OPTIONAL PRONOMINALIZATION AT THE DISCOURSE LEVEL

J. Hinds

In this paper I will examine the notion that pronominalization across sentence boundaries is optional. It is usually stated that, given the following sequence of sentences, either the repetition of the noun phrase or the appropriate pronoun is acceptable in the second sentence.

(1) Last year, $Mary_i$ bought five baby elephants at the baby elephant store. This year $\begin{bmatrix} Mary_i \\ she_i \end{bmatrix}$ bought five more.

It is the thesis of this paper that pronominalization across sentence boundaries is not completely optional, and that the use or nonuse of a pronoun in such cases signals a major difference in structural organization.

First, I assume that grammatical relationships do obtain at the discourse level. Justification for this position may be found in Sanders (1969), van Dijk (1972), Hinds (1973a), among others. A wellformed discourse, as I am using the term, is a complete text, either a monologue or a dialogue, which has a structured beginning and a structured endpoint. I will leave undefined the terms 'structured beginning' and 'structured endpoint' [but see Taber (1966) for some relevant comments]; instead I will provide a few illustrations. In the case of a narrative monologue in English, a structured beginning may be any of the following, as well as similar formulas which are unspecified here.

(2) a. Once upon a time, ...
 b. Let me tell you something ...
 c. I've been asked to speak to you today ...

In the case of dialogue, appropriate structured beginnings are:

(3) a. Hi, how are you.
 b. Let me introduce myself.
 c. Hello. (over the telephone)

Possible structured endings for narratives are:

(4) a. ... and they lived happily ever after.
 b. ... and that's that.
 c. ... this concludes what I want to say today.

and for dialogues:

(5) a. ... well, be seeing you.
 b. ... it's been nice meeting you.
 c. ... goodbye. (over the telephone)

Anything else is a nonwellformed discourse. Two comments are relevant here. One, the study of the form of structured beginnings and structured endings has been completely neglected. We do not even have a taxonomic classification of types or ranges of variations, to say nothing of a workable semantic framework [see, however, Lakoff (1973), Taber (1966)]. Two, nonwellformed discourses may profitably be studied, and in fact, nonwellformed discourses provide the bulk of material under investigation in this study.

In any wellformed discourse, there is usually more than one 'discourse topic', a discourse topic being simply the topic of conversation. At present I can offer no formal definition of discourse topic. Pragmatically, the discourse topic is the answer to the question, What is being spoken (written)about now? Each sequential section of a wellformed discourse which is concerned with the same topic is termed a paragraph. There is obviously a certain amount of overlap between the term paragraph as it is being used here and the rhetorical paragraph, but these terms are not completely synonymous.

In certain languages, paragraph boundaries are formally marked. For instance, in Sarangani Manobo, a language spoken by about 8000 people on the east coast of the Sarangani Peninsula of souther Mindanao in the Philippines, a new paragraph is marked by 'the sentence initial conjunction na or nayan, by multiple time reference, or by a combination of these'. [Dubois (1973:19)] In Angaataha, an Angan language of New Guinea spoken by about 650 people, there is a characteristic intonational feature which signals the termination of a paragraph. This feature is 'a low intonation lower than any other low tone in the paragraph'. [Huisman (1973:31)]

In English, as in other Indo-Europen languages, paragraph boundaries are not obligatorily marked either morphologically or phonologically. There are some conjunctive expressions, however, that can optionally indicate either that a paragraph is continuing or that it is changing. Expressions that indicate that a paragraph is continuing are:

(6) continuing, and, moreover, of course, ...

Some conjunctive expressions that indicate that the paragraph is changing are:

(7) on the other hand, next, first... second... third,
to change the topic a bit, ...

However, even these expressions do not obligatorily indicate the continuation and/or change of paragraph, and there are many conjunctive expressions which are absolutely ambiguous in their function:

(8) but, however, so, ...

