THE ROLE OF PERCEPTUAL STRATEGIES IN THE PROCESSING OF ENGLISH RELATIVE CLAUSE STRUCTURES

Gary D. Prideaux

Department of Linguistics
University of Alberta
Edmonton, Alberta
Canada T6G 2H1

Summary

In order to assess competing predictions made by several different perceptual strategies, an experiment was conducted, using as stimuli English sentences containing a variety of types of relative clauses. The results indicated that of all the strategies investigated, only Word Order and Interruption played significant roles in the comprehension of the sentences. A similar experiment was then conducted using Japanese sentences with relative clauses, and exactly the same two strategies were found to account for the data. The Given-New strategy was also found to play an important role in the two languages.

Introduction

Over the past few years, functional considerations have come to play an important role in the empirical study of language comprehension processes. This functional perspective assumes that hearers employ a set of perceptual,[1] mental,[2] or cognitive strategies to extract semantic information directly from surface structure. Throughout the psycholinguistic literature, a wide variety of strategies has been proposed, some with a syntactic orientation, and others having a semantic or even a discourse basis. Moreover, some strategies appear to be language specific, while others are language independent. The term "perceptual" is somewhat inappropriate in this context since the processes involved typically do not refer to perception in even its most general sense. The more apt term "cognitive" will be used throughout this paper.

In order to establish the viability and relative importance of various strategies, both simplex and complex sentences have been investigated experimentally. English sentences containing relative clauses provide an especially rich source for experimental study, since English sanctions relative clause formation on NPs playing virtually any grammatical role, while the relative pronoun (RP) can itself play a wide variety of grammatical roles with the relative clause (RC). For ar English sentence containing a subject, a transitive verb, and a direct object, a relative clause can be formed on either the subject or object NP. Moreover, within a relative clause containing a transitive verb, the relative pronour can itself be subject or object. For such sentence types, the following four structures are permitted, assuming a single relative clause per sentence:

SSa: NP[RP V NP] V NP
The man that chased the dog saw the boy.
SOa: NP [RP NP V] V NP
The man that the dog chased saw the boy.
OSa: NP V NP[RP V NP]
The man saw the boy that chased the dog.
OOa: NP V NP[RP NP V]
The man saw the boy that the dog chased.

In the coding to the left of each structure, the first letter represents the grammatical role (subject or object) of the NP on which the relative clause is formed, while the second letter represents the grammatical role played by the relative pronoun. The third letter represents the fact that the relative clause is in the active voice.

A parallel set of structures can be constructed in which the relative clause is in the passive voice. These are:

SSp: NP[RP be Ved by NP] V NP
The man that was chased by the dog saw the boy.
SOp: NP[RP NP be Ved by] V NP
The man that the dog was chased by saw the boy.
OSp: NP V NP[RP be Ved by NP]

The man saw the boy that was chased by the dog.

OOp: NP V NP[RP NP be Ved by]
The man saw the boy that the dog was chased by.

Several strategies have been proposed to account for the differential ease of processing of structures containing relative clauses, although most proposals have concentrated on actives and have not addressed the problem of passive relative clauses. When these strategies are gathered together, however, it becomes obvious that they make different predictions. The original purpose of the research reported here was to evaluate four important strategies bearing on relative clause processing in the context of a single experiment with stimuli based on the eight structures discussed above. The incorporation of passive relative clauses was an important added dimension, since it is the passives which serve to differentiate among the strategies. A related goal was to eliminate from consideration those strategies for which empirical support was not forthcoming. A final goal was to establish a hierarchy among the relevant strategies.

The following cognitive strategies are all relevant to the processing of sentences containing relative clauses:

The Parallel Function (PF) Strategy. Comprehension for sentences containing relative clauses is facilitated if the relative pronoun plays the same grammatical role (S or O) as is played by the modified noun.[8]

The Interruption Strategy. A non-interrupted clause is easier to comprehend than an interrupted clause.[10]

The Word Order Strategy. A clause in normal word order is easier to comprehend than a clause in non-normal word order.[10]

The Adjacency Strategy. In parsing a noncompound sentence, start from the left and group together as constituents of the same clause two adjacent NPs (i.e., those not separated by another NP) and an adjacent verb not already assigned to a clause. Interpret the first NP as the subject and the second NP as the object of the verb.[9]

