UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Pragmatic effects in zero anapho r resolution: Implications for modularity.

Permalink

https://escholarship.org/uc/item/5105g5r0

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 17(0)

Authors

Durrant-Peatfield, Mark

Marslen-Wilson, William

Publication Date

1995

Peer reviewed

Pragmatic effects in zero anaphor resolution: Implications for modularity.

Mark Durrant-Peatfield[1] and **William Marslen-Wilson**[2]
Centre for Speech and Language
Department of Psychology
Birkbeck College, University of London
Malet Street, London WC1E 7HX
`ubjtp80@uk.ac.ucl`[1]
`ubjta38@uk.ac.bbk`[2]

Abstract

Fodor (1983) has claimed that informational encapsulation of the parser is the way the language system prevents extra-linguistic factors from slowing down first pass processing. However, in a naming task where the visual probe was an appropriate or inappropriate pronoun continuation to a gerundive phrase following passages in which discourse focus and verb semantics were co-varied (Marslen-Wilson, Tyler & Koster, 1993) we found appropriateness effects which suggest a role for on-line pragmatic inference in top down control of the parser. Fodor, Garret & Swinney (1993) maintain that, though the gerund is marked as requiring a subject NP, the inferential activity underlying referent assignment does not occur until an explicit anaphor (the pronoun target) is encountered. As modularity predicts a cost associated with contacting real world information, assignment times to gerunds should take longer than assignments based on lexical information. A speeded fragment completion task was used to counter Fodor's objection to a pronoun probe and to detect differences in the times taken to make anaphor assignments. The two studies reported here used the original Marslen-Wilson *et al.* (1993) materials. Correct assignments in the gerundive condition ("Running towards...") were cost free with the exception of the condition where the pragmatically most likely subject was not in discourse focus. Latencies to initiate a completion were otherwise similar regardless of whether the to-be-completed fragment contained a gerund or a disambiguating pronoun. Furthermore, in the absence of pragmatic constraints, assignment always favoured the highlighted entity. These results reproduce the critical data from the Marslen-Wilson *et al.* (1993) study which demonstrates context effects on first pass processing.

Introduction

For nearly two decades, several researchers have been concerned to show that there is a syntactic module that operates independently of processes that access general world knowledge. Fodor (1983) has been the most influential advocate of this position. He argued that language interpretation could not happen so quickly if it incorporated constraints based on pragmatic inference. The Modularity hypothesis proposes that human syntactic processing takes place in cognitively impenetrable, input systems which are informationally isolated from non-modular, highly penetrable central processes. The boundary between module and central processes falls approximately at the traditional boundary between semantics and pragmatics: between logical form and a mental model of the current discourse. Informational encapsulation of the parser is the way the system prevents non-linguistic factors from slowing first pass processing.

Two testable claims generated from this approach concern the notion of encapsulation and processing speed. The first claim is that the parser is not sensitive to instructional top-down effects; it should never be controlled in first-pass processing by constraints derived from pragmatic inference. The second claim follows from the implications of the first. Because non-linguistic factors do not slow down first pass processing, language processing activity based on automatised non-inferential module-internal processes should be faster than activity based on pragmatic inference.

These are strong and important claims which have stimulated a great deal of research for and against the modularity assumptions (see Altman & Steedman, 1988, Marslen-Wilson & Tyler, (1987) for context effects on parsing and Mitchell, 1987, for contrary evidence). However, Fodor (Fodor, Garret & Swinney, 1992) has singled out one study in particular as providing "the most striking current evidence for an interactionist view of parsing", where extra-linguistic inferences control syntactic operations . This was a study reported in detail in Marslen-Wilson, Tyler & Koster (1993), which used cross-modal naming as an index of anaphor resolution. Subjects listened to three context-setting sentences containing a single male and female protagonist as in the passage reproduced below.

1. Mary lost hope of winning the race to the ocean when she heard Andrew's footsteps approaching her from behind. The deep sand was slowing her down. She had trouble keeping her balance.

