

# Some Good Reasons for Shallow Pronoun Processing

Peter Bosch  
IBM Germany, Science and Technology – LILOG,  
P.O. Box 800880, D-7000 Stuttgart 8, Germany  
vnet: pbosch@stutvml

## Abstract

In view of what is currently known about the number and complexity of parameters that are involved in anaphoric pronoun resolution the success of purely syntax based algorithms for pronoun resolution is highly surprising. This paper suggests a linguistic, and in part psychological, explanation for why these algorithms work so well. It also proposes a distinction between applications for which syntactic algorithms of pronoun resolution are fully reliable and applications for which they only yield probabilistic results.

## 1. The Problem

Success rates of algorithms for pronoun resolution seem to demonstrate the overall futility of close linguistic analysis for efficient machine processing of natural language. Very simple algorithms that take as their input no more than a parse tree easily yield a success rate of well over 80% (cf. [7]). More sophisticated algorithms that either pay additional attention to some notion of salience (in the form of a syntactically based salience scale as in [14] or in the form of discourse focus [10]) or to some form of semantics [13] would seem to show that the increase in success rate almost seems to decrease in proportion with the additional sophistication of the analysis.

If our concern is the study of human cognition then the practical question as to whether it is really worth all the extra effort, just in order to achieve an only slightly better success rate does not arise. The aim then is not a success rate of one hundred percent, but what we need is something even harder to achieve: the ideal algorithm for pronoun resolution should make the same kind of mistakes as the human processor. But this means that it should employ the same pronoun resolution strategies as the human processor.

In this paper I want to look at just one, hypothetical, human strategy for pronoun resolution that corresponds surprisingly closely to syntax-based algorithms in AI, in particular to Hobbs' [7]. The aim behind taking a closer look at syntactic processing is not to argue against semantically based and knowledge based pronoun resolution. Theoretical motivations apart, the aim is rather to find out about the extent to which 'heavy' processing is avoidable in practice without making the results unreliable. This question is of particular interest in view of the perspective of having language understanding systems acquire the bulk of their world knowledge and lexical knowledge themselves on the basis of a text input.

Supposing that humans employ several different strategies for pronoun resolution, depending on the kind and quality of the input in relation to the amount and structure of relevant knowledge they have at their disposal, we may ask for each strategy what type of pronoun

occurrence it is best suited for. “Type of pronoun occurrence for which a particular strategy is best suited” here means: the type of pronoun occurrence for which the strategy in question will ordinarily yield correct results with less effort than alternative strategies, given the resources that are ordinarily available with a pronoun of that type.

For a syntax-based strategy the relevant type of pronoun occurrence should be one that can regularly and successfully be processed with no appeal to other than syntactic information. In Section 2 I shall characterize such a type of pronoun occurrence.

## 2. Syntactic vs referential processing of pronouns

Earlier research [1] has shown that there is an important distinction that is needed in order to account for a variety of problems in the interpretation of anaphoric personal pronouns: a distinction between pronouns that function purely syntactically and pronouns that function referentially. I have called the two classes *syntactic* and *referential pronouns* respectively (for short, SPs and RPs).

Since SPs occur non-referentially, i.e. with no connection to any domain of reference, all that can matter to their interpretation is their structural relation to their controller. In this respect they behave like agreement morphemes that mark subject agreement in the finite verb. There is no issue as to whether the third person *s* in a sentence like *x leaps* refers to anybody or anything and if so to whom or what. Its function is to mark the fact that *leap* is interpreted as a function that takes *x* as its argument, no matter what you fill in for “*x*”. Likewise, the referential status of an SP-controller is immaterial to the control relation. Hence such relations remain undisturbed by all syntactically appropriate substitutions for the antecedent.

This is different for RPs. They need a referent in order to be interpreted, just like any other referentially occurring NP. If, in a particular syntactic structure, this referent is introduced by means of another definite NP and the RP occurs co-referentially with that NP, then the interpretation of the RP depends on the reference of the NP. Thus if we substitute a non-referential NP for the original NP to which a pronoun is anaphorically related via co-reference, the anaphoric relation must break down. Cf. the following cases for illustration. If we take (1a) in the reading on which *he* is anaphorically related to *Fred*, we can see, after the substitution of the non-referential *nobody* for *Fred* in (1b), that the anaphoric relation between the pronoun and the subject is maintained. Hence also in (1a) the anaphoric relation cannot depend on co-reference and does not require a referential occurrence of *he*. Cf.

(1a) Fred said he was sick.

(1b) Nobody said he was sick.

For (2a), on the reading that links *he* to *Fred*, we get a different result for the substitution of a non-referential NP for the antecedent: in (2b) the anaphoric relation between pronoun and preceding NP has broken down. So we may conclude that, other things being equal, the reference of the antecedent plays a crucial role for the anaphoric relation in (2a), i.e. that the pronoun occurs referentially as an RP and co-refers with the antecedent.

