The Minimal Chain Principle: A Cross-linguistic Study of Syntactic Parsing Strategies

Marica De Vincenzi
Department of Psychology/ University of Massachusetts

The experimental work presented here is part of a series of studies done on sentence processing in Italian. The goal is to provide a cross-linguistic testing of sentence processing models. Most studies in sentence processing have been done in English. Cross-linguistic study is motivated by the possibility that the processing principles identified to date and consequent processing models, may have been biased by language-specific aspects of English. From this last perspective the comparison of English and Italian is particularly interesting because the two languages differ in the setting of a linguistic parameter. Some natural languages allow phonetically null subjects in tensed clauses (henceforth 'pro', a short form for lexically null 'pronominal' element), while others do not. The two types are instantiated by Italian and English, as in (1) (The * means that the English sentence corresponding to the Italian one is ungrammatical. I give the grammatical English translation in parenthesis). It has been shown that other properties systematically correlate with the null subject properties, including a free process of subject inversion (2).

(1) a. pro CORRERA'
    b. *pro WILL RUN.
       (he/she/it will run)
(2) a. pro_i CORRERA' $GIANNI_i$
    b. *pro_i WILL RUN $GIANNI_i$.
       (Gianni will run)

As illustrated in the examples (1) and (2) in Italian there are two types of lexically empty categories that are missing in English: the 'pro' subject in (1) and the 'pro' subject in (2) that is linked to another NP in the same sentence in a filler-gap relationship. (In the standard analysis the post verbal subject is moved form the original pre-verbal position in (2)). The work presented here tests whether these empty elements are processed following the processing principles already identified in English (Crain & Fodor, 1985; Stowe, 1986; Frazier & D'Arcais, submitted; Clifton & Frazier, in press).

SPECIFIC HYPOTHESIS TESTED

The specific processing principle that is tested here uses the notion of chain. A chain is defined as a set of coindexed elements that share one thematic role (e.g., agent, patient, etc.) and one case (e.g., nominative, accusative, etc). The shortest chain is one in which an element is in its unmoved position and it is not linked to any other elements in the sentence, as the NP 'John' in (3):

(3) John runs.

Chains can also have two elements as in the filler-gap relationships of (4) and (5) ('e' designates a phonetically empty structural position):

(4) What $girl_i$ did you call e_i?
(5) $Mary_i$ decided e_i to go.

In (4) the semantic role of the Wh-word (the filler) depends on structural properties of the sentence that follows, namely that there is a 'gap', or empty position in the phrase structure at the direct object position resulting in an empty patient role. Similarly, the

interpretation of the lexically empty structural position (the gap) of the subject of the infinitival 'to go" in (5) is controlled in interpretation by the NP 'Mary' (the filler). In both of these cases, then, there is a chain of length 2 between two elements. Going back to the Italian examples, in (2) there is a chain of length 2 between the Inverted-subject (henceforth, I-subject) and the pre-verbal subject position. In (1) there is a chain of length 1. The lexical empty element in subject position is unmoved and not linked to any other element.

The processing principle tested here is the Minimal Chain Principle (henceforth, MCP):

Minimal Chain Principle:
Avoid postulating unnecessary chain members at S-structure, but do not delay required chain members.

The second clause of the principle amounts to a statement of the "gap-as-first-resort" principle (Fodor, 1978) or the Active Filler Hypothesis (Frazier & D'Arcais, submitted). Here I will focus on the first clause of the principle: it suggests that when a parse is ambiguous between a shorter and a longer chain, the parser will prefer the shorter chain. In the limit, this means that the parser will prefer to have no multiple-element chains at all. Let us take as examples sentences (1) and (2). When a reader gets the verb, she/he knows that there is a lexically empty subject position. But the parse of the pre-verbal subject position is ambiguous between a pro in a singleton chain (as in (1)) or a pro in a multiple-element chain (as in (2)). Several interesting questions arise at this point: does the parser commit itself to a choice of of empty category in the absence of disambiguating information? And if yes, what is the choice that the parser makes? I assume that the parser commits itself on-line to a structural choice without waiting for disambiguation (see Frazier & Rayner,1982) and I will test whether the human parser follows the MCP principle in making choices. This predicts that the parser will initially pursue the singleton chain analysis. This will prove to be correct in (1), but not in (2) where a post-verbal NP occurs. Therefore sentences like (2) should be harder to parse than sentences like (1).