The Parallel Function strategy was initially proposed[8] to account for English acquisition data, although it was later found[9] to be far less operative for adults. There are two fundamental problems associated with this strategy. The first is that it is not clear whether parallel function is to be defined on underlying or surface grammatical roles. In the passives listed above, surface grammatical roles are indicated for all NPs, including relative pronouns. However, each of those could also be interpreted in underlying terms. For example, the relative pronoun is construed as subject in type SSp, although in deep structure terms it is the object. Consequently, two distinct versions of the strategy are possible, one based on surface grammatical relations and the other based on underlying relations. Both versions are tested here. The second problem is conceptual in nature. In particular, the strategy seems to lack any explanatory power, standing only as an isolated statement of certain results, without independent motivation. In fact, even its relevance for the acquisition data has been challenged.[7]

Both the Word Order and Interruption strategies were proposed[10] to deal with language acquisition phenomena, but both can readily be translated into processing terms, as they have been here. Clearly, neither is language specific, although the Word Order strategy assumes the existence of a "basic" or "normal" word order, usually assumed to be that of the simple, declarative, affirmative sentence. On quite independent grounds, Givon[4] has argued that the simple, declarative, affirmative sentence type is the most basic in many, if not all, languages precisely because it is presuppositionally the least burdened type. The Word Order strategy predicts that any clause which deviates from the normal form will be more difficult to process than one which does not. The Interruption strategy predicts that any sentence with an internally embedded (interrupting) clause will be more difficult to process than one with an embedded clause at one extremity or the other.

Sheldon[9] noted that the Adjacency strategy is basically an English parsing device which applies blindly across a sentence from left to right, assigning a surface grammatical role to each full NP it encounters, and leaving relative pronouns unanalyzed. She pointed out that the strategy sometimes fails to assign grammatical roles correctly. For example, in an SSa sentence like "The

man that chased the dog saw the boy, the strategy assigns The man as subject of chased and the dog as object. It then skips over that and incorrectly assigns the dog as subject of saw, and finally specifies the boy as object. Thus, the Adjacency strategy makes one error for type SSa. Sheldon[9] suggests that the number of errors made by the strategy determines the relative processing difficulty of that type of structure.

Each strategy generates predictions as to the ease of processing of the eight sentence types. If both Deep and Surface versions of Parallel Function are tested, five sets of predictions follow. These are listed in Table 1, where ">" signifies "is easier to comprehend than" or "is more natural than."

TABLE 1. PREDICTIONS FOR EACH STRATEGY

Strategy	Predictions
Deep PF	SSa,SOp,OOa,OSp > SOa,SSp,OSa,OOp
Surface PF	SSa,OOa,SSp,OOp > OSa,SOa,OSp,SOp
Interruption	OOa,OSa,OSp,OOp > SOa,SSa,SSp,SOp
Word Order	SSa,OSa,SSp,OSp > SOa,OOa,SOp,OOp
Adjacency	OSp,OSa > SSa,OOa,SSp,OOp > SOa,SOp

The English Experiment

In order to test the predictions, a single experiment was conducted using as stimuli 56 sentences in written form, with seven separate tokens (replications) of each of the eight types. The lexical items were varied across all the sentences. The task of the subjects, eighteen native speakers of English, was to evaluate each of the sentences in terms of relative ease of comprehension or naturalness on a nine-point scale, with "1" the most natural or easiest to understand and "9" the least. Subjects were permitted to work at their own rates and were urged to ignore as far as possible the actual lexical items, focusing their attention rather on the forms of the sentences.

The naturalness data were analyzed on the University of Alberta's Amdahl 470V/6 computer, using a packaged (BMD:08V) three-way analysis of variance program, with the factors of subjects, syntactic type (four levels: SS, SO, OS, OO), and voice (two levels: active and passive). All the main effects were significant ($p<.01$), but more importantly the type by voice interaction was also highly significant ($F(3,51)=20.48$, $p<.001$), with the sentences containing active relative clauses judged significantly more natural than those containing passive relative clauses. Consequently, the two groups were analyzed separately. Planned comparisons were carried out on each group to determine which types within each group were judged to be significantly the more natural. Each of the five strategies was tested within the two groups. Within the group containing active relative clauses, the only significant factor was Interruption, with types OOa and OSa judged significantly more natural than SSa and SOa ($F(3,119)=14.27$, $p<.001$). None of the other strategies was operative in this group.

