Reviewing empty categories: psycholinguistic reality of NP trace and PRO gaps

Shazia Akbar Ghilzai

Abstract

Empty categories have been a focus of much research in sentence processing. Empty categories are phonetically null place holders for locally absent constituents (Chomsky 1981). Different theories of grammar assume different mechanisms for syntactic structures containing extracted arguments. GB (Government and Binding Theory, Chomsky 1981) identifies four types of empty category: NP trace, variable (=wh-trace), PRO and pro. They are characterized in terms of [+/- pronominal] and [+/- anaphoric] features. Thus, NP trace is anaphoric but not pronominal; wh-trace is neither anaphoric nor pronominal; PRO is anaphoric and pronominal; and the fourth one, pro, is pronominal but not anaphoric. The distribution of each of these empty categories is a consequence of the interaction of various principles of universal grammar. The present study discusses the two of such studies related to the theoretical distinctions between NP trace and PRO. McElree et al. (1989) examined the processing of co-referential gaps in syntactic structure as a means of determining whether linguistically defined gaps access their antecedents. Featherston et al. (2000) reported EPR evidence in favour of NP and PRO distinction.

Keywords: empty categories, NP trace, PRO-gap, co-referential gaps

Introduction: The present study deals with two types of empty Categories NP trace and PRO. PPT (principle and parameter Theory, Chomsky, 1981) distinguishes between NP trace and...
and PRO on the basis of different kinds of syntactic dependencies and movement rules. GB captures the anaphoric nature of these constructions. GB asserts that the antecedent of an obligatory PRO gap has an independent thematic role within its matrix clause. The PRO element in the infinitival clause is simply co-indexed to this role. In contrast the antecedent of an NP gap lacks an independent thematic role and receives it only on the basis of the structural position of the gap. On the other hand trace less accounts GPSG (Generalized phrase structure grammar; Gazdar, Klein, Pullum, sag, 1985) and LFG (Lexical functional grammar, Bresnan, 1978; Bresnan & Kaplan, 1982) present feature passing mechanisms without assuming movement. They suggest functional or semantic mechanisms for the reentrance of the missing argument in the gap position.

Psycholinguistic studies have addressed both the general question of the position of empty categories in processing as well as the specific case of NP trace and PRO. Some studies used Prob recognition tasks to investigate the NP-trace and PRO (McElree and Bever (1989), MacDonald (1989) and Bever and Sanz (1997) while other used priming techniques ((e.g. Swinney et al., 1988). Prob-recognition experiments have produced antecedent reactivation effects, but the results from these studies have been criticized on methodological grounds. Priming experiments have not produced effects for NP trace and PRO at gap position.

The present study discusses the two of such studies related to the theoretical distinctions between NP trace and PRO. McElree et al. (1989) examined the processing of coreferential gaps in syntactic structure as a means of determining whether linguistically defined gaps access their antecedents. Featherston et al. (2000) reported EPR evidence in favour of NP and PRO distinction.

**McElree et al. Research (1989)**

McElree et al. (1985, 1988) reported a series of experiments that investigated whether NP/PRO gaps reactivate their antecedents during processing. They used recognition prob
tasks. The sentences were presented in sections. Subjects paced themselves through the sentences by pressing the space bar after reading each section. At a particular region containing a gap a prob word drawn from the antecedent NP was presented. Subjects were required to make a decision quickly by press yes/no.

**Antecedent access following NP-Raising and PRO gaps**

**Experiment 1**

**Material & procedure**

The material for experiment 1 is shown as follows. There were two prob positions. The first prob position P1 occurred at the point where the subject could detect the presence of a gap in constituent structure. The second prob position P2 occurred 4 to 5 items after the gap, at the end of the final clause. The subjects read the constructions in sections denoted by slashes.

<table>
<thead>
<tr>
<th>Internal structure</th>
<th>Early prob (P1)</th>
<th>Late prob (P2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRO gap</td>
<td>928 (.06)</td>
<td>958 (.15)</td>
</tr>
<tr>
<td>NP-raising gap</td>
<td>873 (.04)</td>
<td>919 (.07)</td>
</tr>
<tr>
<td>NP-tough gap</td>
<td>898 (.03)</td>
<td>869 (.07)</td>
</tr>
<tr>
<td>Explicit pronoun</td>
<td>881 (.04)</td>
<td>933 (.07)</td>
</tr>
<tr>
<td>Control</td>
<td>862 (.03)</td>
<td>1.052 (.12)</td>
</tr>
</tbody>
</table>