The passage was followed by a verb fragment initiated either with a personal pronoun (He overtook..) or by a gerund (Overtaking..). At the offset of this fragment a pronoun probe was visually presented which the subject had to say aloud as quickly as possible. This probe (e.g. HIM or HER) was either consistent or inconsistent with the contextual interpretation of the fragment. Any advantage in terms of naming latency for contextually appropriate probes could only arise if the subject of the clause was established at that point and the anaphor resolved. Two other conditions varied the relationship between the discourse highlighted entity and verb bias. The verb could be either neutral or incongruent with respect to the context and targeted antecedent.

In all gerundive phrase conditions, the appropriate probes were named faster than inappropriate probes. Particularly striking was the appropriateness effect in the condition where the non-focused entity was the most likely subject of the gerund (as in the example above). Pragmatic inferencing seems just as effective as an apparently more direct mapping

between an antecedent and an unambiguous pronoun. Pragmatic information instructs the parser which entity to use as subject and which to use as object. Another condition gave an appropriateness effect when no pragmatic constraints existed. Discourse focus alone is able to determine subjecthood.

Acknowledging the theoretical importance of these results, Fodor, Garret & Swinney (1992) launched a counter attack, arguing that processes concerned with establishing pragmatic likelihood are in fact not set in motion until an overt anaphoric device such as a personal pronoun is encountered. The sentence is parsed and the gerund is marked as requiring a subject, but nothing more happens until an overt anaphor is encountered. Only then are central processes consulted to bind the gerunds to their most likely antecedents. Fodor *et al.* maintained that the visual pronoun probe used in the Marslen-Wilson *et al.* (1993) naming task acted as a trigger to initiate contextual processing. Resolution of a gerund therefore has to wait for a trigger. It does not occur on-line. To support this it was necessary to show that, given the appropriate task, there will be no evidence for on-line integration of verb semantics with discourse context to determine the missing subject of the gerundive.

Fodor *et al.* (1992) used a cross modal priming task which enabled them to probe immediately after the critical gerundive phrase rather than as in the Marslen-Wilson *et al.* (1993) study, further downstream at the object pronoun. Subjects listened to passages like (2) below and named visual probes (girl/fireman) presented at key points (indicated by asterisks) in the discourse. The probes had appeared previously in the passage. The rationale here was that, if part of the discourse referred back to an antecedent, then the reactivated information would be available to mediate repetition priming.

2. The fireman carried a rope towards the little girl who was trapped on the ledge, immobilised with fear. He moved very carefully*. Crawling cautiously* towards her* so that he could hand her the rope he managed to save her.

There was no apparent reactivation of the antecedent immediately after the gerund. Naming "fireman" after hearing "crawling cautiously" was not facilitated. In contrast, there was some evidence for reactivation at the later position after the object pronoun. Fodor *et al.* (1992) conclude that inferential processes concerned with anaphor resolution do not occur during or immediately after the gerund.

There are two problems with this conclusion. First, the effects are weak and of uncertain statistical significance. Second, there are problems in the interpretation of the repetition priming task used here. Although the task allows the experimenter to probe during the verb phrase, it does not tell one directly how or whether the processor has assigned subjecthood, but only that certain discourse entities are in some way more or less active at different probe points.

A better task would be one that probed as soon as possible after the verb, did not use a pronoun probe and, most importantly, gave unambiguous evidence about the assignment (or not) of subjecthood. The task should also operate within the time frame of an on-line task. We decided on a speeded fragment completion task. This fulfills the criteria and, in particular, it counters Fodor *et al.* (1992) main objection to the original study by not providing an overt pronoun probe.

This is a task in which subjects hear a short auditory passage followed by a sentence fragment. On the presentation of a neutral visual probe (a row of asterisks), subjects are required to continue or complete the fragment as rapidly as possible, saying aloud their continuation.

The time taken to initiate an appropriate response is taken as a measure of the status and efficiency of the inferential processes leading up to the assignment. Note that the task differs here from the cross modal naming paradigms adopted by both Fodor *et al.* (1992) and Marslen-Wilson *et al.* (1993). Contextual processing is functionally required to perform the task whereas there is no intrinsic requirement for contextual processing in order to perform naming.

