J. Linguistics **24** (1988), 65–87. Printed in Great Britain

Referring and accessibility

MIRA ARIEL

Department of Linguistics, Tel-Aviv University

(Received 4 March 1987; revised 20 June 1987)

The analysis of referring expressions can be divided into two branches for our purposes. The first includes theories of definite descriptions and proper names. The key to the riddle of the appropriate use of such expressions, it was thought, is the notion of presupposition: existence and/or uniqueness. Indeed, this was the question that dominated the literature for many years, starting with the early philosophical analyses of Frege (1982), Russell (1919) and Strawson (1956, 1964), and ending with the much later pragmatically oriented linguistic analyses, such as Liberman (1973), Kempson (1975), Prince (1978, 1981b), Gazdar (1979), McCawley (1979), Hawkins (1974, 1984) and even Loftus (1972, 1974, 1975), although this last approach is more psychological. The second branch of research totally neglected the question of presupposition. Non-syntactic/semantic theories of anaphoric expressions, pronouns especially, were psychologically oriented, and hence saw the issue to be accounted for quite differently. In fact, the objective of these theories has been to elucidate processing procedures by examining anaphoric expressions, rather than to make claims about anaphoric expressions as such. Factors identified as relevant in the processing of anaphoric expressions are listed in (1):

(1) Factors affecting accessibility of antecedents:
  (a) Distance between antecedent and anaphor.
  (b) Number of competitors for the role of antecedent.
  (c) Importance of topicality in antecedent assignments.
  (d) Role of frames in identifying antecedents.

Work in this spirit was carried out by Caramazza *et al.* (1977), Clark & Sengul (1979), Sanford & Garrod (1981), Garrod & Sanford (1985), and others. More linguistic findings or experiments on the distribution of anaphoric expressions in discourse of various languages can be found in Linde (1979) (English), Li & Thompson (1979) (Chinese), Clancy (1980) (Japanese, English), Gundel (1980) (universally), Grosz (1981) (English), Yule (1981) (English), Marslen-Wilson *et al.* (1982) (English), Enç (1983) (Turkish), Givón & associates (1983) (Japanese, Amharic, Ute, Biblical Hebrew, Latin-American Spanish, written English, spoken English, Hausa, Chamorro), and Schiffman (1984) (English).

This artificial division into a 'definiteness' function, attributed to definite

descriptions and proper names on the one hand, and an anaphoric function, attributed primarily to pronouns (and marginally to definite descriptions) on the other hand, seems to derive from the wrong assumption that forms having the former function are always 'antecedents', or 'full NP's' (i.e. expressions used to refer to world entities directly or independently), whereas the latter are necessarily dependent on some linguistic antecedent for referent interpretation. Now, this division of labour is not totally baseless, but it certainly fails in the case of some actual uses. Extending claims independently made by Marslen-Wilson *et al.* (1982), Givón (1983) and Ariel (1985), I will argue that it correctly predicts the distribution of referential expressions in unmarked initial references, but that a better characterization of their use is based on the notion of Accessibility, defined with respect to (at least) the four factors listed above. We begin by pointing out the disadvantages of defining referring expression use by reference to context (1.1), and adduce some empirical evidence supporting the general Accessibility claim (1.2). Section 2 is a more detailed analysis of various referential Accessibility markers, and Section 3 proposes a comprehensive scale of Accessibility marking, incorporating into one scale the separate conclusions regarding relative Accessibility of markers reached in Section 2.

1. From context type to degree of accessibility

1.1. *Against contextual definitions for referring expressions*

Within a given stretch of discourse, when first referring to non-new entities, speakers must rely on one of three possible contexts to supply the antecedent. Indeed, Clark & Marshall (1981) propose a pragmatic theory, defined over three contexts, to account for the use of referring expressions. They suggest MUTUAL KNOWLEDGE between speaker and addressee as a basis for an appropriate use of a referring expression, though it must be pointed out that their main interest lies in explicating the use of definite descriptions. Mutual knowledge can be established on the basis of any of the following:

(2) Mutual knowledge
  (a) *Community membership mutual knowledge*: consists of that knowledge all community members are assumed to possess, plus a system of inferencing.
  (b) *Physical co-presence mutual knowledge*: that knowledge speakers have by virtue of their attending to the physical situation they are in.
  (c) *Linguistic co-presence mutual knowledge*: knowledge shared by speakers because it has been mentioned in the discourse itself.

As Clark & Marshall themselves note, a pragmatic theory that attempts simply to correlate context types and linguistic markers of givenness,

referring expressions among them, cannot be maintained. Many markers can be used on more than one basis, some of which are not even particularly marginal. For example, depending on the text genre, definite descriptions are often used anaphorically more than half the time (see Ariel, 1985: 71 and forthcoming), and demonstratives are more often used anaphorically than exophorically (see Halliday & Hasan, 1976). Even proper names (first names especially) may be dependent for their interpretation on fuller forms. Last, pronouns and even gaps are usable as primary identifying expressions, though this is quite a marginal use. The examples in (3) to (7) illustrate the above:

(3) *The President* (a definite description).
  (a) *The President* will come in any moment now (Herzog, Reagan, etc.).
  (b) Ronald Reagan flew to Japan. *The President* is scheduled to meet with Japanese feminists (probably Reagan, not ???Herzog).

(4) *Gandhi* (a last name).
  (a) There is no one like *Gandhi* (Mahatma Gandhi, Indira Gandhi, Ragib Gandhi).
  (b) When Indira Gandhi was alive, people used to say 'There is no one like *Gandhi*' (probably Indira Gandhi, not ???Mahatma Gandhi, ???Ragib Gandhi).

(5) *This woman* (a demonstrative).
  (a) *This woman* is fantastic (Giora, Prince, Reinhart, etc., depending on the actual pointing).
  (b) Giora worked so hard in preparing for our conference. *This woman* is fantastic (Giora, not ???Reinhart, ???Prince, etc.).

(6) *It* (a pronoun).
  (a) Sherlock Holmes to Watson: The butler did *it* (the murder).
  (b) What a horrible murder. The butler did *it*.

(7) $\emptyset$
  (a) (On a packet of pills): $\emptyset$ contains methanol.

