1 Introduction

1.1 Major themes of the study

The last ten years have seen a tremendous upsurge in work on discourse production and comprehension, correlated with a growing concern in a variety of disciplines with language as it is used in context. Because of its fundamental place in the understanding of memory, discourse structure and semantic interpretation, anaphora has been the focus of much of this research (e.g. Grosz 1977; Reichman 1981; Sidner 1983; Tyler and Marslen-Wilson 1982; Webber 1983; Givón 1983; Halliday and Hasan 1976; Bosch 1983; Linde 1979; Reinhart 1983). Central to this work has been the belief that there is a strong relationship between the flow of information in a text, the structure of the text, and use of anaphora. A recurrent, and intuitively appealing, finding of this work is that referents which are “in focus” or “in the hearer’s consciousness” can be pronominalized, where focus or consciousness are operationalized in terms of the discourse structure (see in particular Grosz 1977 and Reichman 1981).

The present study holds to this interpretation of the relationship between discourse structure and anaphora. One of the themes that runs through this study is that any treatment of anaphora must seek its understanding in the hierarchical structure of the text-type being used as a source of data. Texts may be produced and heard/read in a linear fashion, but they are designed and understood hierarchically, and this fact has dramatic consequences for the linguistic coding employed.

But, while sharing this common background with earlier works, this study departs from them in several critical ways. Because of the (natural) emphasis in cognitive science on information processing, work in the area of discourse structure has tended to view discourse as organized purely in terms of information flow and propositional content. This view is limited in two ways: it ignores the critical role played in all text-types by social, interactional, and affective factors (Linde 1979, van Dijk and Kintsch 1983,
2 Introduction

and Duranti 1984 are important exceptions); and relatedly, it fails to take into consideration the fact that texts are organized as they are not just because of informational limitations (what Dillon (1981) calls channel limitations) but also because of socially accepted conventions (cf. Wittgenstein 1958 on language games). As Dillon (1981: 15) has said:

The general point is that if we abstract ‘conveying propositional content’ as the common property of written discourses, we have woefully impoverished the notion of discourse as a human communicative act. Or, to put it another way, we have created an enormously artificial model of discourse and have obliterated from our view the elaborate sets of conventions governing particular discourse types and the ways these can be employed to signal diverse and complex intentions.

The present study, in attempting to provide a complete account of the distribution of a subset of anaphora – reference to third-person singular humans – in conversation and expository prose, brings out the social as well as informational aspects of the relationship between discourse and anaphora. We will see that interactional and affective factors, as well as genre-specific conventions, do indeed play a significant role in anaphoric patterning in conversation and writing.

I have limited the scope of this study to third-person singular human references. I have narrowed the domain in this way to look at anaphora in its prototypical use: tracking a participant through a discourse (Du Bois 1980; what Prince (1981) would consider textually evoked references). Other uses of anaphora, such as the this in example (1) and they in example (2) below, introduce further complexities: for example, they may refer to previous utterances, rather than to participants – example (1) – or they may not refer at all (in the sense of being non-referential: Du Bois 1980) – example (2).

(1) Americans love apple pie. This is something to wonder at.
(2) This street has gotten very noisy recently. I think they're putting up a new apartment building.

These functions of various anaphoric devices are interesting and deserve further study, but they represent complexities beyond the function of repeated references to an entity over time, and I have thus chosen not to include them in the present study.1

Since I am interested here in the relationships between higher-level discourse considerations and anaphora, the patterns of anaphora I have examined represent what some have called “discourse anaphora” (related to what Bosch (1983) has called referential pronouns), in that the instances analyzed here are not controlled syntactically (see Reinhart 1983 for a
thorough discussion of syntactically controlled anaphora; Bolinger 1977 takes up similar phenomena from a functional perspective). Thus, I did not examine anaphora of the sort illustrated in the following sentences:

(3) Zelda adores her teachers.
(4) Rosa complained that she had a headache. (Reinhart, 1983)

The analyses presented in this study include the patterning of pronouns and full noun phrases in the environment of other referents (both same-gender and different-gender). I have included these patterns because such cases provide crucial information towards understanding how people manage potentially complex and “ambiguous” referring situations. In fact, to some extent it is possible to understand what counts as cognitively complex or ambiguous only by exploring such patterns. I also felt that it was important to determine exactly what kind of referent (in addition to what kind of structural environment) counted as problematic, hence the division of referents into same and different gender.