Fodor's thesis suggests that the pronoun at the beginning of the pronoun fragment triggers contextual processing. The agent of the verb is likely to have been determined before presentation of the probe. For the gerund fragment, where there is no pronoun probe to trigger contextual processing, the inferential activity required to establish the appropriate antecedent in the gerund fragments can only occur on presentation of the neutral visual cue. If there are computational costs associated with contacting real world knowledge (which is an explicit assumption in modularity), then initiation latencies to gerund fragments should be longer, across the board, than initiation latencies to pronoun fragments.

The alternative view supported by the Marslen-Wilson *et al.* (1993) study permits contextual processing to occur as the fragment is being heard. Subject assignment to the gerund is likely to have been completed by the time the probe is presented. Latencies to continue each type of fragment would be expected to be similar and we would be justified in concluding that the processes supporting referent assignment to the gerund occurred "on-line". The task is therefore able to shed light on the speed of discourse mapping and pragmatic inference simultaneously. It offers the chance to distinguish between gerund and pronoun fragments in terms of processing times.

We used this task in a new experiment based on the stimuli and design used in the Marslen-Wilson *et al.* (1993) naming study. There were three conditions in which pragmatic bias and discourse focus were varied. The fragment was initiated either with a gerundive or a pronoun and was phrased in such a way that a personal pronoun was an appropriate continuation given the context. Stimuli representative of the three conditions are given below.

1. Discourse focus with congruent verb bias.

After the surgeon had examined the 12 year old girl with the badly broken leg, he decided that he would have to take immediate action. He'd had a lot of experience with serious injuries. He knew what he had to do next. He injected../Injecting..

In condition 1 the entity focused or highlighted in the discourse is also the most pragmatically likely subject of the gerundive phrase. Off-line written completions to this passage typically place the surgeon in subject position and the girl in object position. The entity focused in the discourse (the surgeon) is also most likely to be the one injecting. The gerund "injecting" is therefore pragmatically congruent with

the entity in focus. Of the two possible pronoun responses, "her" would be the most appropriate.

Fodor's modified thesis suggests that the pronoun "he" in the pronoun fragment triggers contextual processing. The parser is permitted a glimpse of the discourse representation containing two entities, one of which is in discourse focus. With the gerund fragment, the parser is not permitted access to the discourse representation. The advantage to the pronoun fragment is therefore two-fold. First, the inferential activity underlying assignment activity can proceed on the basis of lexically specified information which rules out a parse like "He injected him". Second, the fact that the male protagonist is in discourse focus suggests that "he" will remain in subject position in a future utterance. Assignment times in an on-line scenario are therefore likely to favour pronoun fragments compared to gerund fragments.

The alternative view permits the assignment processes unrestricted access to the discourse representation and real world knowledge. The fragment is interpreted in light of the preceding context and an assignment is made as soon as possible. There is research indicating that an entity in discourse focus enjoys a unique processing status (Garrod & Sanford, 1990) so it is possible that assignment of subjecthood to the gerund would further be facilitated. Assignment times in the SFC task are likely to be similar for gerunds and pronouns in this condition.

2. **Discourse bias with neutral verb bias.** As Bill was buying candy at the cinema, he saw an old girlfriend get in line for a ticket. He had arrived at the cinema especially early. He had wanted to be sure of getting a good seat. He waved at../Waving at..

In condition 2 either Bill, the entity in discourse focus, or the girl friend are equally likely to be waving. There is no pragmatic bias towards either entity. However, in off-line pretests, Bill and not the girlfriend was made the subject of waving. This preference for the focused entity was also evident in the Marslen-Wilson *et al.* (1993) naming study. However, because a modular view of parsing prohibits access to the structure of the discourse representation, there is no reason for the parser to choose one discourse entity in preference to the other. The number of assignments favouring the male protagonist following the gerund fragment would be expected to be matched equally with the number of assignments in favour of the female protagonist. In contrast, the alternative view which allows assignment processes to be influenced by contextual influences would predict that assignments favouring the discourse entity currently in focus would predominate because, as noted above, the focused entity is likely to remain the subject of the discourse. With regard to the response latencies, initiation of completion times for pronoun fragments will, as above be expected to be faster than for gerund fragments. Once again, the alternative view suggests that the structure of the discourse representation is accessible right from the start for both types of fragment. Fragment initiation times should be similar.