Here I will present two experiments that have tested the MCP principle. The first experiment tested the prediction of the MCP sketched above, using sentences like (6) which are structurally ambiguous between the interpretation (6a) and (6b) because the verb is optionally transitive:

(6) HA CHIAMATO GIOVANNI.
  a.*pro CALLED GIOVANNI.
    (he/she/it called Giovanni)
  b. *pro$_i$ CALLED GIOVANNI.
    (Giovanni called)

If the parser obeys the MCP principle, then the interpretation (6a) should be preferred over (6b). Therefore we predict that when disambiguating material occurs later in the sentence, (6b) should be harder to process than (6a) because a revision of the initial assignment is needed.

The same prediction was tested in a second experiment with intransitive verbs belonging to two different classes, ergatives and unergatives. These two verb classes differ in having the subject originating in post-verbal position (7a) or in pre-verbal position (7b) (Belletti, 1988), and are easily identified because they select different auxiliary verbs:

(7) a. E' ARRIVATA pro.
    * IS ARRIVED pro.
      (she/it arrived)

b. pro HA ESITATO.
*pro HAS HESITATED.
(he/she/it has hesitated)

In the case of a lexically expressed subject in post-verbal position, we have a singleton chain in (8a) and a multiple-element chain in (8b):

(8) a. E' ARRIVATA UNA AMICA.
* IS ARRIVED A FRIEND
(A friend arrived)
b. pro_i HA ESITATO UNA $AMICA_i$.
*pro_i HAS HESITATED A $FRIEND_i$.
(A friend has hesitated)

The MCP principle predicts that a structure like (8a) should be easier to process than (8b) because in the latter case the reader must form a chain relating the post-verbal subject to the pre-verbal position. However, if the verb is ergative, as in (8a) no chain will need to be formed.

EXPERIMENTS

Experiment 1 used 20 triplets of sentences with a temporary ambiguity between a pro-subject (as in (6a)) or an Inverted-subject (as in (6b)): Two sentences had an optionally transitive verb (condition 1 and 2) and one had an obligatorily intransitive verb (condition 3). Each continued with a phrase that disambiguated it. Sample sentences for Experiment 1 are the following:

cond.1: Ieri mattina ha chiamato il capoufficio per chiedergli un aumento di stipendio. (*Yesterday called the boss to ask-him a raise)

cond.2: Ieri mattina ha chiamato il capoufficio per offrirci un aumento di stipendio.(*Yesterday called the boss to offer-us a raise)

cond.3: Ieri mattina ha telefonato il capoufficio per offrirci un aumento di stipendio. (*Yesterday telephoned the boss to offer-us a raise).

The rationale clause in the first example pragmatically forces a reading in which "the boss" is the object of "called" and the matrix sentence has a null subject in pre-verbal position (pro). The rationale clause in the second sentence forces the reading in which "the boss" is the subject, occurring in post-verbal position. In the first case, no chain is needed. In the second case, a 2-element chain must be created when the disambiguating phrase is encountered. Following the MCP the disambiguating phrase in condition 1 will not require reanalysis and will be read quickly. The disambiguating phrase in condition 2 does require reanalysis and will be read slowly. When the verb is obligatorily intransitive (cond. 3: "telefonare" is intransitive in Italian) there is no ambiguity of the post-verbal NP. If the parser is sensitive to this fact, it will not initially choose the pro-subject interpretation, and thus it will not have to perform reanalysis when the rationale clause beginning "per offrirci" arrives.

Experiment 2 used 20 pairs of sentences with ergative and unergative verbs that have post-verbal subjects, as in (8). They occurred in two conditions, with ergative and unergative verbs:

cond.1: Questa volta e' venuta una cara amica ad aiutarci a traslocare. (*This time came a dear friend to help us move)

cond.2 Questa volta ha esitato una cara amica ad aiutarci a traslocare. (*This time hesitated a dear friend to help us move).