The data for this study consist of (1) naturally occurring face-to-face and telephone conversations, and (2) small newspaper and magazine articles and segments of a psychoanalytic biography. Because of the theoretical orientation adopted here, which states that anaphora correlates at least in part with hierarchical discourse structure, it was necessary to “parse” these texts into their component units, so that an understanding of the relationships between anaphoric patterning and discourse structure could be achieved. For this task, I chose two different analytic tools, one for each of the two modalities: conversational analysis, developed to structurally analyze spontaneous conversation (Sacks et al. 1974); and rhetorical structure analysis, designed for planned expository prose (Mann et al. 1982; Matthiessen and Thompson forthcoming).

The use of two fairly distinct models for the two types of texts is based on the belief that the modes are fundamentally different in the units that serve to organize them. Written monologue texts are by definition produced by one person, and the units of a descriptive model should reflect this basic one-party-ness; conversational texts, on the other hand, are by definition produced by more than one person, and the units of a text model should also reflect this fact. In addition, written monologue texts are largely information-oriented, and the structural units of an appropriate model should capture the types of informational relationships that can hold between pieces of text, since these relationships give the texts their hierarchical structure (Decker 1974; Graesser and Goodman 1985);
4 Introduction

Conversational texts, on the other hand, are largely interactional, and the units of an appropriate model should capture the social action relationships that hold between pieces of talk, inasmuch as it is these more-than-one-party actions which structure the talk (see Rubin 1980 for a discussion of these and other differences between written and “spoken” language). Thus, while there are clearly interactional and affective aspects of written monologue texts and informational aspects of conversational texts, these are not the fundamental structuring units for that mode. It has therefore been critical in this work to use one model which focuses on the informational relationships between propositions for the written monologue texts and another which focuses on the interactional relationships between utterances for the conversational texts. From the small set of candidates for each type I selected conversational analysis and rhetorical structure analysis. Inasmuch as these approaches are likely to be unfamiliar to most readers, they will each be described in detail in their own chapters.

1.2 Organization

The remainder of the study is organized as follows: Chapter 2 gives an introduction to the method of structural analysis used for the conversational data, conversational analysis. The basic units of the model are discussed, and example passages are analyzed. This chapter is background for Chapter 3, which details the actual anaphoric patterns found in the conversational texts. Chapter 3 formulates the anaphoric patterns found in the conversational material. It lays out the patterns of anaphora in the conversational material in terms of the hierarchic organization of the texts. Structural patterns, as well as non-structural ones (i.e those performing other functions), are explored. Chapter 4 gives an introduction to the method of structural analysis used for the monologue expository written texts, rhetorical structure analysis. The basic units of the model are discussed, and example passages are analyzed. This chapter provides the background needed for the analyses offered in Chapter 5. Chapter 5 formulates the patterns of anaphora found in the expository written texts, using rhetorical structure analysis as the tool for exploring the structural designs of the texts. The patterns of anaphora offered in this chapter, as in previous chapters, are presented in terms of the hierarchic organization of the texts in which they occur. Chapter 6 compares the anaphoric patterning established for the conversational and expository texts using quantitative and qualitative methods of comparison. Evidence is presented which demonstrates that the
Organization 5

two text-types differ fairly strikingly in the distribution of anaphors they display. Chapter 7 offers thoughts on the theoretical issues of localness of pattern and the nature of discourse structure.
2 Conversational analysis

This chapter is meant as an overview of the fundamental concepts of conversational analysis which will be put to work in Chapter 3. It is by no means a complete guide to this approach; I have included only those notions which will be applied in the analyses in Chapter 3. For a more complete introduction to CA, see Levinson 1983 or Atkinson and Drew 1979. Readers already familiar with CA can skip to Chapter 3.¹

2.1 Notation conventions

Before we examine some of the substantive findings of conversational analysis, I would like to present and discuss some of the notational conventions used in CA-style transcripts. Many of these conventions may be unfamiliar to linguists, but they are all in fact fairly straightforward. (The sources of the transcripts reproduced in this book are given in abbreviated form – AD:14, SN:4:30, etc. – at the end of the extracts. Further details of these sources are given in section 3.2.)

