

# WHAT PSYCHOLINGUISTICS TELLS US ABOUT THE SEMANTICS / PRAGMATICS INTERFACE: THE CASE OF PRONOUNS

Massimo Poesio, University of Essex

## Abstract

Whereas most theories of the meaning of pronouns view them as having a single, underspecified lexical meaning, psychological evidence about pronoun resolution and lexical access suggests that the lexicon does not assign pronouns an initial interpretation; rather, their interpretation is specified by competing defaults operating incrementally. We propose an incremental theory of the semantics and pragmatics of pronoun interpretation based on this evidence, that assumes that interpretation involves competing defaults activated in parallel.

## 1 Introduction

Montague's lexical rules for pronouns stipulating that pronouns are infinitely ambiguous at the lexical level (Montague, 1973) are one of the aspects of his theories most clearly at variance with a cognitively plausible view of the mental lexicon. Thus, the current consensus view about pronouns differs from Montague's in that pronouns are assigned a preliminary underspecified interpretation, in the form of a free index / variable denoting a function from assignments to values, sometimes with additional presuppositional content specifying that the entity is female / male, and / or that the variable must denote 'the most salient individual matching the pronoun in gender' (see, e.g., Heim and Kratzer (1998); Roberts (2001); Groenendijk et al. (1996). Jacobson, as well (1999), assumes that pronouns have a single underspecified interpretation.

The psychological evidence about lexical access and pronominal interpretation, however, challenges this view, suggesting instead that lexical access for pronouns involves immediate access to all (contextually salient) interpretations. This might at first seem to indicate that Montague was right after all; but current theories of underspecification allow us to formulate an alternative explanation which does not rely on assuming an infinite number of interpretations at the lexical level. If we allow the semantic interpretation of certain utterances to be left empty by the lexicon, and instead to be supplied by contextual rules operating incrementally immediately after lexical access, we can hypothesize that what happens in the case of pronouns is that contextual resolution rules apply immediately after the lexicon has been accessed, and it is these rules that provide the interpretation of the pronoun. In this paper, we show how this hypothesis about the division of labor between semantics and pragmatics can be made more precise.

## 2 Psychological Evidence about Pronoun Interpretation

The fundamental experimental result about lexical access - reported by, e.g., Swinney (1979) and later confirmed by experiments by Seidenberg et al, Marslen-Wilson et al, and Tanenhaus, among others - is that all lexical entries of a homonym like *bank* are immediately retrieved. A crucial additional fact about lexical access was revealed by experiments by Frazier and Rayner (1990), showing that garden path diagnostics can be used to classify words in two classes. Words like *pitcher* or *bank* have multiple lexical entries, which are accessed in parallel and immediately

pruned, which causes garden path effects when subsequent context reveals that the preferred interpretation is not the right one). However, immediate garden-path effects are not observed with words like *newspaper*, suggesting that they have a single lexical entry assigning them an underspecified meaning covering all the possible interpretations. Frazier and Rayner thus proposed a ‘weaker’ form of the incrementality hypothesis, the Immediate Partial Interpretation Principle (IPIH), stating that while utterances have to be immediately interpreted, the initial interpretation may be underspecified and refined only later.

Experimental results on pronominal interpretation by, among others, Corbett and Chang (1983) and Gernsbacher (1989) show that lexical access for pronouns is more like lexical access for *pitcher* than lexical access for *newspaper*. In sentences like *Karen poured a drink for Emily and then she put the bottle down*, both *Karen* and *Emily* are activated after the pronoun is encountered (but not before it, or when the proper name *Karen* occurs in its place).<sup>1</sup>

### 3 Accounting for the psychological results

As said above, one way of explaining these results is to argue that Montague was right after all, and pronouns are infinitely ambiguous at the lexical level; but there are obvious problems with this idea from a cognitive perspective. Fortunately, current theories of underspecification allow us to formulate the alternative explanation mentioned above.

