Empty Category and the Effect of Teaching in Sentence Processing

Takako KAWASAKI
Hosei University
2-17-1 Fujimi, Chiyoda-ku
Tokyo, 102-8160
JAPAN
takako.k@dream.com

Kiyoshi ISHIKAWA
Hosei University
2-17-1 Fujimi, Chiyoda-ku
Tokyo 102-8160
JAPAN
ling@iij4u.or.jp

Abstract

Different syntactic frameworks have different ways to deal with dislocation constructions. One major disagreement is whether or not empty categories should be assumed. Researchers have been working on this issue on the ground of "psychological reality" (of empty categories), yet have not come to an agreement. The first aim of this paper is to propose an experimental scheme to settle the issue of empty categories. Our second aim is to propose that the application of the experiment to L2 is beneficial to see the teaching effect in L2 acquisition. If native speakers' result supports no-empty-category analysis, yet, L2 learners exhibit the different result, L2 teaching of "wh-movement" created a category which is not in native speakers' mental grammar.

1 Introduction

This paper proposes an experimental scheme which solves two independent problems in sentence processing. The first problem concerns the role of empty categories in sentence processing. The second is regarding the sentence processing in L2 acquisition.

2 Backgrounds

One major disagreement among syntacticians is whether to posit an empty category or not in analyzing dislocation constructions. For example, consider (1):
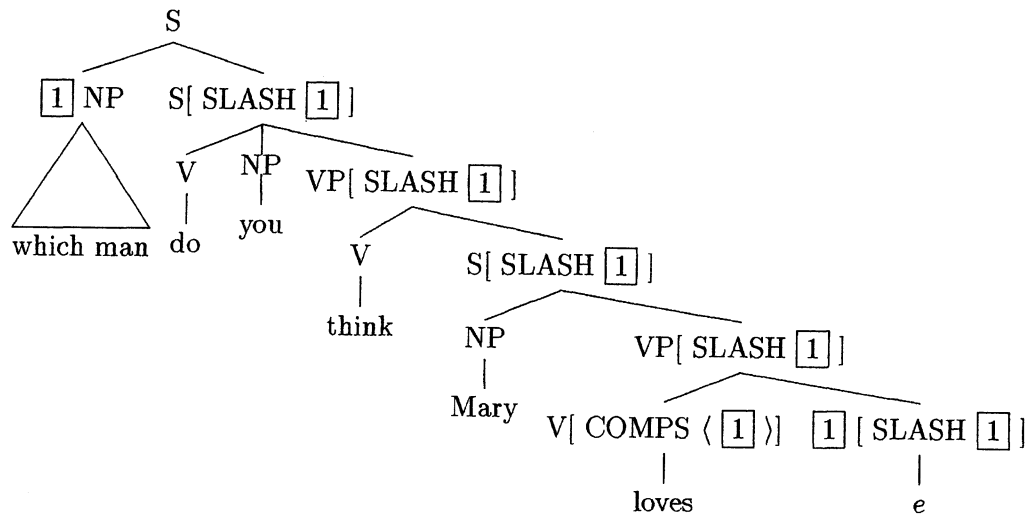
(1) Which man do you think Mary loves?