While in (3) to (5) we see that definite descriptions, proper names and demonstratives can all be referentially dependent on a previously mentioned linguistic antecedent, (6) and (7) show that pronouns and zeros can be referentially independent.

Prince (1981a) proposes a more sophisticated approach to context type. Her analysis would suggest a grading of referring expressions as pointing to varying degrees of givenness or familiarity, but she does not actually attempt to characterize specific linguistic forms, such as definite descriptions, deictics, proper names and pronouns. The categories on her scale, given in (8), are not linguistic categories. Rather, they are properties of referents which are potential discourse entities:

(8) Prince's (1981a) Familiarity Scale (simplified):

$$\left\{\begin{array}{c}\text{EVOKED textually}\\ \text{EVOKED situationally}\end{array}\right\} > \text{UNUSED} > \text{INFERRABLE} > \text{BRAND-NEW}$$

However, for *givenness* to be a proper linguistic term, accounting for the distribution of referring expressions, it should be defined according to the way language CODES this scale. It is not clear how the relationship between linguistic forms and discourse conditions of use should be defined under the terms in (8), for example.

Instead of accounting for reference by the notion of context, I suggest that natural languages primarily provide speakers with means to code the ACCESSIBILITY of the referent to the addressee. Accessibility, in its turn, is tied to context types in a definitely NON-arbitrary way. The three context types (general knowledge, physical surroundings and previous linguistic material) are on the whole hierarchically ordered as to degree of Accessibility to the addressee. In the unmarked case, information stored as 'general knowledge' is not automatically accessible. Information from the physical surroundings, provided it is one that speakers are attending to, is mentally represented with a higher degree of Accessibility. Recent linguistic material is the most accessible source of information, other things being equal. The analogy to long-term memory vs. short-term memory and active memory seems obvious enough though, as we will later see, current models of memory are not rich enough to reflect the variety of distinctions in Accessibility I shall argue for. I suggest that we incorporate into one single system the context-type theories of definite descriptions and the psychological accounts of the interpretation of anaphoric expressions. Specifically, I propose that instead of claiming that an expression type *x* is processed in a certain way, as the psycholinguists have claimed, we view the processing procedure associated with each form as its inherent definition. In other words, referring expressions are no more than guidelines for retrievals.

This move will first of all help us in explaining the apparent difference in choice of referential forms between initial and subsequent retrievals of an entity. Consider first retrievals. Note that when (a) and (b) in (9) and (10) refer to the same world object and the speaker is in a position to utter the (b) sentences, these are indeed the preferred forms:

(9) (a) ##? *That woman over there* is very intelligent.
(b) ## *Rachel* is very intelligent.

(10) (a) ##? There's this woman I know. *She's* very intelligent.
(b) ## *That woman over there* is very intelligent.
(## = Discourse Initial)

(9) shows that first introductions of given discourse entities by proper name are preferred to the use of a demonstrative expression, and (10) shows that a demonstrative expression is better than a pronoun.

https://doi.org/10.1017/S0022226700011567 Published online by Cambridge University Press

In terms of the number of contextual implications in the sense of Sperber & Wilson (1982, 1986), it makes sense that the speaker should refer the addressee to the most general context for the interpretation of given entities, for it is rich in background information about the referent. Optimal relevance, however, dictates that the cost of processing be taken into consideration as well. Obviously, interpreting a referent on the basis of the larger amount of information stored in our general knowledge takes longer. As the examples above show, this weighting is automatically decided with respect to first retrievals. Speakers must supply addressees with the richest possible context (using the (b) versions above), or else, if they use the (a) versions, they may be accused of being misleading, and certainly of being less than fully cooperative. The picture seems to be reversed with subsequent mentions. Note how (11) is bizarre, with a second mention of Geraldine Ferraro by her full name:

(11) Geraldine Ferraro has been an active Democrat for quite a few years. But she/??Geraldine Ferraro ran for Vice-Presidency only in 1984.

Accessibility, I suggest, as measured by the various factors discussed by the psychological researchers and listed in (1) is the determining principle which accounts for the choice of referential forms in both environments. The difference in acceptability between a proper name and a pronoun in subsequent retrievals is due to the fact that pronouns, unlike proper names, mark high Accessibility. Since in (11) Ferraro had just been mentioned, a marker of high Accessibility is more appropriate. The initial referential act, on the other hand, where the speaker is unable to assume a high level of Accessibility, had to be achieved by using a marker of low Accessibility: a proper name in the case of (9) and (11), a demonstrative in (10).

1.2. *Empirical evidence*

In principle, each and every factor identified above as determining degree of Accessibility can and should be tested in order to provide evidence for my claim. Because of lack of space we will concentrate on distance, the factor easiest to quantify; but see the references above, as well as Ariel (1985), for findings regarding other factors. When we examine texts, where the three types of referring expressions refer to entities previously mentioned in the text, we can clearly see the role of Accessibility in determining choice of marker. Though all items refer then to the same context in the Clark & Marshall division of mutual knowledge, and all are textually evoked according to Prince (1981a), consistent differences in distribution can be discerned. The general picture that emerges from counts restricted to anaphoric references is that pronouns are predominantly used when the distances are short, anaphoric demonstratives are used in cases of

intermediate distance, and definite descriptions mostly refer back to antecedents outside the sentence they occur in, when their antecedents are not even close by.[1] Tables 1 and 2 present some statistics collected from four texts, each consisting of about 2200 words.[2]

	Text position									
	Same sentence		Previous sentence		Same paragraph		Across paragraph		Total	
Referring expression	no.	%	no.	%	no.	%	no.	%	no.	%
Pronoun	**110**	**(20.8)**	**320**	**(60.5)**	75	(14.2)	24	(4.5)	529	(100)
Demonstrative	4	(4.8)	**50**	**(59.5)**	**17**	**(20.2)**	13	(15.5)	84	(100)
Definite description	4	(2.8)	20	(14.1)	**65**	**(45.8)**	**53**	**(37.3)**	142	(100)