**Results and discussion**

**Table 1: recognition latencies in (m.sec) and Error rates in proportions**

<table>
<thead>
<tr>
<th>Internal structure</th>
<th>Early prob (P1)</th>
<th>Late prob (P2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRO</td>
<td>928 (.06)</td>
<td>958 (.15)</td>
</tr>
<tr>
<td>NP-raising gap</td>
<td>873 (.04)</td>
<td>919 (.07)</td>
</tr>
<tr>
<td>NP-tough gap</td>
<td>898 (.03)</td>
<td>869 (.07)</td>
</tr>
<tr>
<td>Explicit pronoun</td>
<td>881 (.04)</td>
<td>933 (.07)</td>
</tr>
<tr>
<td>Control</td>
<td>862 (.03)</td>
<td>1.052 (.12)</td>
</tr>
</tbody>
</table>
The recognition times for all of the four types of constructions were compared against the control sentences at each of the two prob positions (Table 1). There was no significant difference at the early prob point. The early prob point was in principle the point where a gap could be detected. There was no facilitation effect at this early prob point. However, at the late prob point P2 all of the sentences showed faster recognition latencies than the control sentences. The results provide the evidence that at the end of the clause all of the constructions accessed their antecedents to a degree comparable to an explicit anaphor.

The raising and tough constructions showed greater activation than the PRO construction. Raising constructions produced faster response time (873, 898) than PRO (928). Raising constructions also showed lower error rates (.04, .03) than PRO (.06).

**Experiment 2**

In experiment 2 the PRO and control constructions were formed around adjectival rather than verb phrases, making both more directly comparable to NP-raising constructions, illustrated below:

<table>
<thead>
<tr>
<th>PRO gap</th>
<th>2a. The conceited actor/ who worked with the leading lady/ was eager [PRO] to (P1)/ rehearse for the entire evening. (P2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP-raising gap</td>
<td>2c. The conceited actor/ who worked with the leading lady/ was sure [r] to (P1)/ rehearse for the entire evening. (P2)</td>
</tr>
<tr>
<td>Adjective control</td>
<td>2e. The conceited actor/ who worked with the leading lady/ was rude to (P1)/ the rehearsers in the evening. (P2)</td>
</tr>
<tr>
<td></td>
<td>Recognition probe: conceited</td>
</tr>
</tbody>
</table>

**Results and discussion**

The Recognition latencies and error rates are presented in table 2 by prob positions. Comparison of the three constructions showed that recognition latencies differed only at the late prob position. At this late point, NP-raising constructions produced faster response time (922) than control (1,022) and PRO (986). In this experiment the PRO didn’t show the strong
evidence of antecedent activation. In both experiments 1 and 2 the PRO constructions produced small facilitation than movement gaps.

<table>
<thead>
<tr>
<th></th>
<th>Early prob (P1)</th>
<th>Late prob (P2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADJ-PRO</td>
<td>909(.09)</td>
<td>986 (.07)</td>
</tr>
<tr>
<td>NP-raising</td>
<td>882(.10)</td>
<td>922 (.0.7)</td>
</tr>
<tr>
<td>ADJ-control</td>
<td>905(.10)</td>
<td>1.022(.16)</td>
</tr>
</tbody>
</table>

Table 2. Recognition latencies in (m.sec) and Error rates in proportions

Antecedent access in the passives

GB assumes that the passive surface subject has been extracted from the underlying direct object position associated with the verb. The resultant post verbal trace is assumed to assign thematic role to the athematic subject NP. McElree et al. predicted that if the processing of NP movement gaps causes strong reactivation of moved NP, then the passive construction must also show the same pattern of antecedent activation. In order to investigate this phenomenon they conducted a third experiment.

Experiment 3

In this experiment passive constructions were compared to the active constructions. Two prob positions were used. For control constructions, rather than using the actual active version of the passive verb (e.g. had suspected), an optionally intransitive verb was used.

Passive

3a. The shrewd lawyer/ who argued for the defense/ was suspected [I] (P1)/ by the judge. (P2)

Active

3b. The shrewd lawyer/ who argued for the defense/ had spoken (P1)/ to the judge. (P2)

Recognition probe: shrewd
Results and discussion

Table 3. Recognition Latencies (in m.sec) and error rates (in proportion) ms

<table>
<thead>
<tr>
<th>Sentence Type</th>
<th>Early prob (P1)</th>
<th>Late prob (P2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passive</td>
<td>928 (.03)</td>
<td>837 (.07)</td>
</tr>
<tr>
<td>Active</td>
<td>887 (.06)</td>
<td>923 (.07)</td>
</tr>
<tr>
<td>difference</td>
<td>41 ms</td>
<td>86 ms</td>
</tr>
</tbody>
</table>

Table 3 shows that at the early prob point (P1), the active constructions produced 41 ms faster response time (887) than the passive constructions (928). But this difference was not reliable statistically. At the late prob position the difference reversed and increased in magnitude. Passive constructions produced 86 ms faster response time (837) than the active construction (923). This difference was statistically reliable. The results were similar to that observed with the NP-raising constructions in the two previous experiments. A fourth experiment was conducted to provide the further evidence for activation of the passive antecedent.