Under the phrase-structure-based theories, whether transformational (GB/MP; Chomsky 1981, 1993) or non-transformational (GPSG; Gazdar et al. 1985), the dependency between *which man* (the filler) and *loves* (the head that selects it) is usually mediated through an empty category (trace) located in the filler's "canonical" position.
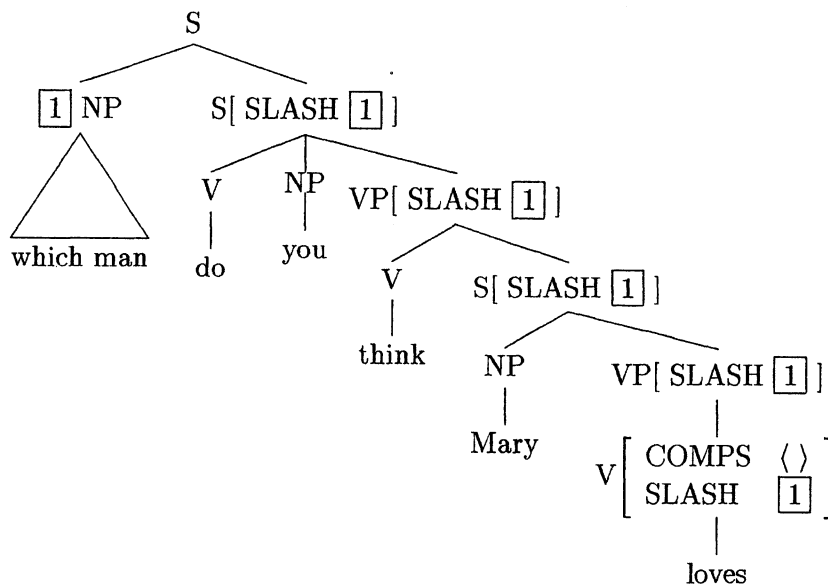
(2) Which man$_i$ do you think Mary loves e_i?

The filler is the antecedent for the trace, which is an argument of the selecting head. On the other hand, non-phrase-structure-based theories such as HPSG (Pollard and Sag 1994, Sag 1997, Bouma et al. 2001), LFG (Kaplan and Zaenen 1989), and various versions of categorial grammars (Steedman 1996, 2000; Pickering and Barry 1991, Pickering 1993) link the filler wh-phrase to the selecting head without the mediation through empty categories. For example, the gist of the GPSG representation (in a more or less HPSG style) of (2), which assumes a trace, would be something like (3a), while the gist of the HPSG representation without a trace would be something like (3b):

(3) a.

```
                    S
            ┌───────┴───────┐
          [1] NP      S[ SLASH [1] ]
          ╱────╲    ┌────┼──────┐
         ╱      ╲   V   NP   VP[ SLASH [1] ]
        ╱        ╲  │    │   ┌────┴────┐
       which man  do  you  V      S[ SLASH [1] ]
                         │    ┌────┴────┐
                       think  NP    VP[ SLASH [1] ]
                              │    ┌──────┴──────┐
                             Mary V[ COMPS ⟨ [1] ⟩]  [1] [ SLASH [1] ]
                                        │                   │
                                      loves                 e
```

b.

```
                    S
            ┌───────┴───────┐
          [1] NP      S[ SLASH [1] ]
          ╱────╲    ┌────┼──────┐
         ╱      ╲   V   NP   VP[ SLASH [1] ]
        ╱        ╲  │    │   ┌────┴────┐
       which man  do  you  V      S[ SLASH [1] ]
                         │    ┌────┴────┐
                       think  NP    VP[ SLASH [1] ]
                              │          │
                             Mary   V [ COMPS  ⟨ ⟩ ]
                                      [ SLASH   [1] ]
                                          │
                                        loves
```

Given that the tag [1] denotes an NP,

    X[SLASH [1]]

denotes an X with an NP missing. In (3a), *loves* is a transitive verb that takes an NP as its complement (indicated by the non-empty COMPS value), which is, in this case, an empty category *e*. In contrast, in (3b), *loves* is an intransitive verb (indicated by the empty COMPS value) with an NP missing (indicated by the SLASH value).[1]

Theoretical parsimony may favor one analysis over another. However, since we are dealing with human language, theoretical parsimony is not the only criterion; we ideally want an analysis with psychological reality. Generally speaking, a particular representation type (data structure) mathematically restricts the possible range of parsing mechanisms (algorithms) and vice versa. If this is also the case in dislocation constructions such as (1), findings on human language processing may decide whether or not an empty category should be posited in human linguistic representation.[2]

---

[1]Steedman's Combinatory Categorial Grammar goes further; put in the present format, *loves* is a transitive verb whose object argument slot is, to borrow Sirai and Gunji's (1998) terminology, *directly bound* by the filler *which man* without a distinguished unbounded dependency mechanism such as SLASH.

