

Pronouns Without Explicit Antecedents: How do We Know When a Pronoun is Referential?

Jeanette K. Gundel*, Nancy Hedberg[†], Ron Zacharski[∧]

*University of Minnesota and NTNU
Minneapolis, MN, USA, 55414 and Trondheim, Norway
gunde003@umn.edu

[†]Simon Fraser University
Burnaby, BC, Canada, V5A 1S6
hedberg@sfu.ca

[∧]New Mexico State University
Las Cruces, NM, USA, 88003
raz@crl.nmsu.edu

Abstract

Pronouns without explicit noun phrase antecedents pose a problem for any theory of reference resolution. We report here on an empirical study of such pronouns in the Santa Barbara Corpus of Spoken American English, a corpus of spontaneous, casual conversation. Analysis of 2,046 third person personal pronouns in fourteen transcripts indicates that 330 (or 16.1%) lack NP antecedents. These pronouns fall into a variety of subtypes. 88 refer to entities that are inferrable from an activated frame or script, or are otherwise easily accommodated. In 110 cases, *it* could refer to a fact, proposition event, activity, situation, or reason which has been evoked by a previous non-NP. 92 cases of *it* were classified as pleonastic. In this paper we focus on some interesting subclasses of pronouns which could be analyzed as either referring to entities of various degrees of abstractness that were introduced by or implied in previous discourse, or as non-referential or pleonastic pronouns. Such cases include possible truncated cleft pronouns, possible truncated extraposition pronouns, and certain non-specific uses of *they*.

1. Introduction

The referent of prototypical pronoun has been recently introduced into the discourse by a noun phrase, i.e. the pronoun has an explicit noun phrase antecedent, as in (1):

- (1) a. My neighbor s Bull Mastiff bit a girl on a bike.
b. It s the same dog that bit Mary Ben last summer.
(Gundel et al., 2000)

It is well known, however, that such an antecedent is neither necessary nor sufficient for appropriate pronoun use, especially in the case of personal pronouns. Examples like (2) show that use of *it* is not always acceptable even when there is a recent NP antecedent:

- (2) a. Sears delivered new siding to my neighbors with the Bull Mastiff.
b. #It s/That s the same dog that bit Mary Ben last summer. (Gundel, et al., 1993)

Moreover, an NP antecedent, or an explicit antecedent of any sort, is not always necessary, a fact which poses problems for any theory of reference resolution (see Cornish 1999, Gundel, Hedberg and Zacharski 2000, Byron 2000, inter alia).

We report here on an empirical study of pronouns without NP antecedents in the Santa Barbara Corpus of Spoken of American English, a corpus of spontaneous, casual conversation. Particular attention will be paid to

cases where it isn t clear whether the pronoun refers to a specific fact, situation, proposition, etc. or whether it is the subject of a truncated cleft (Hedberg 2000), a truncated extraposition construction, or is otherwise possibly non-referential.

2. Methodology

As mentioned above, we obtained our data from the Santa Barbara Corpus of Spoken American English Part-1 (DuBois, et al., 2000). We analyzed the first 790 intonation units in each of 14 transcripts, ranging from 571 to 1493 seconds. The transcripts represent conversations between two to six speakers in a variety of settings. For example one conversation took place when the three participants were preparing a dinner, another took place at a birthday party, and a third was a classroom discussion.

2.1. Coding

We coded each of the 2,046 third-person personal pronouns (excluding false starts) in the corpus for whether or not it had an NP antecedent. Pronouns without NP antecedents were further classified as clearly pleonastic (i.e. lacking a referent), possibly pleonastic, or referential; and referential pronouns were classified into type of referent (e.g. proposition, fact, activity, reason) introduced by a non-NP or as an inferrable. We now turn to a brief introduction of each of these categories.