3. **Discourse bias with opposing verb bias.**
Mary lost hope of winning the race to the ocean when she heard Andrew's swift footsteps approaching her from behind. The deep sand was slowing her down. She had trouble keeping her balance. He overtook../Overtaking..

In condition 3, the less pragmatically likely subject of the gerund is in discourse focus. Given what we have been told about the situation, Mary is not in a position to be overtaking Andrew. Off-line written completions established Andrew as the most likely entity to be overtaking. Since the modular parser is not affected by the structure of the discourse representation or pragmatic constraints, the condition 3 stimuli should be treated in the same manner as stimuli in the other two conditions. However, this condition presents something of a problem to an unencapsulated language processor. A shift in discourse focus is generally signalled in some manner in normal discourse. This condition however makes the entity not focused in the discourse the most likely subject of the gerund without signalling the change in focus. Verb bias is being pitted against discourse focus. Though the Marslen-Wilson *et al.* (1993) study gave appropriateness effects for this condition, it is not clear how this will affect gerund fragment initiation latencies.

Experiment 1

The exact procedure was as follows. A subject heard a short passage comprising three context setting sentences highlighting one of two discourse entities. The passage was followed by a continuation fragment beginning either with an overt personal pronoun or a gerund. A neutral visual cue (comprising a row of asterisks) presented at the offset of the fragment prompted the subject for a voiced continuation. Time taken to initiate a response was recorded as well the actual response. The inter-trial interval (ITI) was 2000ms.

For the timing of these assignment processes to fit within an "on-line" time scale, there should be no costs suggestive of delay or reduced efficiency in the inferential activity supporting referent assignment. Initiation latencies should not differ across anaphor conditions (pronoun vs gerund). If does capture the products of on-line processing, we predict that the same pattern of results seen in the original Marslen-Wilson *et al.* (1993) naming paradigm would be seen here. i) the pragmatic bias exerted by the verb in the condition where the focused entity is the least likely subject of the gerund, should ensure that the non-focused entity becomes the subject. ii) assignment of an antecedent to a gerund in a context where either entity could be a pragmatically acceptable subject, should favour the discourse highlighted entity.

Condition	Gerund	Pronoun
1	99	100
2	99	100
3	80	100

Table 1: Experiment 1: percentage distribution of correct assignments.

First we look at assignment patterns (Table 1) for 34 subjects tested on 36 test items (12 per condition). In conditions 1 and 3 an incorrect assignment is one which does not make sense given what we know about the situation. In condition 2 an incorrect assignment is one which makes the non discourse focused entity the subject of the gerund. Incorrect assignments were very infrequent across the board for conditions 1 and 2 but accuracy drops off in condition 3 where

there is a conflict between discourse bias and verb bias. In Condition 2 where the verb bias favours either discourse entity, subjecthood is almost always assigned to the discourse focused entity. The Marslen-Wilson *et al.* (1993) study also found significant appropriateness effects for condition 2. So far, these results mirror the original Marslen-Wilson *et al.* (1993) study. Both discourse focus and verb pragmatics are sufficient to link an utterance to the discourse context in isolation. However, things are not so clear cut when initiation latencies are considered.

An analysis of variance (ANOVA) by items (F_2) on the response times showed a significant interaction between fragment type and condition, ($F_2[2,33] = 3.68, p<0.05$). A test for simple effects tied down variance due to fragment type solely to condition 3. ($F_2[1,33] = 17.72$, $p<0.005$). There were no other significant effects implying that latencies to initiate a continuation to pronoun fragments are similar to latencies to initiate continuations to gerund fragments in conditions 1, and 2.