(2a) When Fred arrived, he was tired.

(2b) When nobody arrived, he was tired.

These considerations, however, show nothing about properties of occurrences of anaphoric pronouns or about particular pronoun forms. They are strictly concerned with the distinction between the *strategies a linguistic processor employs*: Thus if a pronoun occurrence is tested

with positive result for SP-hood (either by a substitution test based on observations as in (1) and (2), or by means of other procedures (cf. [1]), then we know that it may successfully be processed by a strategy that makes no assumptions with respect to reference. And if the tests show that the pronoun is an RP we know that it can be processed by strategies that make crucial use of knowledge about the reference domain. – Note, however, that there may be pronoun occurrences that can be processed by either strategy and possibly even with indistinguishable results (cf.[2]).

The defining characteristics of SPs, as developed in [1] and [2] on purely linguistic grounds, are that

- (i) SPs exhibit congruence with their controller in gender and number and person, and**
- (ii) SPs (in terms of categorial syntax) either form part of a functor whose argument expression is their controller or they form part of an argument expression whose functor is their controller.**

The controller of an SP can thus be determined merely on the basis of a categorial parse tree that includes the required agreement features.

### **3. Psychological evidence for syntactic processing**

Note that SPs comprise not only personal and possessive pronouns of the third person but also relative pronouns, reflexives, and reciprocals (plus, perhaps, other forms I have not yet investigated). As for personal and possessive pronouns, however, also referential occurrences (RPs) of the same forms are possible. How does the human processor behave in those cases? And how should a processing algorithm behave? The proposal I want to put forward is that the algorithm should make the distinction operational by making the default assumption that all pronouns are SPs and first attempt to process them as such. If this attempt fails, the same algorithm can continue the search for an antecedent NP; but then it will continue outside the control domain and the results will only have probabilistic value.

There is good evidence from experiments by Sanford et al. [8] that this is indeed how the human processor tackles the problem: the search initiated by the syntactic strategy, when it fails to locate an antecedent in the appropriate controller position, still continues and fixes upon another antecedent NP that meets the agreement requirement. Sanford et al. have shown this by demonstrating that pronoun comprehension is delayed in the presence of semantically or pragmatically unsuitable antecedent candidates that do however exhibit the required agreement features. They conclude that such “bonding” of the pronoun by irrelevant antecedents causes the processing delay.

Furthermore there is evidence from developmental studies [5] that shows, at least for a subset of SPs (the remainder not having been studied), that once the appropriate strategy has been acquired at the age of about eight, no significant improvement of the success rate of comprehension takes place, while for RPs the development continues. Similar results that also support the hypothesis of separate syntactic and referential (and thus world knowledge dependent) strategies are found in [12], [9], and [4].

### **4. Why does ‘naive’ pronoun resolution work so well?**

I have hinted at linguistic as well as psychological evidence for the SP-RP distinction and I have suggested, on the basis of experiments by Sanford et al. that the SP-strategy may indeed regularly be the first strategy employed (unless there is semantic information available that

can anticipate the result – as is very clear from the work by Tyler and Marslen-Wilson [11]) even beyond the syntactically appropriate search domain. Corrections of the results become necessary only in the latter case and then only if both criteria (syntactic agreement and controller position) admit a semantically or pragmatically incorrect referent.

Continuation of the search beyond the appropriate control domain yields NP-nodes of the appropriate feature constellation. Already if we place no constraints on the syntactic position of such nodes but consider any old preceding NP the chances that we get the correct antecedent are surprisingly high. Gender and number features alone guarantee 86.5% correct predictions when the correct antecedent is among the first two antecedents considered. Less surprisingly, the number of correct predictions drops very quickly as the number of the NPs increases that are considered and is only barely above chance level (27.4%) when five NPs are reached. These figures refer to abstract probabilities and are calculated on the basis of the distribution of the third person personal pronoun forms in Hobbs' [7] corpus. In actual fact this may account for a total of 56% of pronouns being assigned the correct antecedent, as it did in the case of Hobbs' corpus (cf. [7]) or, in an apparently particularly lucky case, even for 75.9%, as in my own corpus.

What needs accounting for though are not these cases but rather the additional success that is brought about by a continuation of the syntactic strategy beyond the control domain. The syntactic algorithm discussed in Hobbs [7] comes very close to this. It yields a total of 88.3% correct antecedent assignments for the pronouns in his corpus (and 81.2% correct assignments in mine).