Reading Time should be fast for the ergative post-verbal subject sentences (cond.1) as compared to the

Table 1
Average Reading Time for Correct Responses in Experiment 1.

segment #	1	2	3	4	5	%
condition 1 (V_t, pro subject)	1006	1102	950	798	793	.662
condition 2 (V_t, I-subject)	1029	1094	1047	778	853	.718
condition 3 (V_i, I-subject)	991	1124	957	761	812	.929

Note: V_t = verb transitive; V_i = verb intransitive.
*: only in condition 1 the correct response was: "I do not know"

examples: segment #

	1	2	3	4	5
cond. 1	Yesterday/	called the boss/	to ask him/	a raise/	in stipend.
cond. 2	Yesterday/	called the boss/	to offer us/	a raise/	in stipend.
cond. 3	Yesterday/	telephoned the boss/	to offer us/	a raise/	in stipend.

Table 2

Average Reading Time for each segment in Experiment 2.

segment #	1	2	3	4	%
condition 1 (V_e)	989	1043	894	893	.980
condition 2 (V_u)	998	1300	925	881	.883

Note: V_e = verb ergative; V_u = verb unergative.

examples: segment #

	1	2	3	4
cond. 1:	Yesterday/	came a friend /	to help us/	move.
cond. 2:	Yesterday/	hesitated a friend/	to help us/	move.

unergative post-verbal subject sentences (cond.2), because of the necessity of forming a chain in the unergative but not the ergative case.

PROCEDURE AND DESIGN

The 2 experiments were run together. They used phrase-by-phrase self-paced reading task. Subjects pushed a key with the right hand to start each trial, and they pushed the same key to see each successive phrase of the sentence. Only one phrase was visible at a time. When each phrase appeared, the previous one (to the left) disappeared. Subjects were instructed to read at a natural rate, while

maintaining good comprehension. After each sentence, a comprehension question appeared on the screen all at once and the subject had to respond to the question by pressing a button. The question queried the subject of the sentence (for condition 1 and 2 of both experiments) or some randomly chosen aspect of the sentence (condition 3 of experiment 1 and fillers).Reaction Time and question-answering time recorded.

The subjects were sixty college students of the University of Padova (Italy). They received written instructions, then saw 20 practice trials, followed by the experimental and filler trials in randomized order. No subject saw more than one version of any sentence and, accross subjects, every experimental sentence appeared in each condition.

RESULTS

Experiment 1: the data for Experiment 1 are presented in Table 1; the mean reaction times were computed for each segment, after eliminating times that were longer than 5000 msec. and shorter than 100 msec. Reading times associated with erroneous answers to the questions were discarded and the data were normalized by eliminating those subjects (nine) who made over 80% errors. In the critical disambiguating region (segment 3), there was a significant difference in reading time in three conditions, (F_1 (1,50)=3.16, $p<.04$) and (F_2 (1,19) =3.45, $p<.04$). Pairwise comparisons showed that condition 1 was faster than condition 2 (F_1 (1,50)=3.89, $p<.05$) and (F_2 (1,19) = 4.67, $p<.04$) as was condition 3 (F_1 (1,50)=3.75, $p<.05$), (F_2 (1,19)=5.20, $p<.03$). The difference between condition 1 and condition 2 remained significant ($p<.06$) when a regression analysis was used to adjust for length and frequency differences between the verbs. On the 5th segment there was still a significant difference in reading time (F_1 (1,50)=,2.92 $p<.05$), (F_2 (1,19) =2.97, $p<.06$). A pairwise comparison showed that the only significant difference was between condition 1 and 2 (F_1 (1,50)=4.90, $p<.02$), (F_2 (1,19)=4.65, $p<.04$). The questions after the sentences were answered more accurately in condition 3 than in conditions 1 and 2, (F_1 (1,50)=46.62, $p<.001$), while the latters did not differ significantly from one another (F_1 (1,50)=2.68, $p<.11$).