This hypothesis about the division of labor between semantics and pragmatics can be made more precise using the theory of incremental semantic interpretation and underspecification developed in (Poesio, 1994, 1995; Poesio and Traum, 1997; Poesio and Muskens, 1997; Poesio, 2002). According to this theory, semantic processing is best viewed as a process during which oncoming utterances are incrementally assigned an interpretation, in part by the lexicon, in part by syntactic and contextual rules. As in sign-based theories, this interpretation is characterized in terms of functions specifying the syntactic and semantic classification of utterances, among which a function **sem**(*u*) specifying an utterance’s meaning. We briefly introduce the theory in this section; the interested reader is referred to the papers just cited for details.

**The common ground as a record of the occurrence of utterances** The theory we assume hypothesizes that the common ground includes a record of both the occurrence of utterances and their interpretation. This information is characterized using Compositional DRT (Muskens, 1995, 1994), modified as suggested in (Poesio and Muskens, 1997) to recover anaphoric accessibility. According to the theory, an utterance of the sentence *Kermit croaked* is incrementally interpreted as follows. First, the occurrence of an utterance of the word *Kermit* is recorded in the common ground: this amounts to updating the previous common ground *K* with the information that an event  $u_{pn}$  of uttering the word *Kermit* occurred, as follows:

$$(1) \quad K' = K; [u_{pn} | u_{pn} : \mathbf{utter}(a, \text{“Kermit”})]$$

$K'$  is then further updated as a result of lexical access and incremental parsing. As in sign-based theories, lexical access assigns a syntactic and semantic classification to  $u_{pn}$  in terms of the functions **cat** and **sem**; parsing hypothesizes that  $u_{pn}$  is a constituent of larger utterances—such an NP  $u_{np}$ .<sup>2</sup> The result of lexical access and parsing are the interpretations in (2) and (3), respectively.

<sup>1</sup>If only one antecedent matches the pronoun, the other interpretation is immediately discarded; otherwise both competing interpretations remain active until the end of the sentence is reached.

<sup>2</sup>We assume here for simplicity a purely bottom-up parsing.



The rules that eliminate the underspecification are not part of the lexicon but operate at a later stage, after the sentence is completed, as suggested by Frazier and Rayner (see also (Lascarides and Copestake, 1998)). Thus, no reanalysis is required when the disambiguating context is encountered.

## 4 The semantics and pragmatics of pronouns

The lexical entry for *he*, on the other hand, only specifies its syntactic properties (again, we omit all but the essential information):

```
U:utterance(X, "he") : cat(U) = pro, gender(U) = male,
                      num(U) = sing
-----
                      cat(U) = pro, gender(U) = male, num(U) = sing
```

Because this lexical entry does not specify the meaning for the pronoun, this must be derived from contextual reasoning in order to satisfy Frazier and Rayner's IPIH. Our contextual interpretation rules for pronouns use a conceptual vocabulary derived from Centering Theory (Grosz et al., 1995). More specifically, we assume that the local focus is updated by C-UTTERANCES, each of which introduces into context a number of ranked FORWARD-LOOKING CENTERS (CFs). A complete formalization of Centering involves both 'weak' rules activated in the absence of a CB, and stronger rules that operate when a CB has been identified. We only discuss here the weak rule for pronoun resolution, that hypothesizes as a meaning of a pronoun utterance  $U_{pro}$  occurring in c-utterance  $U_{n+1}$  the discourse entity assigned as a meaning of utterance  $U_{np}$  occurring in the previous c-utterance  $U_n$ , provided that the two CF-introducing utterances match. We write **cf-utt**( $U, U'$ ) for "U' is an utterance part of the c-utterance U that introduces a CF", and **prev-utt**( $U, U'$ ) for "U' is the c-utterance preceding U".

$$\frac{\begin{array}{l} [|\text{cat}(U_{pro}) \text{ is } pro, \\ \text{cf-utt}(U_{n+1}, U_{pro}), \text{cf-utt}(U_n, U_{np}), \\ \text{prev-utt}(U_{n+1}) \text{ is } U_n, \\ \text{agr-match}(U_{np}, U_{pro}), \text{sem}(U_{np}) \text{ is } x ] \end{array}}{[|\text{sem}(U_{pro}) \text{ is } x]} \quad \text{PRO-MATCH}$$

Under the assumption that the example sentences in Corbett and Chang's experiments involve two c-utterances, when the pronoun *she* is encountered in the second c-utterance, in a context in which both *Karen* and *Emily* are introduced in the previous c-utterance, the rule above is activated twice, generating two conflicting extensions. We discuss in the longer version of the paper the implications of these findings concerning the form of the 'stronger' rules implementing centering preferences.