Experiment 4

In this experiment the passive constructions were compared with an adjectival form of the passive. Only one prob point

was used following the final adverb. Both constructions are illustrated in 4a. and 4b. below:

| Passive       | 4a. The ragged drifter/ traveling the land/ was resented [r]/ constantly. (P1) |
| Adjectival    | 4b. The ragged drifter/ traveling the land/ was resentful/ constantly. (P1) |
|              | Recognition probe: ragged          |
Results and discussion

Table 4. Mean reaction times in ms and error rates in proportion

<table>
<thead>
<tr>
<th>Sentence type</th>
<th>Prob Position (P1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passive</td>
<td>998 (.07)</td>
</tr>
<tr>
<td>Adjectival</td>
<td>1.115 (.08)</td>
</tr>
<tr>
<td>Difference</td>
<td>117 ms</td>
</tr>
</tbody>
</table>

Table 4 presents the mean reaction times and error rates across the two conditions. Passive constructions were 117 ms faster than the than the adjectival control. The results were similar to experiment 3. Once again, passive constructions showed greater availability of initial NP than the control constructions. The difference was statistically significant (p< .005). There was no difference in error rates.

Experiment 5

McElree et al. speculated that the previously observed differences in recognition performance can be the result of differences in the amount of time subjects spend reading the passive and adjectival phrases. They conducted the fifth experiment to remove this artifact and controlled the reading times for both passive and adjectival phrases. In experiment 5 again the passive constructions were contrasted with adjectival constructions, and adjectival constructions were used as baseline control. Again the two prob positions were used. As illustrated in 5a. and 5b.

Passive
5a. The dazed cabbie/who drove the beat-up taxi/
   was resented[t] (P1)/constantly.(P2)

Adjectival
5b. The dazed cabbie/whodrove the beat-up taxi/
   was resentful (P10/constantly.(P2)

Results and discussion

Table 5 shows that the reaction times were not significantly different at the early prob point. But like the previous experiments, passive constructions produced the faster response time
(923) than the control constructions (1,008) at the late prob position. Passive constructions were 85 ms faster than the baseline control. The error rates at both prob positions were not different significantly. McElree at al. argued that on the basis of the above finding their results with the passive cannot be attributed to differences in the amount of time that subjects spend processing the final phrase before receiving the recognition prob.

<table>
<thead>
<tr>
<th>Probe position</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence Type</td>
<td>Early (P1)</td>
<td>Late (P2)</td>
</tr>
<tr>
<td>Passive</td>
<td>901(.05)</td>
<td>932 (.09)</td>
</tr>
<tr>
<td>Adjectival</td>
<td>909 (.06)</td>
<td>1,008 (.11)</td>
</tr>
<tr>
<td>difference</td>
<td>8 ms</td>
<td>85 ms</td>
</tr>
</tbody>
</table>

Table 5. Mean reaction times in ms and error rates in proportions

**Conclusion**

Summing up the results, McElree et al. (1989) found that both NP movement gaps and PRO gaps access their antecedents during processing. They reported faster reaction times for movement gaps as compared to PRO gaps. They suggested that NP movement gaps activate their antecedents to a greater degree than PRO gaps. They supported the GB assumptions that distinguish the PRO and NP movement gaps. GB asserts that the antecedent of an obligatory PRO has an independent thematic role within its matrix clause. The PRO element in the infinitival clause is simply coindexed to this role. In contrast the antecedent of an NP gap lacks an independent thematic role and receives it only on the basis of the structural position of the gap. McElree et al. suggested that processing of a gap formed by NP movement produces more activation because it cannot be coindexed to a previously built structural representation of the antecedent but rather it has to build a structural representation for its antecedent itself.
Featherston et al. Research (2000)

**Brain potentials in the processing of complex sentences: An ERP study of raising and control conditions.**

Martial

Featherston et al. (2000) examined three types of sentences. First, sentences with subject control verbs such as *hoffen* “to hope” (a). Second, sentences with the raising verb *scheinen* “to seem” (b). Third, transitive sentences with verbs such as *erkennen* “to recognize”(c).