[2]One terminological note: Steedman's Combinatory Categorial Grammar explicitly rejects the "representation" status of his Surface Structure. However, his "derivation" *is* given a psychological interpretation (see Steemand

Furthermore, processing of a first language (L1) could be different from processing of a second language (L2), which could be affected either by the first language parser or the (explicit) instruction of the grammar in the classroom. The investigation of L2 processing, which is itself of some theoretical interests, could shed some light on the interaction of the representation type and the parsing mechanism in L1.

3   Problems with the Past Experiments

Despite the long debate over the role of empty categories in sentence processing (Pickering and Barry 1991, Gorrell 1993, Gibson and Hickok 1993, Pickering 1993, and Sag and Fodor 1994), researchers have not yet succeeded to settle the psychological reality issue of empty categories. In this paper, we will point out several problems in the past discussions and suggest a possible testing scheme to advance the debates over the issue. Our scheme not only contributes to settle the representational issue of L1 grammar, but also can be used to investigate L2 processing in the light of teaching effects.

According to Pickering and Barry (1991), the intuitively felt parsing ease of (4a), in contrast with the parsing difficulty of (4b), argues against positing an empty category:

(4) a. In which box did you put the very large and beautifully decorated wedding cake bought from the expensive bakery?

   b. Which box did you put the very large and beautifully decorated wedding cake brought from the expensive bakery in?

Their reasoning is something like the following.

(A) If an empty category is assumed, the filler receives its thematic role only via the empty category, which receives its thematic role from the selecting head (*put* in (2a) and *in* in (2b)).

(B) This means that the parser would have to retain the filler (wh-phrase) in working memory until encountering the empty category.

(C) Given that the object NP

   the very large and beautifully decorated wedding cake bought from the expensive bakery

   is quite complex and should put some burden on working memory, (2a–b) should both cause parsing problems.

(D) However, this is not the case.

(E) Hence, an empty category should not be assumed; the filler is directly assigned its thematic role by the selecting head.

Their argument appeals to intuition, but the thing is not that simple. (A) is indeed assumed in COMPETENCE theories such as GB/MP and GPSG, but a particular design of a competence theory does not automatically implicate a parsing mechanism. One obvious possibility is that a parser, when encountering a head, automatically projects its complements' nodes, in which case the parser will have built an empty category node in (4a) but not in (4b) immediately after encountering the verb *put* (i.e. before the complex object NP).

Such a top-down parsing strategy, or a parsing strategy in which the order of processing is different from that of derivation, needs to be assumed in GB/MP frameworks anyway. In GB/MP, derivation begins with selection of lexical items. An initial syntactic representation

---

2000) and it *is* one of its characteristics that a trace is not employed in the "derivation" of Surface Structures of such dislocation sentences. We mean the term "representation" broadly enough so that a theoretical entity counts as (part of) a "representation" when it is employed some way or other in human linguistic processing.

is built from the selected lexical items (D-structure or the set called Numeration, etc.), which is modified to another syntactic representation through various syntactic operations. Then finally a PF representation is constructed. However, if the representations constructed through derivation are psychologically real, parsing will have to begin with the final PF representation and the initial syntactic representation at the same time; the input is a sound wave, from which the parser has to cut out each word to construct a PF representation, a task possible only with simultaneous access to the lexicon.

Note also that if gap positions are posited when they are "licensed by the grammar" (the "first resort" gap-positioning principle by Gibson and Hickok 1993), the intuitive processing ease/difficulty presented in (4) can also be explained under the empty category approaches. *In* in *in which box* could be a good indicator of its thematic role; the PP filler alone will lead the parser to assume temporarily that a PP with *in* should bear a Location role to some head. In contrast, the filler in (4b) has no suggestive information concerning its thematic role. Hence, if parsing is a process of constructing predicate-argument structures from the input, (4a) should be easier to process than (4b) irrespective of whether an empty category is assumed or not, if such a default thematic role strategy is real.

Thus, the past experimental designs fail to differentiate the predictions of the empty category and the non-empty category approaches.

4 Our Proposal

To overcome this problem, we propose the following experimental design.