Table 1
Breakdown of anaphoric expressions as to text positions

	Expression							
	Pronoun		Demonstrative		Definite description		Total	
Text position	no.	%	no.	%	no.	%	no.	%
Same sentence	**110**	**(93.2)**	4	(3.4)	4	(3.4)	118	(100)
Previous sentence	**320**	**(82.1)**	50	(12.8)	20	(5.1)	390	(100)
Same paragraph	**75**	**(47.8)**	17	(10.8)	**65**	**(41.4)**	157	(100)
Across paragraph	**24**	**(26.7)**	13	(14.4)	**53**	**(58.9)**	90	(100)

Table 2
Popularity of anaphoric expressions in text positions

Table 1 shows that both pronouns and demonstratives favour a position where the antecedent occurs in the previous sentence. The difference between them is mainly reflected in their secondary environments, pronouns being preferred within the same sentence, and demonstratives within the same paragraph. Definite descriptions favour the two most distant positions (same paragraph and across the paragraph). In fact, we can say that the majority of occurrences of each expression type (about 80 per cent) centre on the two positions indicated in bold for each referring expression. Table 2 analyses the

[1] As Antecedent for a Given entity I counted the last reference made to the same entity, rather than the full NP to which it is anaphoric.

[2] Two of the texts examined were fictional discourse. The first is the opening section of a novel (Young, 1934: 7–13). The second is a short story (Paley, 1980: 43–52). For non-fiction I chose 'Trouble on the set' (pp. 1–12), and Tucker, 1985: 1–6.

same data from the perspective of textual positions: i.e. given a certain textual position, what likelihood is there for a certain referring expression to occur? Not surprisingly, co-reference in intra-sentence position almost always means use of pronouns (93 per cent). Previous sentence position clearly favours pronouns as well, but at a somewhat lower rate (82 per cent). A position further away in the paragraph seems to favour pronouns and definite descriptions (89 per cent together). Finally, the across-the-paragraph position is predominantly filled by definite descriptions, although quite a few pronouns also occur in this position.

With respect to pronouns, Table 2 seems something of a counter-example. Even if we take into consideration, which indeed we must, that demonstratives (in English) are simply not very popular, the high percentages of pronouns in the two distant positions seems to go against our claim that pronouns mark high Accessibility. However, as mentioned in (1) above, distance is not the only factor affecting Accessibility. Another factor is the salience of the referent, mainly as related to its being a topic or a non-topic.[3] Indeed, once we leave out the anaphors referring to discourse topics in our counts, the picture changes quite dramatically. Whereas the short-distance positions are hardly affected, the longer-distance positions now conform to the Accessibility claims. Table 3 presents the same data, but omitting references to the discourse topics (520 references out of the 755 in Tables 1 and 2):

	Expression							
	Pronoun		Demonstrative		Definite description		Total	
Text position	no.	%	no.	%	no.	%	no.	%
Same sentence	**69**	**(89.6)**	4	(5.2)	4	(5.2)	77	(100)
Previous sentence	**189**	**(73.1)**	50	(19.3)	20	(7.7)	259	(100)
Same paragraph	34	(29.3)	17	(14.7)	**65**	**(56)**	116	(100)
Across paragraph	2	(2.9)	13	(19.1)	**53**	**(78)**	68	(100)

Table 3
Popularity of non-topic anaphoric expressions in text positions

Thus, when we neutralize the factor of topicality we can clearly see that at short distances anaphoric expressions of the High Accessibility type are most

[3] For a definition of sentence topic, see Reinhart (1981). Since the texts I analysed were simple prose in the sense that they were all about specific human characters about whom predications were constantly added on, establishing the discourse topics was intuitive and easy. For a more general attempt to outline a discourse-topic extraction procedure, see Giora (1985). The discourse topic referred to above would correspond to the NP part of the topic proposition, as argued by Giora.

https://doi.org/10.1017/S0022226700011567 Published online by Cambridge University Press

popular, or least marked. At larger distances, anaphoric expressions of Low Accessibility are most popular. Since demonstratives (Mid Accessibility markers) are not widely used in English compared to other languages, they never predominate, not even in the middle-distance position, but as Table 1 clearly reveals, the majority of their occurrences fall in the two intermediate distances. Actual distance was, of course, only partially taken into account in these findings. I preferred to consider sentence and paragraph boundaries, since textual unit boundaries are crucial factors in themselves, in addition to the indication of distance. Memory scope is crucially related to textual units, and it seems clear that at least some material signalling the closure of a unit is released from short-term memory.

Yule (1981) presents similar findings. He too reaches the conclusion that a distinction of given vs. new is not enough to characterize the actual use of definite forms. 'Clearly, the use of more attenuated forms (e.g. *it*, Ø) is much more frequent for current entities' (Yule, 1981:49). Thus the degree of attenuation increases after the first mention of the new entity, but it decreases after a while. Whereas 92 per cent of the references to non-recent entities in his data were achieved by demonstrative pronouns and definite descriptions, only 36 per cent of the references to current entities used such markers. The numbers for pronouns and Øs are the mirror image. Whereas in 64 per cent of the references made to current entities such markers were used, non-recent entities were referred to by attenuated forms in only 8 per cent of cases. Clark & Sengul (1979) present experiments showing that people took less time comprehending a sentence when the antecedents of definite NPs and anaphoric pronouns were mentioned in the previous sentence, as opposed to two and three sentences back. They assign no importance to the difference between the processing time of definite descriptions and pronouns when antecedents appeared three sentences away. Nevertheless, their findings show that pronouns took 5 per cent longer than definite NPs when the antecedent was not very close. Purkiss (1978), as reported in Sanford & Garrod (1981), achieved similar results. Whereas pronouns are better retrievers for topics, even distant ones, definite descriptions are better retrievers for non-topics two and three sentences away. Indeed, this is what is predicted by a theory of Accessibility. A definite description should be a better retriever when the antecedent is not highly accessible, as when it is a non-topic or when it is a distant one; a pronoun is a better retriever when the antecedent is highly accessible, i.e. a topic or a close-by entity.

Segmentation into paragraphs plays an important role in decisions about the contents of current working memory. It is therefore not surprising that we find that initial paragraph position of texts, even ones whose discourse topics have long been established, tends to reintroduce the topic in a full NP form anyway, rather than in pronominal form. In three Hebrew texts I examined, containing a total of 60 paragraphs, 53 mentioned the continuing

discourse topic in the first sentence of the paragraph.[4] Of these, 75.5 per cent chose a Low Accessibility marker, rather than a High Accessibility marker. Clancy (1980) and Marslen-Wilson *et al.* (1982) also found textual boundaries to be relevant.