Constructions (a) and (b) contain an infinitival complement and has an understood covert subject. Difference between (a) and (b) is that a contains a PRO gap which doesn’t form a movement chain with the subject but might only be coindexed with the subject matrix. While (b) contains a movement trace (NP trace) denoted by ti. Which forms a chain with the subject matrix. While in (c) there is just a simple direct object but no understood subject. The adjunct clause was inserted to maximize the length of dependency.

- **a. Subject Control [PRO gap] condition**
  
  [Der Sheriff], hoffte [als die Witwe plötzlich in das Zimmer kam] [IP PRO, [NP den Täter] endlich verurteilen zu können]  
  [the sheriff], hoped [as the widow suddenly into the room came]  
  [IP PRO, [NP the offender] at last sentence to can]  
  (= The sheriff hoped, as the widow suddenly came into the room, to be able to sentence the offender at last.)

- **b. Raising [NP gap] condition**
  
  [Der Sheriff], schien [als die Witwe plötzlich in das Zimmer kam] [IP ti, [NP den Täter] endlich verurteilen zu können]  
  [the sheriff], seemed [as the widow . . . ti [NP the offender] at last sentence to can]  
  (= The sheriff seemed, as the widow suddenly came into the room, to be able to sentence the offender at last.)

- **c. Transitive [no gap] control condition**
  
  [Der Sheriff] verurteilte [als die Witwe plötzlich in das Zimmer kam] [NP den Täter] endlich im Scheinwerferlicht  
  [the sheriff] sentenced [as the widow suddenly came into the room] [NP the offender] at last in the spotlight

The sentences had a stereotyped structure as follows:

Subject NP  ➔ matrix verb  ➔ seven word adjunct  ➔ direct object NP
This direct object NP (*den Täter*) was the critical position for the ERP measurements because it immediately followed the gap positions and disambiguated the structure of the verb complement.

**Procedure**

Sentences were presented word by word on the screen. The words were shown for 300 ms duration at 500 ms intervals. There was a fixation point for the remaining 200ms. After the last word of the sentence a yes/no question was flashed on the screen. Subjects had 3500 ms to response to the question and then a 2200 pause before the next sentence.

**Results:**

*Figure 1.* illustrates the Grand average ERPs for the three conditions (i.e. transitive, subject control and raising) at all electrodes of the 10/20 system.

Figure1. Shows the Grand average ERP to the critical noun phrase. The ERPs were time locked to the onset of the critical noun phrase and contained the responses to the two following words. All three conditions gave rise to different ERP patterns, starting at around 250 ms and showed maximal effects at central and parietal electrodes.
About 300 ms after the onset of the article of the critical noun phrase, the ERP’s to the subject control and raising conditions are showing more positive ongoing ERP, than the transitive condition. While, the raising condition is showing a more positive-going ERP than the subject control condition.

Figure 1 also shows LAN effects. LAN is an early negativity in 200-400 ms time range. It is typically left-Lateralized with a fronto temporal distribution. There was no difference among the three conditions at F7 electrode. However, a comparison of left hemisphere site F7 with right hemisphere site F8 shows a more negative ERP over the left hemisphere. Which suggests that a LAN was present in all three conditions. This effect can be observed more clearly in figure 3.

a) Subject control sentences show a small difference then the transitive sentences (300-500 ms)
b) Raising sentences were more positive than transitive
c) Raising sentences were more positive than subject control conditions

Figure 2. illustrates the pairwise comparison of the three conditions for the Cz electrode.

Figure 2. shows Pairwise comparison of the three conditions. The transitive and subject control conditions ERP wave forms are compared in Fig 2(a). The two conditions diverge at about 300 ms where the subject control conditions were more positive than the transitive conditions. The effect lasts until 600 ms. In fig 2(b) ERPs from raising and transitive
conditions are compared. In 400-600 ms time window, ERPs to the raising condition were more positive. There was also a significant difference between these two conditions in 800-1000 ms and 1000-1200 ms time windows. Fig 2 (c) shows the comparison of raising and control conditions. Raising condition produced a more positive going ERP waveform after 500 ms than the subject control condition. The difference was more significant in the subsequent 600-800 ms, 800-100 ms and 1000-1200 ms time windows. In all these three time windows raising condition were more positive than subject control conditions.

![Figure 3](image)

**Figure 3.** Shows a comparison of left and right hemisphere anterior temporal sites for the article of the critical noun phrase and a control article. Only the critical article shows a left anterior negativity for all three conditions.

