

The Relative Contribution of Syntactic and Semantic Prominence in Pronoun Reference Resolution

Ralph L. Rose <rose@ling.northwestern.edu>, Northwestern University

This paper presents the results of a series of psycholinguistic experiments which compare the influence of syntactic prominence and semantic prominence on the salience of antecedents for pronominal reference.

Subjects are preferred antecedents for pronominal reference (Mathews and Chodorow, 1988). Many pronoun resolution algorithms account for this by assuming a hierarchy of syntactic prominence for antecedents as in (1) (cf., Lappin and Leass, 1994). The upper part of this hierarchy has been validated in numerous psycholinguistic (e.g., Hudson-D’Zmura and Tanenhaus, 1997) and computational (Lappin and Leass, 1994) studies. However, for many verbs in English, syntactic role is conflated with semantic role: That is, syntactic SUBJECTS are often semantic AGENTS and so on. So what appears to be the result of the prominence of syntactic SUBJECTS might actually be the result of the prominence of semantic AGENTS with respect to a hierarchy of semantic roles as in (2).

(1) SUBJECT > OBJECT > OBLIQUE

(2) AGENT > PATIENT > OTHERS

The first experiment was designed to examine the lower part of the syntactic and semantic hierarchies by using constructions which allow alternation of their internal arguments—*spray/load* verbs as in (3)a-a’. Here we are interested in the relative salience of the two nonhuman entities for subsequent reference by the pronoun *it*. In (3)a, syntactic and semantic prominence converge and the utterance should therefore have a single salient entity—namely *paint*. This is the CONTROL condition. However, in (3)a’, syntactic and semantic prominence diverge onto separate entities: as a surface OBJECT, *wall* is the syntactically more prominent entity, but it is less semantically prominent (as a LOCATION) than the PATIENT, *paint*. Hence, this condition—called here the SPLIT condition—allows observation of the relative influence of syntactic and semantic prominence.

(3) a. John sprayed some paint_i on a wall_j. CONTROL

a’. John sprayed a wall_j with some paint_i. SPLIT

b. It_i dribbled down and made a mess. COREF_{pat}

b’. It_j was big and needed two coats. COREF_{loc}

Experiment participants read one of the two context sentences (i.e., (3)a or a’) followed by a continuation beginning with a pronoun. The continuations were written to refer unambiguously to one of the two entities in the context sentence: *paint* or *wall* (see (3)b and b’, respectively). The experiment was self-paced with participants pressing a button to continue reading. Whole-sentence reading times of the continuations were recorded. Results show only a main effect of continuation sentence [by subjects $F(1, 31) = 11.13, p < 0.005$; by items $F(1, 47) = 5.61, p < 0.05$] and no significant interaction. In the CONTROL condition, participants read the continuation which referred to the syntactically and semantically prominent entity faster (by subjects $t(31) = 3.61, p < 0.005$; by items $t(47) = 3.00, p < 0.005$). Hence, when syntactic and semantic prominence converge, there appears to be a clearly salient entity in the discourse for subsequent pronominal reference. In the SPLIT condition, participants also read the continuation which referred to the semantically prominent entity faster: However, this difference was not significant. Thus when syntactic and semantic prominence diverge, participants seem to have more difficulty deciding which entity is more salient, although an apparent preference is given to the semantically more prominent entity. These results suggest that semantic prominence contributes to the discourse salience of entities for pronominal reference and that syntactic prominence makes a relatively weaker contribution. Similarly, it seems OBJECTS are perhaps only slightly more prominent than OBLIQUES—a finding which suggests revision to current notions of the syntactic prominence hierarchy.

Converging evidence from similar experiments using *tough*-constructions (another argument-reordering construction) and the *repeated-name penalty* technique of (Gordon et al., 1993) will also be presented.

The experimental evidence suggests that both syntactic and semantic prominence contribute to the discourse salience of entities: neither factor alone determined participants’ preferences in the SPLIT condition. However, results suggest that semantic prominence is the stronger of the two factors. Furthermore, the results call into question the validity of the lower end of the syntactic prominence hierarchy (which has long been assumed but never explicitly verified experimentally). Finally, implications for the design of discourse salience representations and pronoun resolution algorithms will also be discussed.