2. Degrees of accessibility: finer distinctions

We have up to now concentrated on a rough three-way distinction between High, Mid and Low Accessibility Markers, but Accessibility comes in a much richer variety of shades. We turn now to examine markers within each category, at the end of which I shall attempt to present a full scale of Accessibility markers.

2.1. *Low Accessibility markers*

Low Accessibility markers are those commonly uttered on the basis of general knowledge in unmarked initial retrievals, e.g. proper names and definite descriptions. Now, although all Low Accessibility markers refer to entities currently not in working memory, distinctions do exist as to how inaccessible that entity is. The less accessible it is, the lower the Accessibility marker chosen. This is understandable, since lower Accessibility markers normally contain more lexical information, which is needed for successful retrieval (cf. *the man who left* vs. *the man*; *this hat* vs. *this*, etc.). Immediately accessible entities do not require so much wording for the entity to be evoked. Two kinds of distributional facts can help us establish degree of Accessibility: rate of popularity among first-mention given entities, and distance between the referring expression and its antecedent. Low accessibility markers tend to retrieve initially, and to refer back to antecedents further away. I decided to apply only the second criterion, for the first one demands the choice of texts to be limited to those containing well-known figures. I checked two articles and the opening section of a short story.[5] First, Table 4, which presents a comparison between definite descriptions and full proper names (i.e. names consisting of both a first and a last name).

The counts reveal that full names are lower Accessibility markers than definite descriptions. In the same sentence, and in positions further on in the same paragraph, their pattern of distribution is not significantly different, but definite descriptions have a significantly higher percentage of occurrence in previous-sentence position, while proper names occur much more often in the most distant position, across the paragraph. In passing, we should mention

[4] The three articles are: 'The weakness of the strong man', *Haaretz*, 1 July 1983; 'Was he or was he not?' and 'Starting all over', both in the special Rosh-Hashana supplement of *Haaretz* 7 Sept. 1983. All were in Hebrew and all were character portraits.

[5] The two articles appeared in the supplement of *Haaretz*, 16 Aug. 1985. The short story is H. Ben-Yehuda's Ha-Kadish, in Y. Berlovich (ed.) (all in Hebrew).

	Same sentence		Next sentence		Later in same paragraph		Across paragraph		Total	
	no.	%	no.	%	no.	%	no.	%	no.	%
Definite descriptions	6	(3.7)	60	(36.6)	42	(25.6)	56	(34.1)	164	(100)
Full names	2	(3.4)	12	(20.7)	15	(25.9)	29	(50)	58	(100)

Table 4
Occurrence of markers in various textual positions

that a combination of a proper name and definite description (no matter in which order) is the lowest Accessibility marker. Only a negligible percentage of these occur in subsequent mentions. Similarly, a special V.I.P. introductory expression consisting of a proper name and an adjacent free relative, which I have identified for Hebrew (see Ariel, 1983), is restricted to first retrieval only. This is probably the closest we ever get to having a purely referential expression.

But in cultures where more than one type of name is used, different types of proper names can also be distinguished as to degree of Accessibility. First names are shortest, and most equivocal. They are normally used when the referent is not only familiar, but also an intimate,[6] and hence permanently 'accessible'. Last names are intermediate. They are (in western cultures, at least) less ambiguous than first names, and they do not imply intimacy. A typical paragraph might very well start with a full description or a full name, preferably both, and then change to a last name alone, and afterwards a pronoun.[7] The examples in (11) are typical:

(11) (a) (i) ha + max a zay, yehošua sobul, ...
The playwright, Yehoshua Sobul, ...
(ii) sobul...
Sobul...
(iii) hu...
He... (*Haaretz*, 14 May 1985)
(b) (i) ha +sofer ha +yehudi ha +mizrax–germani,
The writer the Jewish the east German,
stefan haym.
Stephan Heim.
(The East German Jewish writer, Stephan Heim...

[6] The above claim is appropriate mainly for men. The fact that women, children and minorities (such as Arabs in Israel) are referred to by first names even when they are not intimates of the speaker is to be explained as a special derogatory use.

[7] References change back to fuller forms after a while, when the speaker wishes to remind the addressee who the referent is. Moreover, repeated references by full name increase the likelihood that an item will be stored in long-term memory (Kintsch, 1970).

(ii) haym...
Heim...
(iii) ...sifro...
...his-book... (*Haaretz*, 14 May, 1985).

A statistical examination of the same three texts mentioned above indicates clearly that proper names differ in Accessibility, so that full names are lowest and first names are highest. Table 5 shows the percentage of each name type in first retrieval position. Full names predominate, whereas there are no first names initially.

Full name		Last name		First name	
no.	%	no.	%	no.	%
132	(85.2)	23	(14.8)	0	(0.0)

Table 5
First retrieval by name

Comparing distances between the names and their co-referring entities (i.e when names are used in subsequent retrievals) confirms the above on the whole, as can be seen in Table 6.[8] The popular positions for each name type are in bold print and the same hierarchy of names is revealed.

	Same sentence		Previous sentence		Later in same paragraph		Across paragraph		Total	
Name type	no.	%	no.	%	no.	%	no.	%	no.	%
Full name	2	(3.4)	12	(20.7)	15	(25.9)	**29**	**(50)**	58	(100)
Last name	0	(0.0)	**31**	**(39.2)**	**24**	**(30.4)**	**24**	**(30.4)**	79	(100)
First name	3	(6.1)	**21**	**(42.9)**	**20**	**(40.8)**	5	(10.2)	49	(100)

Table 6
Names in textual positions

2.2. *Mid Accessibility markers*

As we have already stated, the degree of Accessibility of the marker used should correspond to the Accessibility of the referent it is used to refer to.

[8] Table 6 contains three unexpected findings, but all of them can be explained. First, the fact that names recurred within the same sentence is due to a direct quotation. The relatively high percentage of last names across paragraphs was limited to the discourse topic and to other extremely famous personalities. The very high rate of first names in the two distant positions (51 per cent together) is due to the fact that one of the articles is about an Arab (and her Arab husband). Members of minorities are often referred to by first names (see Ariel (forthcoming)).