Condition	Gerund	Pronoun	Diff
1	1119	1103	+16
2	1115	1065	+50
3	1178	1041	+137*

Table 2: Experiment 1: item means per condition for all correct responses initiated with a personal pronoun. "*" denotes p<0.005

Discussion

The data for conditions 1 & 2 challenge a central claim modularity makes. Responses involving inferential activity based on real world knowledge should be slower and less efficient than responses based on look-up information. However, the interpretation here is complicated by the finding that there was a difference in condition 3, where responses were slower following a gerundive fragment than following a pronoun fragment. Does this cost show that inferential activity based on real world knowledge is indeed slower and less efficient as predicted by modularity? We think not.

It is not that pragmatics is being used in condition 3 and not the others. Rather, the subject has to switch from the focused entity to the alternative entity not in focus (e.g. in the example 1 given above, the listener has to switch from Mary to Andrew). Pragmatic evaluation is immediate and will be applied first to the most discourse salient candidate for subjecthood. The effect in condition 3 is in fact evidence that in conditions 1 and 2 pragmatic evaluation is also going on. If pragmatic evaluation was not proceeding before or at the cue points in all three conditions then there would be no effect in condition 3. The language processor could not identify the unsuitability of the highlighted entity until the pragmatic implications of the assignment have been realised. When confronted by the gerundive fragment in condition 3 the language processor assumes the subject of the gerundive will be the subject of the prior discourse. Hobbs (1979) has noted that 92% of the time this is an accurate heuristic. The use here of this heuristic is supported by the comparatively large number of mistaken assignments in favour of the focused

entity (20%) in condition 3. The longer latencies here reflect the time taken to discard the entity suggested by discourse focus and consider the alternative one.

In summary, the data did not distinguish between the 2 types of fragment in terms of initiation latencies. With the exception of condition 3 there was no overall difference between initiation latencies for pronoun and gerund fragments. Is this because each can be related to the current discourse equally rapidly or do the long latencies conceal possible differences in timing? If so then what we are seeing might have little to do with inferring referent assignment and much to do with those processes concerned with preparing an utterance. If long latencies conceal real differences between pragmatic and non-pragmatic assignment activity, then speeding up the responses should increase the differences between continuation times for gerund and pronoun fragments. This should be especially noticeable in condition 3.

The experiment was repeated but with modifications to make the initiation of continuation latencies much shorter. Subjects were instructed to say as quickly as possible just the first word that occurred to them on seeing the visual cue. They were advised that the word should follow on naturally from the preceding context and the fragment.

Experiment 2

The assignment patterns (Table 3) from 36 subjects tested on the 36 test items shows that the two critical effects observed in the last study were evident also in this study. In Condition 2 where the verb bias favours either discourse entity, subjecthood is assigned to the focused entity. In Condition 3 where the verb bias favours the non-highlighted entity, 74% of the assignments were made towards the pragmatically most likely entity.

Condition	Gerund	Pronoun
1	95	100
2	92	100
3	74	100

Table 3: Experiment 2: percentage distribution of correct assignments.

Turning to the reaction time data (Table 4), there was no significant difference between gerund and pronoun fragment initiation latencies. However, a significant interaction between fragment type and condition ($F_2[2,32]=3.14$, $p<0.05$), was seen in condition 3 ($F_2[1,32]= 5.27$, $p<0.05$) using a comparison of simple means. Latencies to supply an utterance to a gerundive phrase in condition 3 were longer than for utterances based on lexically specified information.

Condition	Gerund	Pronoun	Diff
1	694	710	-16
2	703	703	0
3	707	674	+33*

Table 4: Experiment 2: item means per condition for all correct responses initiated with a personal pronoun. "*" denotes p<0.003

367

Discussion

Thus, despite much quicker latencies in Experiment 2, the same pattern of results was found as in Experiment 1. There are still no significant differences between initiation latencies for pronoun and gerund fragments. The latencies in Experiment 2, compare favorably with naming or other on-line tasks and it was clear from subject reports that they could not perform any faster.

Finally and inconsistent with the claim that the long latencies in Experiment 1 conceal actual differences between the gerund and pronoun conditions, the differences were now much smaller. Although there is still a significant difference between the pronoun and gerundive condition in condition 3, the cost has been reduced to 33 msecs. This change is most probably a reflection for the time required to prepare an utterance. In the first task, subjects had to complete the fragment whereas in the second task, subjects had only to provide a single word continuation.