Construct sentences, say (5), with extraction and use them in an associative priming experiment (similar to the one by Swinney et al. 1988). The stimulus here should contain a filler whose "canonical" position is far from the selecting verb.

(5) In which box did you put [1] the very large and beautifully decorated wedding cake bought from the expensive bakery [2] last summer in England?

Under this experimental design, approaches with and without traces predict the following:

(i) If there does exist a trace in the "canonical" position [2] but the parser does not "look ahead" so as to project the trace's node immediately after encountering the selecting verb *put*, a priming effect should be observed only at the position [2].

(ii) If there does exist a trace and the parser does "look ahead" and projects the trace's node immediately after the verb, a priming effect should be observed both at [1] and at [2].

(iii) If there does not exist a trace and the filler is directly linked to the verb, a priming effect should be observed only at [1].

Our experimental design thus distinguishes the predictions by the trace and tracelesss approaches.

5 Application to L2 Teaching

Our experimental design pays attention both to the representational and the processing issues. In our view, the most fundamental flaw in the past literature is that only the representational issue was discussed without attention to the processing issue.

In addition, the issue can be indirectly investigated through a comparison of L1 and L2. (Carroll, 2002) proposes the Uniform Parsers Hypothesis as in (6) below.

(6) The Uniform Parsers Hypothesis
    Linguistic stimuli are processed by the same parsers regardless of their linguistic 'origin'. Initially, L1 parsing procedures apply automatically to L2 stimuli. (Carroll 2002: 228)

According to the Uniform Parsers Hypothesis, the input which cannot be dealt with by the L1 parser will be the input to L2 acquisition. In addition, she assumes that L2 acquisition of sentence processing is i-learning ("implicit") and explicit instructions will have no effects. The hypothesis can account for transfer errors frequently observed in the course of L2 acquisition. However, her assumptions have not yet been experimentally supported. As a second point of our paper, we propose that the same scheme we used in (5) can be used to test the hypotheses above in L2 acquisition.

Provide the same experiment in (5) to native Japanese speakers learning English as their L2, and see whether or not the priming effects at [2] (and at [1]) are observed. Our underlying assumptions are:

(i) Japanese wh-question constructions exhibit no "overt" movement, as shown in (7),

    (7) a. anata-wa Marii-ga   Ken-o   aisiteiru to   omoimasu ka
          you-TOP Mary-NOM Ken-ACC loves   COMP think   INTE
          'Do you think that Mary loves Ken?'

       b. anata-wa Marii-ga   dono-hito-o   aisiteiru to   omoimasu ka
          you-TOP Mary-NOM which man-ACC loves   COMP think   INTE
          'Which man do you think that Mary loves?'

    and hence have no empty category.[3]

(ii) Japanese speakers identify wh-questions in English with wh-question in Japanese and hence tend to employ the same processing strategy in both constructions.

(iii) In English grammar classes in Japan, wh-questions are generally taught as wh-movement from canonical position.

According to Carroll (2002), the L2 acquisition of sentence processing is implicit learning, and teaching intervention will not affect the processing procedure. If Carroll's hypothesis is correct, and if we observe the priming effect only at [1] in the experiment in (5) for native speakers, then, we should expect the priming effect only at [1] in experiments targeting Japanese speakers. In other words, we should not expect the priming effect at [2], the position where empty category is supposed to reside. On the other hand, if we find the priming effect in [2] only in the case of Japanese speakers, then we must conclude that the explicit grammar teaching had created the empty categories in L2 acquisition. Thus, our experimental design can be used to test Carroll's hypothesis as well (given that the result of native speakers support the no-empty-category approach).

6 Conclusion

As discussed in Section 3, even under the empty category (trace) approaches, the possibility that the parser "looks ahead" when it hits the selecting head is not eliminated. Because of this

---

[3]INTE here is borrowed from Gunji and Hasida (1998: 61) and denotes "interrogative," something that marks the clause as interrogative.

On the other hand, we confess that we have not carefully studied syntactic literature on the claimed Subjancency effects observed in such constructions. For example,

(i) anata-wa Ken-ga   dare-o   aisiteiru ka   siritai   no
    you-TOP Ken-NOM who-ACC loves   INTE want-to-know INTE
    'Do you want to know who Ken loves?'