Complicating the problem a bit more is the fact that within any paragraph, there will be 'subparagraphs'. These subparagraphs are developments of the major topic of discussion. A good example of this is a typical newspaper article. The article is usually 'about' only one topic; that is, it consists of a single paragraph. Often, but not always, this topic is the headline that accompanies the article. The elucidation of that topic,

however, contains a number of specific points of particular importance.
Each of these specific points of particular importance will be termed the
'peak of the subparagraph', or more simply, the 'peak'. A hierarchical
disgram of the structure of a paragraph is presented in (9):

(9) P(ARAGRAPH)

$Peak_1$ $Peak_2$ $Peak_3$ ... $Peak_n$

(S) (S) (S)

[+pro] [+N] [+pro]
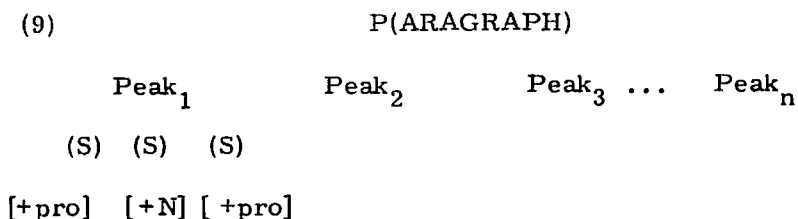
This diagram illustrates that a paragraph consists of a number of peaks,
and that each peak consists of a number of sentences. The features under
the S nodes will be discussed below. Although this diagram is in the form
of a branching tree diagram, no particular importance should be attached
to that fact, since I am not concerned with the 'generation' of a paragraph.
Pragmatically, the peaks are determined by asking the question, What are
the major points made in the paragraph?[1] The peaks are in a semantically
coordinate relationship with one another. They are always manifested by a
sentence (or a sentence fragment). Any other sentences (or sentence frag-
ments) in the paragraph will be semantically subordinate to the peaks.[2]
    I turn now to pronominalization. Pronominalization may occur
across sentence boundaries if the following conditions are met.[3]

(10) Within a given paragraph, pronominalization is
      possible if (a) the referent is properly registered
      in the discourse registry, and (b) there will be no
      vagueness or ambiguity created because of multiple
      antecedents.

A few comments on these conditions are in order. First, it is necessary
to constrain the limits of pronominalization to paragraph boundaries to pre-
vent situations such as the following. Assume a proper discourse. The
topic of the first paragraph is Richard M. Nixon. The paragraph concludes
and the topic of the next paragraph is established as Martha Mitchell.
Within the second paragraph, assuming no mention of Nixon, it is necessary
to produce (11) and to block (12).

(11) She [i. e. Martha] meets Nixon twice a week.

(12) She [i. e. Martha] meets him twice a week.

    Second, the referent must be properly registered in the discourse
registry. This is usually done by the introduction of the referent at some
prior point, but it may also be done by implication [see Phillips (1973) for
thoughts on this].
    Third, pronominalization must usually be blocked if there is more
than one possible referent.[4] For instance,

(13) *Bill and John are always late for work.
He's been late all his life.

There are a number of complicating factors here, such as semantic rela-
tions and grammatical functions of possible referents, which make it diffi-
cult to specify precisely when pronominalization is blocked in case there
are two possible referents of the same sex and number in the preceding
sentences of a paragraph. For instance, both of the following are perfectly
acceptable and unambiguous, even though there are two possible referents
in each case.

(14) John loaned Bill$_i$ his car. When he$_i$ returned it,
it had a big dent in the front fender.

(15) John$_i$ went all the way to Seattle to meet his uncle.
Then he$_i$ had to go back to Newark to meet his cousin.

In the following, I consider only instances in which there is one
possible antecedent. This is a proper limitation since it is under precisely
this condition that pronominalization at the discourse level is termed op-
tional. I examine two newspaper articles. It will be shown that both news-
paper articles, although markedly different in form and content, exhibit
the same structural properties, and indicate that pronominalization at the
discourse level, far from being optional, signals major structural proper-
ties of the paragraph.

(16) FAMOUS ARTIST IWATA DIES AT 73
(Japan Times: 9 February 1974)

(1) Sentaro Iwata, a celebrated artist and 1961
winner of the Medal of Honor with Purple Ribbon
known for the sensuous illustrations of slim, kimono-
clad women he did for periodicals, died of cerebal
hemorrhage in Tokyo Tuesday. (2) He was 73.
(3) He complained of a severe headache and nausea
at about 8 p. m. Monday while working on magazine il-
lustrations at his house in Shibuya-Ward, Tokyo, and
soon fell unconscious.
(4) He was taken to the Keio University Hospital in
Shinano-machi, Tokyo, where he died at 10:35 a. m.
Tuesday.
(5) Born as the son of a printer in Asakusa, Tokyo, in 1901,
Iwata became one of Japan's most popular illustrators
when, at 25, he worked for the famous novel "Ako Roshi"
(The Tale of 47 Ronin) written by the late Jiro Osaragi.
(6) In 1955, Iwata won the Kan Kikuchi Prize, an
award for those having done outstanding work in art and
journalism.