All talk is transcribed in a pseudo-phonetic system, using the basic orthographic symbols of written English; that is, if the speaker pronounces a word in a way that is not the only possible pronunciation for that word, then special care is taken to transcribe that particular pronunciation (but without using a special alphabet). This practice makes the transcripts somewhat difficult to read (especially for the non-native speaker), but since it brings out useful information I have not normalized the transcripts (except for crucial pronouns, which have been normalized for ease of reading and exposition).

The double slash (//) indicates the place at which a speaker’s utterance is overlapped by talk from another speaker.

M. No, they’re all thin.
C. They’re not

(AD:14)

Thus with this notation we can see that C’s utterance starts after the a in M’s
Notation conventions

_all._ The right-hand bracket (]) indicates the place in the utterance at which the overlap ends (so C’s overlap lasts until after M’s thin).

An utterance which has more than one double slash in it is overlapped at more than one place, and the utterances which do the overlapping are given in sequential order after the overlapped utterance.

G. they’re all Keegans like the ones around Greensprings
   they’re all kind’v, // bout five five, five sl/x,
M. They’re all from around Greensprings//gs.
C. Yeh

(AD:14)

Here, M’s utterance overlaps G’s starting at _bout_, and C’s utterance overlaps G’s starting with the _x_ in _six_. Notice that C’s utterance also overlaps the very end of M’s.

A left-hand bracket at the beginning of two lines indicates that the two utterances begin simultaneously.

M. Yeh.
   [C. Long time ago it reminds me

(AD:14)

The equals sign (=) indicates latching, that is, the next speaker begins without the usual “beat” of silence after the current speaker finishes talking. In this case there is an equals sign at the end of the current speaker’s utterance and another equals sign at the beginning of the next speaker’s utterance. If two speakers simultaneously latch onto a preceding utterance (that is, they begin talking simultaneously), this is indicated in the transcript with a left-hand bracket preceded by an equals sign.

(R) (h)hh (h)uh (h)uh (h)uh! =
(S) hhh(h) H(h)m
   =[
K. Which la:mpost?

(SN-4:30)

Here S and K simultaneously latch onto R’s laughter.

Numbers given in parentheses indicate elapsed silence, measured in tenths of seconds. Single parentheses with a raised dot between them represent a silence that is less than a tenth of a second but still longer than the usual beat of silence. These figures are not arrived at with a stop-watch, but are calculated with a verbal counting technique which takes into account
8 Conversational analysis

the tempo of the preceding talk. Thus a silence which is timed at 0.3 of a second in one stretch of talk might well be timed at something else in another stretch of talk if the tempo of the preceding talk is different. The numbers thus indicate “experiential time” rather than chronological time.

Certain facts about the production of the talk are given through the orthographic symbols used. Punctuation is used to suggest intonation; underlining indicates stress. A colon after a letter means that the sound represented by that letter is somewhat lengthened; a series of colons means that the sound is increasingly lengthened. Anything preceded by a degree mark is quiet.

The letter h within parentheses indicates “explosive aspiration,” and usually means some type of laughter is being produced. A series of hs preceded by a raised dot represents an inbreath (where number of hs is meant to correspond to the length of inbreath), while the same series preceded by nothing represents exhaling. Capitalization is used to indicate increased volume.

Questionable transcriptions are enclosed within single parentheses; the transcribers thereby indicate that the exact form of the utterance is not clear. Speaker’s initials given in single parentheses means that there is some question about the speaker’s identity. Double parentheses – e.g. ((clears throat)) – represent non-transcribed material (i.e., noise which is non-linguistic).

These are the major transcription conventions which will be used in the data fragments in this and later chapters. For a more detailed guide to CA notational conventions, see Sacks et al. 1974.2

To help the reader identify the structure of some of the passages, I have created a simple system of labeling which indicates which adjacency pair an utterance belongs to, what level of structure it is to be heard at, and whether it is a first-pair part, a second-pair part, or a third-position utterance. This simple system works in the following way.

The label has three slots: the first indicates the position in the adjacency pair – first-pair part, etc. – that the utterance fills (first-pair part is abbreviated fpp, second-pair part spp). So, for example, the following invented example has two utterances, the first labeled fpp, the second spp.