Within the group containing passive relative clauses, the only significant factor was Word Order. The types OSp and SSp, which have a relative clause word order of SVO, were judged significantly easier than types SOp and OOp, which have relative clause word order of OSV ($F(3,119)=64.60$, $p<.001$). These results are all summarized in Table 2.

TABLE 2. ENGLISH NATURALNESS JUDGEMENTS

Voice	Naturalness	Type	Word Order
Active	1	OSa	S V O[SVO]
		OOa	S V O[OSV]
	2	SOa	S[OSV] O V
		SSa	S[SVO] V O
Passive	3	OSp	S V O[SVO]
		SSp	S[SVO] V O
	4	OOp	S V O[OSV]
		SOp	S[OSV] V O

Interpretation

The first question to address in interpreting these data is why those sentences containing passive relative clauses were judged significantly more complex and less natural than those containing active relatives. To unravel

this complex issue, it is instructive to examine several factors associated with the active/passive distinction. Within an active clause, the subject of a transitive action verb is typically, although not always, interpretable as a semantic agent, while for the passive, it is certain that the subject is not an agent. In fact, Givon[4] has suggested that an important function of the passive construction is to place a non-agent NP into subject position. A passive clause therefore deviates from the "normal" case of subject as semantic agent. Consequently, while both active and passive clauses can be characterized as having a surface SVO word order, the subject of the passive is non-agent, and the object is the object of the preposition by.

Another important difference is that passives are far less frequent than actives. In discussing text counts made over a broad spectrum of genres, Givon[4] reported that some 90% of the affirmative, declarative sentences were actives, and only 10% were passives. Furthermore, only 20% of the latter were "full" passives with overt agentive by phrases, while 80% were truncated passives. Accordingly, only 2% of affirmative, declarative sentences are "full" passives. Givon[4] suggested the special discourse properties and presuppositions associated with the passive might account for their low frequency. For example, the subject of a passive clause is not a potential agent, but is more typically a semantic patient. However, the subjects of declarative sentences tend to contain Given information, with the objects more likely to be New. This follows from the Given-New strategy,[2] according to which Given information normally precedes New in a sentence. Some 90% of the subject NPs in active sentences are definite, while 93% of the subject NPs of passive sentences are definite.[4] In general, then, the subjects of both active and passive sentences tend to be definite and Given. In active sentences, however, only about 56% of the direct object NPs are definite, and the direct object is far more likely to contain New information than is the subject. Similarly, if there is an overt object (agentive) phrase in a passive, it is almost invariably New information.

Based on these observations, we can extract the following general facts concerning the distinction between active and passive clauses. Actives are more frequent, have subject as agent, and object as patient. Furthermore, the subject of an active is typically definite and Given, while the object is less likely to be definite and more likely to be New information. The passive, on the other hand, has a definite, non-agentive, typically Given subject and if it has an object at all, the object is agent, typically definite, and New. The differences are clearly signalled by the word order and morphological factors associated with the passive. Consequently, it would appear that the global characteristic of voice has associated with it a host of syntactic, semantic, and discourse properties, with the passive being by far the more "marked" and less expected form.

Once the two groups have been separated by voice, we can examine the factors within each group which give rise to further subgroupings. Within the actives, the Interruption strategy separates the four groups into two, demonstrating that comprehension is more difficult when a relative clause interrupts the main clause. Within the passives, Word Order was the only significant factor: a passive relative clause with SVO word order was judged easier to comprehend than one with an OSV word order. Interruption played no role among the subgroupings of the passives.

The experiment demonstrated the importance of Interruption and Word Order, while neither version of Parallel Function nor Adjacency emerged as significant. However, the two operative strategies are not equally salient, and in particular, it appears that Interruption is important only for the active structures, in which the normal expectation of subject as agent is met. For the non-normal (passive) cases, Word Order is very important. Consequently, it appears that Word Order must be satisfied before Interruption can be called into play.