General Discussion

Initiation latencies are rapid enough to pick up the products of on-line processing. If there are significant differences in processing times for pronouns and gerunds then surely they would have shown up here. There are two possible scenarios which account for these data. Either the inferential activity concerned with gerundive assignment occurred as the fragment was being heard or, assignment activity occurred immediately on presentation of the neutral visual cue. The latter case suggests that, all things being equal, resolving gerundive fragments occurs as rapidly as pronoun fragments even when contextual processing is required. The former admits the possibility that contextual information is being contacted as the fragment is being parsed. Either of these possibilities is inconsistent with the modular framework.

The other results replicate those observed in the Marslen-Wilson *et al.* (1993) study which demonstrated context effects on syntactic (intra-modular) representations and enable us to reject the kind of counter attack made by Fodor and others. There are no linguistic cues in the gerundive phrases, yet agency is assigned on the basis of pragmatic inferential activity in condition 3, or in condition 2, on the basis of discourse focus alone.

There is the question of whether and how the pronoun probe used in the Marslen-Wilson *et al.* (1993) study differs from the neutral probe used in the current experiments. Both probe types prompt a response from the subject but the similarities end there. An overt personal pronoun is bound to have more significance to a modular language processor than a row of asterisks. Fodor *et al.* (1992)'s modified thesis makes the appearance of an overt pronoun a necessary part of first pass processing. The instruction to consider the constituents and structure of the discourse representation on encountering an overt anaphor is presumably built into the language processor. A modular parser will act only on elements from its grammar. It is likely then that only an overt pronoun will suffice to initiate the instruction. A modular parser expecting but denied an overt pronoun as in the SFC task would have to wait until central processes took over and tried to make some sense out of the discourse representation. The unlikely alternative is to allow the parser to be sensitive to neutral probes as well as overt pronouns. Clearly then, the neutral probe cannot be considered equivalent to the pronoun in its effect on the language processor.

Marslen-Wilson *et al.* (1993) argue that if there is any informational encapsulation, it is not in a module that has logical form as its output, but rather one that outputs a mental model or some other form of discourse representation. Such output requires background knowledge for its construction and if the output is to be "on-line" then the knowledge has to be accessed almost immediately. The experimental evidence suggests we abandon the strictly modular view, and assume that, as listeners proceed in the sentence, they attempt to construct a discourse interpretational structure directly. There remains the issue which was the original motivation for modularity. How can inferential activity based on real world information occur so rapidly? It is remarkable given the unconstrained nature of the data-base which must support such activity that there does not yet appear a processing penalty in using it. There might be constraints operating to make language processing as rapid as possible but there does not appear to be a constraint on the type of information available to a putative parser in the first few hundred milliseconds of processing.

Acknowledgements

References

Altmann, G. T. M. & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, 30, 191-238.

Fodor, J. A. (1983). Modularity of mind. Cambridge, Ma: MIT Press.

Fodor, J. A., Garrett, M., & Swinney, D. (1993). A modular effect of parsing. Manuscript, Program in Philosophy, CUNY Graduate Center.

Garrod, S.,& Sanford, A. (1990). Referential processes in reading: Focusing on roles and individuals. In Comprehension processes in reading. Eds Balota, LEA.

Hobbs, J. R. (1979). Coherence and coreference. *Cognitive Science*, 3(1), 67-90.

Marslen-Wilson, W., & Tyler, L. K. (1987). Against Modularity. In J. L. Garfield (Eds.), *Modularity in Knowledge Representation and Natural Understanding*. Cambridge: MIT Press.

Marslen-Wilson, W., Tyler, L. K., & Koster, C. (1993). Integrative processes in utterance resolution. *Journal of Memory and Language.*, 32, 647-666.

Mitchell, D.C. (1987). Lexical guidance in human parsing: Laws and processing characteristics. In M. Coltheart (Ed.), *Attention and Performance XII: The psychology of reading.* London:Lawrence Erlbaum Associates Ltd.