Experiment 2: the 2nd segment, containing the verb plus the post-verbal subject, was read faster for ergative than for unergative verbs. The difference was significant both in a simple analysis of variance (F_1 (1,59)=52.28, $p<.001$), (F_2 (1,19)=72.39, $p<.001$) and in a regression analysis, adjusting for length and frequency differences among the verbs ($p<.04$). Similarly, questions were answered more accurately following ergative sentences than following unergatives (F_1 (1,59)=13.31, $p<.001$.).

DISCUSSION

Both experiments confirm the hypothesis that the parser follows MCP, preferring shorter chains. In experiment 1 the sentences whose pragmatics are consistent with the theoretically preferred pro-subject analysis are easier to read than sentences whose pragmatics force the theoretically unpreferred Inverted-subject reading. This result is especially surprising because a pro has a contextually given antecedent in normal usage, but it does not in the experimental setting. In context, therefore, there should be an even stronger preference than the one found here. When the sentence contains an obligatorily intransitive verb (as in condition 3) eliminating the ambiguity of the post-verbal NP,

reading time for the critical segment "per offrirci" is fast relative to "per chiedergli", as if it does not occasion any reanalysis.

Experiment 2 shows that sentences with a subject after an ergative verb are easier to process than similar sentences with unergative verbs. This confirms the MCP principle that the parser prefers analysis that require chains of the smallest possible length. It also provides support for an analysis of the syntax of Italian ergative verbs like that presented by Belletti (1988).

CONCLUSION

We have shown that the human sentence parser follows a structural principle of Minimal Length of Chain in parsing lexically empty elements and in establishing filler-gap relationships. This principle can be seen as a generalization of the Active Filler principle (assign an identified filler as soon as possible) which has received support from experiments done on English and Dutch (Clifton & Frazier, in press; Frazier & D'Arcais, submitted). Thus, we may offer a uniform account of processing empty categories which holds for languages varying in their inventory of empty categories.

This principle, finally, is consistent with a view of language processing as operating under time pressure and short-term memory limitations. One can assume that constructing a multiple-element chain is costly in terms of time, because there is a delay in assigning an element to its argument position. Keeping elements in memory in an unstructured form is generally assumed to be more costlythan mantaining them in an unstructured form (Miller, 1965). Similarly it can be argued that a multiple-element chain is costly in terms of effort, when several elements must be kept in short-term memory (Wanner & Maratsos, 1978). The consequence is that chians containing the smallest possible number of elements will be constructed before chians involving more elements because the former are faster and place less burden on the language processor.

REFERENCES

Belletti, A. (1988). The case of unaccusatives. Linguistic Inquiry, 19, 1, 1-34.

Clifton, C., Jr. & Frazier, L. (in press) Comprehending sentences with long distance dependencies. In M. Tanenhaus & G. Carlson (Eds) Linguistic Structure in Language Processing. Dordrecht: Reidel.

Crain, S. & Fodor, J. (1985). How can grammar help parser? In D. Dowty, L. Kartunnen & A. Zwicky (eds), Theoretical Perspectives on Natural Language Parsing. Cambridge: Cambridge University Press.

Fodor, J. (1978). Parsing strategies and constraints on transformation. Linguistic Inquiry, 9, 427-474.

Frazier, L. & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. Cognitive Psychology, 14, 178-210.

Frazier, L. & Flores D'Arcais, G. B. (submitted). Filler-driven parsing: A study of gap-filling in Dutch.

Miller, G. (1965). The magical number seven, plus or minus two, or some limits on our capacity for processing information. Psychological Review, 63, 81-96.

Stowe, L. (1986). Parsing Wh-constructions: evidence for on-line

gap location. Language and Cognitive Processes, 1, 227-245.

Wanner, E. & Maratsos, M. (1978). An ATN approach to comprehension. In M. Halle, J. Bresnan & G. A. Miller (Eds) Linguistic Theory and Psychological Reality. Cambridge, Mass: MIT Press.