At this point, something further must be said about the consequences of the Given-New strategy within relative clauses. Since it has an antecedent, a relative pronoun typically represents Given information. Consequently, in a relative clause with the structure [RP V NP], the relative pronoun is subject and is in the position associated with Given information. For relative clause structures of the form [RP NP V], the relative pronoun is again the first NP in the clause, and satisfies the Given

position. Furthermore, it is just this Given RP which can be successfully deleted. However, the subject NP, which is also typically Given information, now finds itself in the New position, according to the Given-New strategy. Consequently, this type of relative clause structure, with word order OSV, might be expected to be somewhat less natural in terms of the Given-New strategy, than the former, with the word order SVO. No such significant difference was found in these data, although such a result was reported by Lynkowsky,[6] who conducted a similar experiment, but used only active relative clauses. In her experiment, there was no deviation from the expected, normal case of subject as potential agent, and consequently it would be expected that the dominant Word Order strategy would be the first strategy to be called into operation.

Within the passive group, the issue is somewhat more complex. The passive relative clauses can have one of two forms, either [RP be Ved by NP], with the word order of SVO, or [RP NP be Ved by], with an OSV word order. In both cases, the normal expectation of subject as agent is violated. The results of the experiment indicate that it is the former, SVO, passives which are judged far more natural than the latter, OSV structures. In terms of the Given-New strategy, this makes a great deal of sense. In the SVO case, the relative pronoun appears to be Given for two reasons: it is a subject and it is a relative pronoun. The agentive NP object is precisely where New information should be. In the OSV case, however, the relative pronoun should be Given since it is a relative pronoun, but it should be New since it is the object of the preposition by. Furthermore, the subject NP is in the New position, but as subject it should be Given. Consequently, the Given-New strategy seems to be working against itself in the passive relative clauses with OSV word order.

One final aspect of the Given-New factor must also be mentioned. Typically, definite NPs are construed as Given information. In the stimuli for the present experiment, however, all NPs were definite. If a relative clause is formed on a definite NP, there may be a tendency to view the relative clause as adding to the definiteness or specificity of the NP, thereby making it even more "Given," regardless of where it is placed in the sentence. Consequently, within the active group, the Given-New strategy might be viewed as a "force"[3] favoring structures with the relative clauses on the subject NP, while Interruption would be a force favoring structures with relative clauses on object NPs. Accordingly, these two forces are in competition within the actives. In the passive group, however, the Word Order strategy favors relative clauses of the form SVO, and the Given-New strategy also favors such structures. Here, the two strategies work together.

The Japanese Experiment

At this point, brief mention will be made of a Japanese experiment similar to the one reported above, but focusing only on active relative clauses. The same methodology was employed, with 24 native speakers of Japanese serving as subjects. The stimuli were twelve sentences, with three replications of four types. Japanese is an SOV language, and the relative clause precedes the modified NP. Furthermore, Japanese relative clauses do not contain relative pronouns, but rather exhibit deletion of the relativized NP. The following four structures were tested:

SS [NP+o V]NP+ga NP+o V
okane o mitsuketa otoko ga sharei o moratta
money OM found man SM reward OM received
"The man who found the money received the reward."

SO [NP+ga V]NP+ga NP+o V
kodomo ga karakatta inu ga te o kanda
child SM teased dog SM hand OM bit
"The dog that the child teased bit (his) hand."

OS NP+ga [NP+o V]NP+o V
shoojo ga sakana o tabeta neko o tataita
girl SM fish OM ate cat OM spanked.
"The girl spanked the cat that ate the fish."

OO NP+ga [NP+ga V]NP+o V
gakusei ga kyooju ga kaita hon o yonda
student SM professor SM wrote book OM read
"The student read the book that the professor wrote."

The data were analyzed on the University of Alberta's Amdahl 470V/6 computer, using the BMD:08V two-way analysis of variance program with subjects and types (four levels: SS, SO, OS, OO) as factors. The only

significant factor was type ($F(3,69)=115.54$, $p<.001$). Planned comparisons were used to test for each of the strategies except for the English-specific Adjacency strategy. The first comparison indicated that types SS and OS were significantly easier and more natural than types SO and OO ($F(3,69)=318.27$, $p<.001$). There was no significant difference between OO and SO, although type SS was significantly easier than OS ($F(3,69)=28.25$, $p<.001$). These results are represented in Table 3.