A. Do you have a computer at home? [fpp]
B. No. [spp]

The second slot in the label reflects a sequential numbering of the adjacency pairs. The first adjacency in a cited passage is given the number 1,
Notation conventions

the second adjacency pair the number 2, and so on. The numbering is done strictly by temporal ordering. All parts of a single adjacency pair (i.e. the first-pair part and the second-pair part, along with any third-position utterance) are given the same number. In the following example, we have two adjacency pairs:

B.  hh Hey do you see v- (0.3) fat ol’ Vivian [fpp (1)]
    anymouh?
A.  No, hardly, en if we do., y’know, I just say [spp (1)]
    hello quick’n ‘hh y’know, jus’ pass each othuh
    in th//e hall.
B.  Is she still hangin arous (with) Bo::nny? [fpp (2)]
    (TG:14)

Notice that the fpp and the spp of the first pair are both assigned the number 1, while the fpp of the next pair is assigned the number 2.

The third slot of the label indicates the level of structure at which the utterance fits in (members of a single adjacency pair are treated as being at the same level of structure). For example, if we have an announcement pair followed by a post-elaboration question–answer pair, the announcement pair will be labeled as being at a higher level of structure than the post-elaboration (lower-case letters are used for this slot, beginning with a). This labeling is meant to capture the impression that the announcement in such a situation is somehow nuclear, or core, and the post-elaboration is somehow embedded, or subordinate, or adjunct. An example of this labeling follows:

M.  A:nd ( ) as far as that goes my father’s on his [fpp (1,a)]
    honeymoon. =
    (y:ah ha ha”ha)
K.  (Oh:::) Very nice= [spp (1,a)]
K.  = Where’d he go. [fpp (2,b)]
    (SN-4:10)

M’s utterance is a first-pair part (fpp), and is in the first adjacency pair of the fragment (1). In addition, it is at the highest level of structure of the fragment (a). K’s first utterance possesses all of the same features, except that it is a second-pair part (spp). K’s second utterance, on the other hand, is different: it is the beginning of the second adjacency pair, and it is at an “embedded” level of structure with regard to the first pair (indicated by the letter b).

This notation system is meant to provide a “map” for some of the
10 Conversational analysis

fragments presented in Chapter 3. It has no theoretical status, and is offered only as a simple schematic guide for the reader.

2.2 The turn-taking system

One of the most critical aspects of conversational structure is turn-taking. How are turns allocated in informal conversation? To answer this question we need to establish some rudimentary facts about turns.

We need to first establish that there are units out of which turns can be constructed. These units have been referred to as turn-constructional units (TCUs), and can be single lexical items (yes, for example), phrases, clauses, or sentences. According to the turn-taking system, each speaker is at first allotted one of these TCUs. The end of such a unit constitutes a place where speaker-change could occur: that is, at this point another person could begin talking. The end of a TCU is thus a transition-relevance place (TRP), since it is a place at which a transition from one speaker to another can (but need not) occur.

The following turn-taking rules, which are based on these concepts, are taken verbatim from Levinson (1983, which is based on Sacks et al. 1974):

Rule 1 – applies initially at the first TRP of any turn:

(a) If the current speaker selects a next speaker in current turn, then current speaker must stop speaking, and that next speaker must speak next, transition occurring at the first TRP after next-speaker selection.

(b) If current speaker does not select next speaker, then any (other) party may self-select, first speaker gaining rights to the next turn.

(c) If current speaker does not select next speaker, and no other party self-selects under option (b), then current speaker may (but need not) continue.

Rule 2 – applies at all subsequent TRPs:

When rule 1(c) has been applied by the current speaker, then at the next TRP rules 1(a)–(c) apply, and recursively at the next TRP, until speaker change is effected.

These rules provide a foundation for making sense out of two related phenomena: simultaneous talk and silence.

Simultaneous talk obviously occurs when two (or more) speakers talk at once. But not all simultaneous talk represents a violation of the turn-taking system rules. Rather, there are two basic types of simultaneous talk (i.e. overlap): competitive overlap and non-competitive overlap. In one type of non-competitive overlap, called terminal overlap, the current speaker approaches the end of a turn-constructional unit and, as that is happening,