My hypothesis for an explanation of the success of this algorithm is related to how Hobbs' algorithm and parse trees reflect the *aboutness-structure* of sentences. The background for this hypothesis is provided by a conjunction of two principles which I defended in [1]: the *Aboutness Principle of Anaphora* and the *Principle of Natural Sequential Aboutness*. The former claims that each independent or embedded sentence S that contains a (referential) anaphor A must be interpretable as being about the referent of A. The latter expresses the default assumption that each sentence says something about referents introduced in previous sentences. The obvious problem with both principles (as with similar notions proposed in [6]) is the fact that the notion of aboutness is not easily made operational. There is however a probabilistic operationalization in terms of Hobbs' parse trees: each (sub-)tree S is more likely about the referent of its highest NP node than about the referents of the highest NP-node in the next higher S, and so on following the order of NPs on the search path of Hobbs' algorithm. I found this rather rough and ready operationalization in agreement with the predictions made by a more principled (but much harder to handle) operationalization, which I defended for a different type of syntactic structure in [1] in over 80% of all cases in my corpus (an exact figure is hard to give because of some uncertainties I had about Hobbs' parse trees and some indeterminacies in my own account). Thus discounting for some unclear cases and allowing for exceptions from the Principle of Natural Sequential Aboutness (in my corpus it worked out correctly when applied to S (sub-) trees in 78.6%), the success of Hobbs' algorithm, and more generally of syntactic processing continued beyond the control domain, would seem explicable eventually on the basis of theoretically defensible regularities and tendencies of discourse structure.

It is therefore not a priori implausible, and on the basis of Sanford's results even a very reasonable conjecture, that a syntactic strategy is the first strategy a linguistic processor would employ for pronoun resolution, at least in the absence of semantic or pragmatic expectations (cf. above): it sets out inside the control domain on the default assumption that the pronoun is an SP (in my corpus 26.4% of all pronouns were SPs – 89 out of 337 occurrences) and may

continue beyond the control domain with no other difference than with respect to the reliability of the results and the grounds for the reliability: syntactic control mechanisms inside and the aboutness-structure of discourse outside the control domain.

Finally a word of caution. There should be no misunderstanding that I have considered the hypothesis for just one particular syntactic strategy for pronoun processing, which in case I am right is just one of a number of strategies which the human processor would ordinarily employ in interaction with each other. The abstraction in particular from all semantic considerations and all pragmatics makes this enterprise particularly problematic. I believe none the less that every now and then we must permit ourselves such abstraction in order to slowly reach an understanding of the various single factors involved in complex cognitive processes and not be confronted with a mere mess of parameters.

## References

- 1 Bosch, Peter: **Agreement and Anaphora. A Study of the Role of Pronouns in Syntax and Discourse.** Academic Press. London. 1983
- 2 Bosch, Peter: Pronouns under Control? **Journal of Semantics** 5:65-78 (1987)
- 3 Charniak, Eugene: **Towards a Model of Children's Story Comprehension.** Diss. MIT. 1972.
- 4 Clark, Herbert H. and Susan E.Haviland: Comprehension and the Given-New Contract. In: Roy O.Freedle (ed.): **Discourse Production and Comprehension.** Ablex. Norwood, N.J. 1977.
- 5 Deutsch, Werner, Charlotte Koster, and Jan Koster: What Can We Learn From Children's Errors in Understanding Anaphora? **Linguistics** 24:203ff. (1986).
- 6 Givon, Talmy (ed.) **Topic Continuity in Discourse.** Benjamins. Amsterdam. 1983.
- 7 Hobbs, Jerry R.: Resolving Pronoun References. **Lingua** 44:311ff (1978).
- 8 Sanford, Anthony J., S.C.Garrod, A.Lucas, R.Henderson: Pronouns Without Explicit Antecedents. **Journal of Semantics** 2:303-318 (1983)
- 9 Schaner-Wolles, Chris and Christine Kasem: Who is Who? On the Acquisition of Pronominal Interpretation. **Wiener Linguistische Gazette** 37:39ff (1986).
- 10 Sidner, Candace Lee: **Towards a Computational Theory of Definite Anaphora Comprehension in English Discourse.** Diss. MIT. 1979.
- 11 Tyler, Lorraine K. and William Marslen-Wilson: Processing Utterances in Discourse Context: On-line Resolution of Anaphors. **Journal of Semantics** 1:297-314 (1982)
- 12 Wexler, Ken and Y.C.Chien: The Development of Lexical Anaphora and Pronouns. **Papers and Reports on Child Language Development** 24:138ff (1985).
- 13 Wilks, Yorick A.: Preference Semantics, in: Edward L. Keenan (ed): **Formal Semantics of Natural Language.** Cambridge UP. London. 329ff. 1975.
- 14 Winograd, Terry: **Understanding Natural Language.** Edinburgh UP. Edinburgh. 1972.