Sentences such as (i) cannot be interpreted as meaning something like:

(ii) *Who do you want to know whether Ken loves?

Based on such observations, some GB theorists have claimed that such construction involve covert movement (LF movement or surface empty operator movement).

Obviously, theories with no movement operation (GPSG, LFG, HPSG, categorial grammars, etc.) will have to explain such observations in terms other than covert movement. We have not explored a syntactic analysis of such observations.

possibility, the results of past researches remain unconclusive. In this paper, we have proposed an experimental scheme which solves the problems of past researches. By incorporating Swinney et al.'s (1988) cross-modal priming method, we will be able to settle the long-lasting problem of empty category.

If the result of the experiment proposed in this paper supports traceless approaches for native speakers, the same experiment can be used to see the effects of overt teaching. Japanese native speakers are generally taught English wh-questions as wh-movement. If Japanese L2 learners of English exhibit priming effects in [2] in (5), it means that the grammatical instruction created the category (i.e., trace) which is not in native speakers' mental grammar.

References

Bouma, G., R. Malouf, and I. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory*, 19(1):1–65.

Carroll, S. E. 2002. Induction in a modular learner. *Second Language Research*, 18(3):224–249.

Chomsky, N. 1981. *Lectures on Government and Binding*. Foris, Dordrecht.

Chomsky, N. 1993. A minimalist program for linguistic theory. In K. Hale and S. J. Keyser, editors, *The View from Building 20*. MIT Press, Cambridge, MA, pages 1–52.

Gazdar, G., E. Klein, G. Pullum, and I. Sag. 1985. *Generalized Phrase Structure Grammar*. Harvard University Press, Cambridge, MA.

Gibson, E. and G. Hickok. 1993. Sentence processing with empty categories. *Language and Cognitive Processes*, 8(2):147–161.

Gorrell, P. 1993. Evaluating the direct association hypothesis: A reply to Pickering and Barry (1991). *Language and Cognitive Processes*, 8(2):129–146.

Gunji, T. and K. Hasida. 1998. Measurement and quantification. In T. Gunji and K. Hasida, editors, *Topics in Constraint-Based Grammar of Japanese*. Kluwer, Dordrecht, pages 39–79.

Kaplan, R. M. and A. Zaenen. 1989. Long-distance dependencies, constituent structure, and functional uncertainty. In M. R. Baltin and A. S. Kroch, editors, *Alternative Conceptions of Phrase Structure*. University of Chicago Press, Chicago, pages 17–42.

Pickering, M. 1993. Direct association and sentence processing: A reply to Gorrell and to Gibson and Hickok. *Language and Cognitive Processes*, 8:163–196.

Pickering, M. and G. Barry. 1991. Sentence processing without empty categories. *Language and Cognitive Processes*, 6:229–259.

Pollard, C. and I. A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. University of Chicago Press, Chicago.

Sag, I. A. 1997. English relative clause constructions. *Journal of Linguistics*, 33(2):431–484.

Sag, I. A. and J. D. Fodor. 1994. Extraction without traces. In R. Aranovich, W. Byrne, S. Preuss, and M. Senturia, editors, *Proceedings of the Thirteenth West Coast Conference on Formal Linguistics*, pages 365–384, Stanford. CSLI/SLA.

Sirai, H. and T. Gunji. 1998. Relative clauses and adnominal clauses. In T. Gunji and K. Hasida, editors, *Topics in Constraint-Based Grammar of Japanese*. Kluwer, Dordrecht, pages 17–38.

Steedman, M. 1996. *Surface Structure and Interpretation*. MIT Press, Cambridge, MA.

Steedman, M. 2000. *The Syntactic Process*. MIT Press, Cambridge, MA.

Swinney, D., M. Ford, J. Bresnan, and U. Frauenfelder. 1988. Coreference assignment during sentence processing. In M. Macken, editor, *Language Structure and Processing*. CSLI, Stanford.