## References

- Brewka, G., 1991. *Nonmonotonic Reasoning: Logical Foundations of Commonsense*. Cambridge: Cambridge University Press.
- Corbett, A. and F. Chang, 1983. Pronoun disambiguating: Accessing potential antecedents. *Memory and Cognition*, 11:283–294.
- Frazier, L. and K. Rayner, 1990. Taking on semantic commitments: Processing multiple meanings vs. multiple senses. *Journal of Memory and Language*, 29:181–200.
- Gernsbacher, M. A., 1989. Mechanisms that improve referential access. *Cognition*, 32:99–156.
- Groenendijk, J.A.G., M.J.B. Stokhof, and F. Veltman, 1996. Coreference and modality. In S. Lappin (ed.), *The Handbook of Contemporary Semantic Theory*, chapter 7. Blackwell, pages 179–213.
- Grosz, B. J., A. K. Joshi, and S. Weinstein, 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21(2):202–225. (The paper originally appeared as an unpublished manuscript in 1986.).
- Heim, I. and A. Kratzer, 1998. *Semantics in Generative Grammar*. Blackwell.
- Hobbs, J. R., M. Stickel, P. Martin, and D. Edwards, 1993. Interpretation as abduction. *Artificial Intelligence Journal*, 63:69–142.
- Jacobson, P., 1999. Towards a variable-free semantics. *Linguistics and Philosophy*, 22:117–184.
- Kameyama, M. and J. Jaspars, 1996. Preferences in dynamic semantics. In M. Stokhof and P. Dekker (eds.), *Proc. of the 10th Amsterdam Colloquium*. ILLC, Amsterdam.
- Lascarides, A. and N. Asher, 1993. Temporal interpretation, discourse relations and commonsense entailment. *Linguistics and Philosophy*, 16(5):437–493.
- Lascarides, A. and A. Copestake, 1998. Pragmatics and word meaning. *Journal of Linguistics*, 34(2):387–414.
- Montague, R., 1973. The proper treatment of quantification in English. In K.J.J. et al. Hintikka (ed.), *Approaches to Natural Language*. Dordrecht: D. Reidel, pages 221–242. Reprinted in Thomason (1974).
- Muskens, R., 1994. A compositional discourse representation theory. In P. Dekker and M. Stokhof (eds.), *Proceedings of the 9th Amsterdam Colloquium*.
- Muskens, R., 1995. Tense and the logic of change. In U. Egli, P. Pause, C. Schwarze, A. von Stechow, and G. Wienold (eds.), *Lexical Knowledge in the Organization of Language*. Amsterdam / Philadelphia: John Benjamins, pages 147–183.
- Poesio, M., 1994. *Discourse Interpretation and the Scope of Operators*. Ph.D. thesis, University of Rochester, Department of Computer Science, Rochester, NY.
- Poesio, M., 1995. Disambiguation as (defeasible) reasoning about underspecified representations. In P. Dekker and M. Stokhof (eds.), *Proc. of the Tenth Amsterdam Colloquium*. ILLC.

- Poesio, M., 2002. *Incrementality and Underspecification in Semantic Interpretation*. Lecture Notes. Stanford, CA: CSLI. To appear.
- Poesio, M. and R. Muskens, 1997. The dynamics of discourse situations. In P. Dekker and M. Stokhof (eds.), *Proceedings of the 11th Amsterdam Colloquium*. University of Amsterdam, ILLC.
- Poesio, M. and D. Traum, 1997. Conversational actions and discourse situations. *Computational Intelligence*, 13(3):309–347.
- Roberts, C., 2001. Uniqueness presuppositions in english definite noun phrases. Available on the Semantics Archive.
- Swinney, D. A., 1979. Lexical access during sentence comprehension: (re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, 18:545–567.
- Thomason, R. H. (ed.), 1974. *Formal Philosophy: Selected Papers of Richard Montague*. New York: Yale University Press.