In order to observe the LAN effect (fig.3) more clearly the left and right anterior temporal electrodes were overlaid for the article of the critical noun phrase (*Der Sheriff hoffte* “the sheriff hoped”) and the first article of the subordinate clause (*als die Witwe* “as the widow). The difference between the left and right hemisphere site at F7 and F8 electrodes was only found for the critical article but not for the control article. Between 200-400 ms time window LAN was observed for the critical article in all three conditions. The difference was statistically significant (p < 0.001).
Conclusion
Summing up the results of the ERP studies, Featherston et al. (2000) found that raising constructions are difficult to process than control constructions. They found the P600 immediately after the trace position in the raising construction. NP traces elicited P600 of higher amplitude than PRO. Amplitude of P600 reflects cost of syntactic processing. With higher processing cost correlated with higher amplitude (Osterhout, 1994: 18). Therefore, raising construction elicited stronger P600 because it requires extra computational operation due to the movement operations.
In addition to P600 they also found LAN (Left anterior negativity) but unlike the P600, LAN doesn’t differentiate between the three sentences. They argued that this is because of the intervening adjunct clause that was inserted in to separate the subject and the critical object NP. When the parser assigns a thematic role to the critical object NP (e.g. Täter), it also accesses the subject which puts a stress on the working memory capacity because the system access the subject across the intermediate material. As the adjunct clause was present in all of the three sentences, that’s why LAN was observed in all of three conditions. They suggested that LAN reflect the working memory operations.

Discussion
On the basis of their results from behavirol data McElree et al. suggested that PRO is difficult to process than NP movement because they found faster response time and lower error rates for processing NP constructions than PRO. While the Featherston et al. presented opposite results for the distinction between NP trace and PRO. They proposed that NP traces are difficult to process than PRO. They reported stronger P600 for NP movement gaps than PRO.
Which one result one should believe to be correct? Let’s have a critical look at both of them separately.
McElree et al. (1985, 1988) reported that PRO gaps produced a smaller amount of facilitation than movement gaps. They related these differences to the movement rules asserted by GB, which assumes a distinction between NP and PRO gaps. The reaction times and error rates reported by McElree et al. show that there were faster response times and lower error rates for NP constructions than PRO constructions. The results suggest that movement gaps produced greater facilitation and hence showed greater activation of antecedents than PRO gaps. But the difference between these two constructions was only significant at the late probe position P2. The late probe position was the end of the sentence. The claim was made about the gap positions but the late probe position P2 was not the place where the gaps have been postulated or movement of particular constraints has been assumed.

There was no significant difference between these two constructions at the early probe point. The early probe point was in fact the point at which a gap could in principle be detected. McElree et al. attributed the absence of facilitation at the early point to the simple delay on the part of the subjects. They argued that readers do not immediately access the antecedent after processing the region containing the gap. They either delayed in postulating a gap or assigning a filler until subsequent input confirmed the validity of the analysis. However, speculating the sensitivity of their prob task they performed a series of experiments. But in all of the 5 experiments there was no significant facilitation at Prob P1 except the experiment 4. In experiment 4 there was only one Prob point P1. Where the passive showed the greater facilitation than the adjectival control. On average passive was 117 ms faster than adjectival control. But looking at the material for experiment 4, Prob position P1 was again the end of the sentence.
In short there were no differences in reaction times/ error rates at the actual prob positions where the gaps (PRO or movement gaps) are in principle assumed. Therefore, the results from the prob recognition tasks reported by McElree et al. are not reliable.

Featherston et al. (2000) used ERP technique and found larger P600 for raising constructions (containing NP trace) than the Control (constructions containing PRO gaps). They argued that the P600 they found in the raising constructions after the trace position is related to syntactic differences. In a raising construction the parser reconstitutes and copies the overt subject when the gap is encountered, setting up the filler dependency as movement chain. This extra computational operation is not required in the control sentences. Because of this difference NP trace elicited a P600 of higher amplitude than PRO. Unlike the McElree et al. they found the difference between the between the raising and control sentences at the position immediately following the gap. MEcElree et al. self paced prob recognition tasks were more an offline experiment than an online. While Featherston online experiment was a carefully designed online experiment with specific time durations for response.

Featherston et al. suggested that raising is harder to process because in a raising construction the parser reconstitutes and copies the overt subject when the gap is encountered, setting up the filler dependency as movement chain. There seems a contradiction here, because it is customarily assumed that semantic representations ought to “limp slightly behind syntactic ones” ( Fiebach, Schlesewsky, and Friederici 2001: 328: see Fodor 1983). How is it possible
for control, a semantic issue involving coreference, to be easier and faster to process than a semantically governed raising?

In short, neither the Meclree et al. (1989) results nor the featherston et al. (2000) results are decisive experimental evidence for NP trace and PRO. Further experimentation is therefore, required.

References