TABLE 3. JAPANESE NATURALNESS JUDGEMENTS

Naturalness	Type	Word Order
1	SS	[OV]S O V
2	OS	S [OV]O V
3	OO	S [SV]O V
	SO	[SV] O V

The factor separating the first two types from the second pair is clearly Word Order: sentences containing relative clauses with OV word order are preferred over those containing relative clauses with the SV word order. Moreover, within the pair containing OV relative clauses, Interruption determined that the non-interrupted SS type was preferred over the interrupted OS type. Consequently, in Japanese, as in English, the Word Order strategy is the more important of the two, with Interruption only called into operation when the normal word order is met. In her study of the acquisition of compound and complex sentences in Japanese, Kawashima[5] found that left-branching structures were mastered before those containing center embeddings, thus supporting Interruption as the major factor in the acquisition of relative clauses.

Only speculation can be offered as to why the OV word order is preferred to the SV. Japanese, like English, appears to obey the Given-New strategy, but unlike English, Japanese permits the omission of an "understood" (e.g., Given) NP, making sentences without overt subjects quite common. Within relative clauses, the Given, relativized NP is omitted. But since Given subjects may also be omitted, it follows that the OV clause type would be more natural and common than the SV type when the verb is transitive. In short, the Japanese data support the cross-linguistic viability of the Word Order, Interruption, and Given-New strategies.

Conclusions

Originally, this research was undertaken in an attempt to sort out competing cognitive strategies relevant to the processing of complex sentences with relative clauses. That goal was achieved, and the importance of both Word Order and Interruption was demonstrated, while both versions of Parallel Function and Adjacency were rejected. However, other strategies were also found to be operable. The semantic strategy which associates subject with agent is in fact a version of Word Order. In addition, and perhaps most importantly, the Given-New strategy was found to be of extreme importance, both for English and for Japanese. Finally, a hierarchy of strategies has suggested itself. The Given-New strategy is obviously of great importance for sentences in context, and although the stimuli used in the two studies reported here were presented in isolation, the strategy still seems to be operative. Furthermore, within a particular sentence, it appears that Word Order criteria must be satisfied before Interruption is employed.

There are still numerous loose ends to be investigated. The definiteness of NPs must be varied experimentally to tighten up the tentative suggestions concerning the Givenness of definite NPs; sentences must be carefully studied in contexts; text counts should be undertaken to establish the proportion of relative clauses and their positions, in accordance with such factors as definiteness, Interruption, and the Given-New strategy. The present study has only scratched the surface, but at least the results are encouraging in that they accord well across two very different languages, providing a start on the problem of the interaction of cognitive strategies.

Acknowledgements

I am grateful to Wm. J. Baker for his comments and advice in matters statistical, to Matthew S. Dryer for his insightful metaphor of competing strategies as "forces," and to Michiko Kawashima for collecting the Japanese data. I am responsible for all errors. This manuscript was formatted on the University of Alberta "Textform" system.

References

1. Bever, T. G. The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), Cognition and the development of language. New York: John Wiley & Sons, 1970, 279-352.

2. Clark, H. H., & Clark, E. V. Psychology and language: An introduction to psycholinguistics. New York: Harcourt Brace Jovanovich, 1977.

3. Dryer, M. S. The positional tendencies of sentential noun phrases in universal grammar. Canadian Journal of Linguistics, forthcoming.

4. Givon, T. On understanding grammar. New York: Academic Press, 1979.

5. Kawashima, M. The acquisition of Japanese relative clauses. Unpublished M. Sc. Thesis, University of Alberta, 1980.

6. Lynkowsky, P. E. The development of relative clauses: Comprehension strategies in English and Ukrainian. Unpublished Ph. D. dissertation, University of Alberta, 1980.

7. Prideaux, G. D. The acquisition of relative clauses: A functional analysis. Canadian Journal of Linguistics, 1979, 24, 25-40.

8. Sheldon, A. The role of parallel function in the acquisition of relative clauses in English. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 272-281.

9. Sheldon, A. On strategies for processing relative clauses: A comparison of children and adults. Journal of Psycholinguistic Research, 1977, 6, 305-318.

10. Slobin, D. I. Cognitive prerequisites for the development of language. In C. A. Ferguson & D. I. Slobin (Eds.), Studies of child language development. New York: Holt, Rinehart & Winston, 1973, 175-208.