The article has been reprinted with its original paragraphing, al-
though I have numbered each sentence for ease of reference. It might be

- 148 -

pointed out that newspaper paragraphs coincide neither with rhetorical paragraphs nor with paragraphs as I have been using the term.

Notice initially that it is indeed possible to substitute the noun Iwata for the sentence initial occurence of He in sentences (2), (3), and (4). In addition, it is possible to substitute he for Iwata in sentences (5) and (6). If these substitutions are possible, how can it be denied that discourse pronominalization is optional? The reason is that the substitutions, while possible, indicate a distinct difference in organizational content.

The major point of this article is to announce Iwata's death. This is reflected in the headline, as well as being stated explicitly in sentence (1). Sentence (2) presente information which is subordinate to this major point, as do sentences (3) and (4) since they develop or expand the circumstances surrounding Iwata's death. Sentence (5) departs from making a point about Iwata's death by noting one of the major achievements of his life. This constitutes the second peak. Sentence (6) presents a second major achievement in Iwata's life, parallel to, but not subordinate to, the first. Sentence (6) therefore constitutes the third peak. Substituting Iwata for He in sentences (2), and (3) or (4) creates the impression that the sentence in which the substitution is made constitutes a peak, or coordinate idea, rather than a subordinate idea. This would be most likely, I believe, for sentence (2) in case Iwata's age were something particularly noteworthy. For instance, if sentence (2) were to read

(16:2') Iwata lived until the grand old age of 145.

there would be more motivation for assuming that sentence (2) should be semantically coordinate with sentence (1).

If we compare this with the following article, we find that similar results obtain.

(17) GIRL SUBDUES ATTACKER
          (Mainichi Daily News: 14 March 1974)

(1) A brave young woman turned the tables on a robber and beat him with an iron pipe which she had wrested from him, then handed him over to the police in Osaka Wednesday night. (2) About 11:25 p. m. Wednesday, a man attacked Miss Mayumi Sanda, 23, of Oyodo-cho, Oyodo-ku, Osaka, on a street in the same ward. (3) The man struck her several times on the head with an iron pipe and tried to strangle her. (4) When the man demanded money, she asked him to slacken his grip on her neck. (5) When the man released his grip, Miss Sanda snatched the iron pipe from him and beat him. (6) Astounded by her counterattack, the man was discouraged and quickly subdued. (7) She took him to the Oyodo Police Box and the man was immediately arrested. (8) The police identifed the man as Mitsuo Henmi, 24, who was born in Tokyo and has no fixed abode. (9) Henmi told the police that he had wanted to get some money to go back to Tokyo. (10) However, he said, he was extremely hungry and lost interest when she hit back at him.

The major points in this article are: (I) a woman was attacked; (II) she surprisingly fought back and subdued her attacker; and (III) a justification for the attack. Now, if we examine the pronominalization phenomena, we find that they correspond to these major points. There are some complications, but they are explainable in a realistic fashion.

Sentence (1) has the purpose of summarizing the entire article in one sentence. This is a standard newspaper technique [cf. sentence (16:1)] and might be considered a special type of structured beginning, although I will not pursue this line of investigation here. The sentence clearly has the purpose of establishing the discourse topic. Sentence (2) is the actual beginning of the paragraph, since it (re)introduces the main characters. Notice that it provides most of the essential information: it specified when, where, and what happened, and who was involved. The fact that the female character is introduced by name and the male character is referred to simply by the epithet man informs us that (at least) this section of the paragraph is concerned primarily with the female character, rather than the male. This sentence constitutes the first peak. Sentence (3) is an elaboration of sentence (2). For this reason her is more appropriate than a repetition of the name.

Sentence (5) constitutes the next peak. It is the one event which is most surprising in this whole sequence of events. Sentence (5) refers to the female character as Miss Sanda, rather than she. Sentence (4) is semantically subordinate to sentence (5) since its purpose is to set the stage for the major event. The female character is referred to by a pronoun in this sentence. Sentences (6) and (7) are also semantically subordinate to (5), since they report the consequences of the second peak. The female character is referred to in each of these sentences by a pronoun.