The less accessible the representation of a given entity, the lower the Accessibility marker used. Applying this principle to markers initially used to point towards entities within the physical environment suggests that *that*, for example, should be categorized as a lower Accessibility marker than *this*, and so should all distal demonstratives. Indeed, these differences show up not only in pure deictic uses. When anaphoric, distal demonstratives tend to refer to more remote entities, while proximal demonstratives refer to more immediate antecedents. This is the difference between Latin *hic* vs. *ille*, German *dieser* vs. *jener*, Spanish *este* vs. *ese*, French *celui-ci* vs. *celui-là*, Turkish *bu* vs. *o*, and Hebrew *ze* vs. *ha+hu* (see Fillmore, 1975; Lyons, 1977). Claiming that *that* is lower on the scale than *this* also fits the fact that it is *that* and not *this* in English, *ille* and not *hic* in Romance, *az* and not *ez* in Hungarian (as well as many other languages) that developed into a definite article (see Greenberg, 1978; Harris, 1980). It is the distal demonstrative, an already lower Mid Accessibility marker, that tends to develop into a full-fledged Low Accessibility marker.

A comparison between demonstrative pronouns and missing NPs, where the deletion is dictated by the physical presence of the objects referred to, seems to point to another Mid Accessibility gradation. Missing NPs require more accessible entities than demonstratives. The former are usually restricted to references to the physical objects (packets, bottles, etc.) on which the sentence appears. This goes very well with the claim about lexical information being a factor here. The emptier the marker the higher the Accessibility rate. Gaps are obviously as empty as one can get. Note the example in (13a) vs. (13b):

(13) (a) Open Ø with care.
(b) This (meeting) is quite interesting.

Since the more informative the marker the more help it is in retrieving less accessible information, we should also distinguish between demonstratives used as substantives on the one hand and as adjuncts on the other (*this/that* vs. *this/that book*). In fact, once the speaker adds more information, for instance in the form of a relative clause, Mid Accessibility markers turn into Low Accessibility markers. (14) is such as example:

(14) That holiday we spent in Cyprus was really something.

To sum up, we may say that Mid Accessibility markers in their central uses connect discourse to given entities from the physical surroundings. However, since most references in actual discourse are made to textual units, Mid Accessibility markers often also function anaphorically. As such, they mark less accessible information than High Accessibility markers, such as pronouns. Moreover, since Mid Accessibility markers are themselves distinguished as to Accessibility, proximal and distal markers serve to refer to entities over shorter and longer textual distances respectively.

2.3. *High Accessibility markers*

High Accessibility markers are those normally thought of as requiring a linguistic antecedent (pronouns, gaps). Indeed, such referential expressions more often than not do have a linguistic antecedent but, as we have seen above, even gaps can refer extra-textually. Still, the point remains that these expressions are used only when speakers can assume a very high level of Accessibility for the referent. We begin with the anaphor vs. pronominal distinction, which we suggest is dictated by Accessibility differences. We then propose that the difference between unstressed and stressed pronouns can also be accounted for in terms of differences in Accessibility.

2.3.1. *Anaphors and pronominals.* Government and Binding theory (Chomsky, 1981) distinguishes between anaphors, pronominals and names based on their difference with respect to deictic reference. The claim is that anaphors are interpreted only as bound anaphora, i.e. the anaphor must be c-commanded by an antecedent within the same governing category, while pronouns and names are not so restricted. Pronouns are 'free' (i.e. they lack a co-indexed c-commanding antecedent) within the same governing category, while names are free in the whole S domain. In our terminology, anaphors should be categorized as higher Accessibility markers than pronominals (and names, of course). Pronouns can refer independently or deictically, but most often they refer to an antecedent within the previous clause or two, whereas anaphors are restricted to finding their antecedents within the same minimal domain as that in which they occur. These differences can then account for choices of $\emptyset$ versus pronoun, whether grammatically or pragmatically performed. In Ariel (forthcoming), we devote a lengthy discussion to arguing that Binding Conditions are actually a grammaticalized version of Accessibility theory. Here we only mention the possibility of using Accessibility theory to motivate the pragmatic principle of 'Avoid Pronoun', assumed by all proponents of GB. Chomsky (1981) attributes the fact that (15a) does not have a co-referential reading (as opposed to (15b)) to an instruction to the speaker to avoid pronouns when co-reference can be achieved without them:

(15) (a) John bought a book for him to read.
(b) John bought a book to read.

Reinhart (1983, 1986) also considers sentences where bound anaphora could have been an alternative, and argues that in such cases if the speaker did not choose that alternative, the addressee does not get the co-reference reading. This is so because the addressee reasons that the speaker would have used bound anaphora had s/he meant a co-reference reading. Bound anaphora, argues Reinhart (1983:167), 'is the most explicit way available in the language to express co-reference, as it involves dependency. So, when co-

reference is desired, this should be the preferred way to express it'. This principle, which Reinhart derives from Grice's Maxim of Manner, is very much in line with the claims about the correspondences between referring expression types and antecedent availability. Assuming that anaphors mark a higher degree of Accessibility than pronominals, preferring a 'free' pronoun over a bound anaphor (a reflexive, for example), where the latter could have been used, means avoiding the Accessibility marker reserved for the most highly accessible entities. Since the antecedent in bound anaphora is within the same governing category, using a lower Accessibility marker is misleading. The latter should be saved for cases where the antecedent is not that salient (or when the grammar forbids bound anaphora). The problem, however, with the 'Avoid Pronoun' principle is that it only dictates preferences in one direction (favouring higher Accessibility markers), relegating preferences in the opposite direction to 'marked' uses. My claim is that such preferences are not necessarily 'marked'. They simply arise under circumstances where mental representations are assumed to be less accessible to the addressee.[9]

The anaphor/pronominal contrast manifests itself in many ways. Clark & Haviland (1977) present data showing that reflexives (anaphors) took less time to interpret than regular pronouns. Enç (1983) shows that Turkish has three forms potentially usable for third person pronouns: *o*, *kendisi* and *kendi*+inflexions; *o* is a Mid Accessibility marker, *kendi*+inflexion a reflexive anaphor, and *kendisi*, she argues, is an intermediate case. It behaves like a pronominal (it is free in its governing category), and it can refer across a sentence boundary, but it cannot refer deictically (similar, probably, to unstressed pronouns in English – see below). Similar points have been made about zero pronouns in Chinese and Arabic. Li & Thompson (1979) claim that in Chinese, 3rd person pronouns are optional. Still, there is a tendency to use zero pronouns when the clause containing the antecedent is more 'conjoinable', i.e. semantically or pragmatically more related (thus making the antecedent more available). Eid (1983) argues that in Arabic, subject pronouns are generally omitted. If an embedded clause contains a pronoun, it will most likely be interpreted as non-co-referential with the main clause subject (this is a very similar point to that made by Chomsky with reference to (14) above). Claims in this spirit have been made with respect to Turkish and Japanese as well (Enç, 1983).

