comprehension. These include morpho-syntactic information, verb selectional restrictions, structural biases and the associated assignment of grammatical functions and thematic roles, the plausibility of role fillers, and the roles depicted in the actions of visual scenes. Furthermore, the nature of *what* precisely is predicted remains an important empirical and theoretical issue.

Conclusions

It has long been accepted that human language comprehension is highly incremental, integrating words into a compositional interpretation of the utterance as they are encountered. Our findings demonstrate that the human parser is even more eager than this, constructing anticipatory hypotheses about what is likely to follow, while also revealing how local ambiguities have been resolved. We suggest these findings can best be modelled by probabilistic parsing frameworks.

Acknowledgements

The previously unpublished experimental findings reported here involved Pia Knöferle, Sandra Muckel, and Christoph Scheepers. This research was supported by DFG SFB-378, Pia Knöferle is supported by DFG IGK-715.

References

- Altmann G. and Kamide Y. (1999) Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73, 247-264.
- Crocker M. (1999) Mechanisms for Sentence Processing. In: Garrod & Pickering (eds). *Language Processing*, Psychology Press, UK.
- Crocker M. and Brants T. (2000) Wide Coverage Probabilistic Sentence Processing. *Journal of Psycholinguistic Research*; 29(6), pp. 647-669.
- Jurafsky D. (1996). Probabilistic Model of Lexical and Syntactic Access and Disambiguation. *Cognitive Science*, 20, pp. 137-194.
- Knöferle P., Crocker M., Scheepers C. and Pickering M. (2002). Anticipatory Eye movements in initially ambiguous sentences: There's more to it than meets the eye. Poster presented at: *AMLaP-2002*, Tenerife, Spain.
- Scheepers C., Crocker M. and Knöferle P. (2002) Constituent order priming from reading to listening Paper presented at: *AMLaP-2002*, Tenerife, Spain.
- Tanenhaus M., Spivey-Knowlton M., Eberhard K. and Sedivy J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, pp. 632-634.
- Kamide Y., Scheepers C., Altmann G. and Crocker M. (2002) Integration of syntactic and semantic information in predictive processing: Anticipatory eye movements in German. Paper presented at: *The CUNY Conference on Sentence Processing*, New York, NY.