Sentence (8) switches the prime concern of the article from the female character to the male by introducing his name and giving some background information on him. Sentence (9) constitutes the third peak, since it gives a justification for the attempted robbery. For this reason the man is referred to again by name, rather than by a pronoun. Sentence (10) merely elaborates on this, and so may be considered semantically subordinate. The man is referred to by a pronoun in this sentence.

One problem which I will only mention here involves the use of she in sentence (10). I find this usage a bit awkward, and would prefer to see a full noun phrase such as the woman or Miss Sanda in its place. If my intuitions are correct, and this usage is really a mistake on the part of the author [see again note 1], there would be sufficient justification for considering this article as having two closely related paragraphs; one, comprising sentences (2) - (7), describing the events in relation to the female character; the other, sentences (8) - (10), describing the events in relation to the male character. With the use of she, I am forced to conclude that there is only one paragraph, but that there are two different foci, the female and the male [cf. Kuroda (1973) for some relevant comments].

In summary, I have attempted to show that cases of what appear to be optional pronominalization at the discourse level actually are subject to specific constraints. The nonuse of a pronoun, assuming the conditions in (10) are met, indicates a structural peak in a paragraph. The use of a pronoun indicates that the sentece containing that pronoun is semantically subordinate to a peak. While this is not the final answer to optional pronomin-

alization at the discourse level, it is an attempt to provide some order to an otherwise chaotic situation.

Notes

1. In order to answer this question, a specific 'competence', in the sense of ability is assumed. It should also go without saying that individuals may differ in their degrees of competence. That is, it is altogether possible that two native speakers of a language, looking at the same material, will arrive at different conclusions about what the peaks of a given paragraph are.
2. It is possible that a peak could be established by implication rather than by an actual occurence of a sentence. I will not consider this situation here [see, however, Phillips (1973)].
3. These conditions are not absolutely correct. First, items may be registered in a discourse registry by implication [see Karttunen (1968), Phillips (1973)]. Second, some items may exist in a 'permanent registry' [see Hinds (1973a), Takahashi et al. (1972)]. Third, there are situations in which ambiguity or vagueness is desired. Fourth, there may be cases in which a particular effect is desired, as for instance, the use of a pronoun with no antecedent to build suspense, etc. [see Hinds (1973b)]. Fifth, the speaker's assumptions may cause a violation of these principles.
4. Since a speaker may assume that the hearer can successfully disambiguate a sentence, even in a case in which the hearer cannot, there may be accidental violations of this constraint.

Bibliography

van Dijk, Teun. 1972. Some Aspects of Text Grammars. The Hague: Mouton.

Dubois, Carl. 1973. 'Connectives in Sarangani Manobo narratives'. Linguistics 110:17-28.

Hinds, John. 1973a. 'Japanese discourse structure'. Unpublished Ph. D. Dissertation: SUNY/Buffalo.

_____. 1973b. 'Passives, pronouns, and themes and rhemes'. Paper presented at LSA Winter Meeting, San Diego.

Huisman, Roberta. 1973. 'Angaataha narrative discourse'. Linguistics 110:29-42.

Karttunen, Lauri. 1968. 'What makes definite noun phrases definite?' RAND Corporation Reprint.

Kuroda, Sige-Yuki. 1973. 'When epistemology style and grammar meet-- A case study from Japanese' in A Festschrift Presented to Morris Halle. NY: Holt, Rinehart, and Winston, 377-91.

Lakoff, George. 1972. 'Structural complexity in fairy tales'. The Study of Man I:128-50.

Phillips, Brian. 1973. 'Topic analysis'. Paper presented at 1973 International Conference on Computational Linguistics, Pisa, Italy.

Sanders, Gerald. 1969. 'On the natural domain of grammar'.
Indiana University Linguistic Club Reprint.

Taber, Charles. 1966. The Structure of Sango Narrative. Hartford CT:
Hartford Seminary Foundation.

Takahashi, Hirose, Osamu Fujimura, and Hisao Kameda. 1972.
'A behavioral characterization of topicalization: What is topical-
ization from a computer point of view'. Research Institute of
Logopedics and Phoniatrics Annual Bulletin 6:93-110. Tokyo
University: Tokyo, Japan