Gundel (1980) argues that zero pronouns are always bound by topics, but she still discerns differences among languages with respect to the freedom of zero pronoun occurrence. Chinese and Japanese are among the freest. Russian, Arabic, Hebrew, Spanish, Turkish and Latin are examples of intermediate languages, and English, French and German are among the

[9] For more discussion of the 'Avoid Pronoun' principle, see Levinson (1987) and Smith (this volume).

most restrictive ones. These contrasts are important to note, for they have implications for the relative Accessibility predicted by the theory for the neighbouring referring expressions. Thus, pronouns in languages which also have free zero options should be lower on the Accessibility scale, and should then serve for more marked cases. This has in fact been noticed by Eid (1983) for Arabic, by Enç (1983) for Turkish and Japanese, and by Bates *et al.* (1980) for Italian. For example, both English and Japanese have third person pronouns. However, only in Japanese is there in addition an understanding that a personal relationship exists between the speaker and the referent of the pronoun (see Hinds, 1978).[10] This, perhaps, renders the pronouns in Japanese lower Accessibility markers closer to Mid Accessibility markers. Similarly, Longacre (1979) mentions a few languages where pronouns are restricted to paragraph scope (Gurung of Nepal and Sanio-Hiowe of New Guinea). This is not the case in English, Hebrew, and other languages. The former are obviously higher Accessibility markers than the latter. In what follows I will therefore compare High Accessibility markers only within and not across languages. In fact, I will mainly use examples from English.

2.3.2. *Stressed pronouns*: *High Accessibility markers on the way to Mid Accessibility markers*. Most pronouns are unstressed, and in fact the pronouns I have referred to so far are all unstressed. However, note the examples in (16) (where small capitals indicate a stressed word):

(16) (a) Jane$_i$ kissed Mary$_j$ and then SHE$_j$ kissed Harry.
(b) Jane$_i$ kissed Mary$_j$ and then Harry kissed HER$_i$.

If we compare (16) with (17), where the pronouns are unstressed, we get a different co-reference pattern. Thus:

(17) (a) Jane$_i$ kissed Mary$_j$ and then she$_i$ kissed Harry.
(b) Jane$_i$ kissed Mary$_j$ and then Harry kissed her$_j$.

Various explanations have been suggested for this shift in pronoun interpretation due to stress differences (see Schmerling, 1976; Bardovi-Harlig, 1983), but a general explanation is possible (See Stenning, 1978; Solan, 1983), along the lines suggested by Accessibility theory. Stressed pronouns mark a lower Accessibility basis. They therefore signal the addressee to reject the unmarked co-reference interpretation, reserved for Øs and pronouns.[11] Now, such differences in Accessibility need not be marked

[10] This is reminiscent of the use of proper names in English.
[11] The same is true for full NPs. Note:
John$_i$ thought that JOHN$_i$ / *John$_i$ should put the cat out.
Without heavy stress we get disjoint reference, which the stress serves to cancel (Stenning, 1978: 170).

by word stress. Hebrew, for example, prefers to use a demonstrative where English uses a stressed pronoun above. Other languages may have different lexical items for the two pronoun functions. Gundel (1980) presents examples from Polish and Irish, where third person pronouns are not differentiated as in English or Hebrew. English unstressed pronouns correspond to Polish *go*, Irish *e*; stressed pronouns, including those accompanied by a pointing gesture, are lexically distinct (though obviously morphologically related): Polish *jego*, Irish *eisen*.

However, there is another class of stressed pronouns. When a pronoun is stressed and accompanied by some paralinguistic feature – a pointing gesture of some sort – the pronoun refers deictically, not requiring a linguistic antecedent (see Lyons, 1977; Bardovi-Harlig, 1983, and others). Thus (18) should be accompanied by a nod or other gesture in order to be acceptable.[12]

(18) Who's HE?

In other words, stressed pronouns (already lower Accessibility markers than unstressed pronouns) are transformed into fully-fledged Mid Accessibility markers when accompanied by special pointing gestures. This is only natural, given the gradedness of Accessibility. In fact, the Hebrew distal demonstrative is a clear development of a High Accessibility marker (*ha*+*hu*, literally 'the he/it').

In the next and final section, I should like to suggest a few criteria by which to determine the relative Accessibility height of referring expressions in general. I will then try to incorporate the three types into one graded Accessibility scale.

3. THE ACCESSIBILITY HIERARCHY

We have claimed that the three types of referring expressions (High, Mid and Low Accessibility markers) are a simplification of a more complex Accessibility system, which correlates with various statuses of Accessibility of mental representations. Low Accessibility marked entities are often those stored in long-term memory, while High Accessibility marked entities are normally those held in short-term memory. Extremely high Accessibility marked entities are probably more often than not those currently held in the phonological buffer. Then, using the same notion of Accessibility or saliency, we were able to draw further distinctions within each of the major referential categories. Just as Low Accessibility markers correlate with low accessibility in comparison to Mid and High Accessibility markers, so lower Accessibility

[12] However, when the referent is extremely salient, unstressed pronouns may be used on a physical basis, as in:

(i) [In the middle of a meeting someone suddenly leaves]: He must be upset.

markers WITHIN each category are associated with lower Accessibility. Thus, Low Accessibility markers, across the whole scale correlate with long(er)-term memory, and with a status of relatively low Accessibility for the entities (no matter where they are stored). High Accessibility markers, on the other hand, correlate with short(er)-term memory, and with a higher status of Accessibility for the entities, again, no matter where they are stored. In Table 7 below I have tried to map various referring expressions as to their relative degree of Accessibility. We stick to the three-way distinction in Table 7, mainly to represent unmarked initial retrievals.

The vertical line indicates unmarked memory storage type, which normally corresponds to unmarked context type. Top forms tend to retrieve from long-term memory, and bottom forms from short-term memory. The horizontal line indicates degree of Accessibility within each memory type. Higher Accessibility appears on the right, lower Accessibility on the left. Hence the lowest Accessibility markers appear on the left-hand side of the top of the table, where long-term memory and lowest accessibility converge (*Joan Smith*, *the president*, for example). Highest Accessibility markers appear on the right-hand side of the bottom of the table, where short-term memory and high Accessibility meet (*herself*, or ∅, for example).

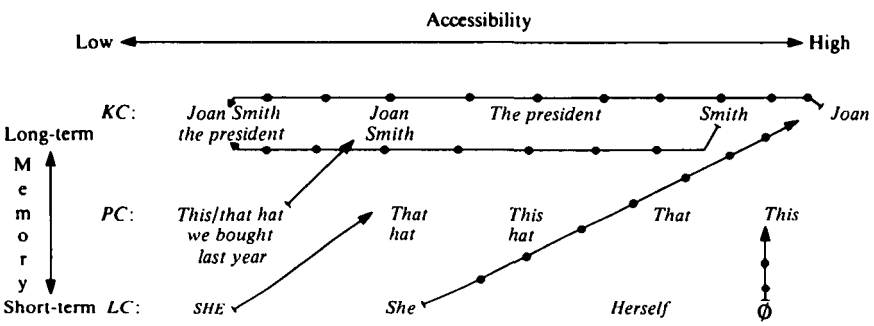

Table 7
Initial Accessibility Marking (KC = General knowledge context, PC = Physical context, LC = Linguistic context)

Note that we only have arrows going up. This is due to a universal constraint on historical changes within givenness markers, referring expressions among them, which we cannot go into here (but see Ariel (1985) and (to appear)). Regular arrows indicate a grammaticalized change in Accessibility rate. Thus, stressed pronouns, for example, belong both in the High Accessibility block (linguistic context) and in the Mid Accessibility block (physical context). The same goes for modified deictics such as *that hat we bought last year*. Dotted arrows signal *ad hoc*, extremely context-dependent uses of High Accessibility markers to refer to what are normally

considered inaccessible entities. The inherent difference between the dotted arrows and the regular arrows is that regular arrows are potential historical grammaticalization changes, whereas dotted ones are not. The dotted arrows upgrade High Accessibility markers which are not good candidates for a permanent Accessibility change, the reason being that they are not lexically informative enough. And this brings us to characteristics which very often distinguish between Higher and Lower Accessibility markers. In fact, in Ariel (forthcoming), I make the claim that three criteria are operative in the linguistic codification systems of the world languages when translating the cognitive concept of Accessibility into linguistic dress: Informativity, Rigidity and Attenuation.

In conformity with the principle of Relevance, the lower the Accessibility marker, the more lexical information it normally incorporates. Whereas at the right end of the bottom line we have $\emptyset$, we have quite a bit of information in expressions like *Joan Smith*, *the president*. In between are: *this/that hat*, *this/that* (+pointing or nodding), and even *she* (where at least person and number are marked). It is only natural that lower Accessibility markers should contain more information. After all, they are information retrieval devices which refer addressees to less accessible sources. In order to be effective, the marker has to be a good search guide, even if it takes time and effort to process the long expression. The more information it imparts the better retriever it is – other things being equal, of course. In contradistinction to the above, higher Accessibility markers are semantically emptier. Since the entities referred to are very salient, the guidelines need not specify too much. A detailed description would in this case take up a lot of processing effort with no extra benefit to the addressee. Closely related is the criterion of Rigidity. Proper names, being (more) rigid, are thus distinguished from all other markers, but additional distinctions can be motivated by this criterion. Last names (in western societies) are more numerous, hence less ambiguous than first names. They are then more Rigid, signalling lower Accessibility.

Degree of attenuation and/or phonological size can distinguish between stressed and unstressed forms (of pronouns as well as full NPs), and between shorter and longer expressions which are not necessarily distinguished with respect to the information they transmit. Thus, based on textual counts in Hebrew, for example, I suggest (see Ariel (forthcoming)) that full pronouns, cliticized pronouns and Agreement markers which are equally informative (each marking gender, number and person) signal different degrees of Accessibility, as dictated by their descending phonological size. An additional difference found among marker types is that High Accessibility markers are also much more formally constrained in distribution, even though they may be extremely frequent. In many cases they are syntactically constrained. Hardly any syntactic constraints pertain to the relationship between antecedent and Mid Accessibility or Low Accessibility markers.

As far as antecedent properties characteristic of each marker are concerned, most High Accessibility markers refer to unmarked, contextually

salient entities (especially discourse topics). Low Accessibility markers refer to more marked, less accessible antecedents. Furthermore, only entities which are easily and uniquely determined can be retrieved using a High Accessibility marker. In ambiguous contexts, where more than one referent 'competes' as the proper interpretation, a lower Accessibility marker must be employed. Accepting Sperber & Wilson's (1982, 1986) account of context, according to which there is an initial context, typically composed of information derived from the previous utterance(s), and various possible extensions involving information from earlier utterances in the discourse, from the physical environment or from general encyclopedic knowledge as the most extended type of context, we can say that High Accessibility markers call for a minimal contextual extension, whereas Low Accessibility markers call for a maximal contextual extension. As a result, High Accessibility markers contribute to the Relevance of the utterance at a lesser cost. (First retrieval) referential acts which refer the addressee to her/his general knowledge, are more costly in terms of processing, but they contribute more to long-term implications. Preference should then be given to High Accessibility markers, whenever contextual implications are unimpaired.

We should also keep in mind that the Accessibility scale is, to some extent, language-specific. The degree of Accessibility of synonyms may vary from one language to another. Each language assigns a slightly different slot to its referring expressions, mainly on the basis of the other referring expressions available in that language. Thus, the space on the scale occupied by a referring expression (i.e. the limits on its maximal and minimal Accessibility) will be determined, at least partly, by its neighbouring markers. For example, languages with a free distribution of zero pronouns (e.g. Chinese, Japanese) have less use for lexical pronouns. These will then refer to more marked antecedents. Languages with no special definiteness marker (e.g. Slavic languages) have more use for demonstrative pronouns. Communities without last names (e.g. the Amish) probably have more use for definite descriptions instead (*X, the daughter of Y*, for example). In other words, the size of the slot on the scale is dependent on other competing forms.

Finally, I should like to suggest (19) as the most precise scale; it is arranged by degree of Accessibility only, with no reference being made to context-type. In principle, (19) says we can use any marker initially, as well as in subsequent retrievals, provided the degree of Accessibility and/or the markedness of the referent matches that signalled by the specific marker. Now, whereas Table 7 represents unmarked initial retrievals more clearly, the following is not only more precise on the whole, it is certainly the more relevant Accessibility scale for subsequent retrievals.[13]

[13] *Ad hoc* differences may alter the order in (19), as well as in Table 7. For very famous personalities (Reagan, for example), a last name may be a sufficient retriever from general knowledge, even more so than a definite description. This is more rarely true for the first name, though Israeli Golda would be an example.

(19) *Joan Smith, the president* > *Joan Smith* > *The president* > *Smith* > *Joan* > *That/this hat we bought last year* > *That hat* > *This hat* > *That* > *This* > SHE > *she* > *herself* > Ø

While (19) was constructed to account for English data, the Accessibility claim is universal, in that the same hierarchy is predicted to hold for all languages, provided they possess the forms in question. The only differences which are allowed pertain to degree of markedness, i.e. the relative freedom of occurrence of the various expressions in different languages. Indeed, based on a cross-linguistic study, Givón (1983) proposed a gradation of referring expressions (interwoven with syntactic configurations) very much in the spirit of (19). (20) reproduces item (10) in Givón (1983).[14]

(20) *Most continuous/accessible topic*
↑ zero anaphora
unstressed/bound pronouns or grammatical agreement
stressed/independent pronouns
R-dislocated DEF-NPs
neutral-ordered DEF-NPs
L-dislocated DEF-NPs Y-moved NPs
('contrastive topicalization')
cleft-focus constructions
↓ referential indefinite NPs
Most discontinuous/inaccessible topic

Though the theoretical status Givón attributes to the above categories is not sharply defined (it is never made clear whether only topics, or any referring expressions were considered), and though I believe that the syntactic categories alluded to above are to be differently accounted for (using the Keenan–Comrie NP Accessibility Hierarchy), the data collected by Givón and his associates confirm the Accessibility claim, as I view it. They thus find that in quite a few unrelated languages (English, Ute, Early Biblical Hebrew, colloquial Spanish, Hausa and Chamorro), distance and antecedent competition determine the choice of a referential form along the scale in (19). More discussion of this issue is to be found in Ariel (forthcoming).

Summing up, the treatment of referring expressions should not be separate for expressions which serve as antecedents as opposed to those which are always anaphoric. Nor are referring expressions divided into ones that presuppose and ones that do not. Instead, I have argued that referring

[14] Givón's theory on topic continuity is very much in the spirit of Accessibility theory, though I was not aware of it when I first published such claims (Ariel, 1985). This is mainly due to a most unfortunate misnomer: topic continuity, rather than referent accessibility continuity. This misnomer may in fact be more than a wrong name choice, but even if it is, and the expressions counted were only topics, the findings still corroborate the Accessibility claim.

expressions in all languages are specialized as to the degree of Accessibility they mark. As to presuppositionality, I would claim that whether or not an existential presupposition is attached to a particular occurrence of a referring expression is determined by discourse factors, most notably, by considering whether or not it is an initial retrieval. Only those entities initially retrieved from the 'general knowledge' or from the physical context are assumed by the speaker to exist. Entities retrieved not for the first time, that is those that rely on a previous linguistic mention, are not guaranteed to exist. In other words, presupposition is a derived, rather than a primitive notion. NP forms are categorized in the mental lexicon as to the degree of Accessibility they mark, and not, as is commonly assumed, with the feature [±presuppositional].[15]

REFERENCES

Ariel, M. (1983). Linguistic marking of social prominence: the Hebrew *mi she* introducer. *JoP* 7. 389–409.

Ariel, M. (1985). *Givenness marking*. Ph.D. thesis submitted to Tel-Aviv University.

Ariel, M. (forthcoming). *Accessing NP antecedents* (tentative title). London: Croom Helm.

Bates, E., Kintsch, W., Fletcher, C. R. & Giuliani, V. (1980). Recognition memory for surface forms in dialogue: explicit vs. anaphoric reference. In Kreiman, J. & Ojeda. A. E. (eds). 41–49.

Caramazza, A., Grober, E., Garvey, C. & Yates, J. (1977). Comprehension of anaphoric pronouns. *JVLVB* **16**. 601–610.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Clancy, P. M. (1980). Referential choice in English and Japanese narrative discourse. In Chafe, W. L. (ed.), *The pear stories*, vol. 3 in Freedle, R. O. (ed.), *Advances in discourse processes*. Norwood; Ablex. 127–202.

Clark, H. H. & Marshall, C. R. (1981). Definite reference and mutual knowledge. In Joshi, A. K. *et al.* (eds). 10–63.

Clark, H. H. & Sengul, C. J. (1979). In search of referents for nouns and pronouns. *Memory and cognition* 7. 35–41.

Eid, M. (1983). On the communicative function of subject pronouns in Arabic. *JL* **19**. 287–303.

Enç, M. (1983). Anchored expressions. MS, University of California.

Fillmore, C. J. (1975). *Santa-Cruz Lectures on Deixis 1971*. Indiana University Linguistic Club.

Frege, G. (1982). On sense and reference. In Zabeeh, F. *et al.* (eds). 117–140.

Gazdar, G. J. M. (1979a). A solution to the projection problem. In Oh, C. K. & Dinneen, D. (eds). 57–90.

Gazdar, G. J. M. (1979b). *Pragmatics: implicature, presupposition and logical form*. New York: Academic Press.

Giora, R. (1985). *Informational function of the linear ordering of texts*. Ph.D. Thesis, Tel-Aviv University.

Givón, T. (ed.) (1979). *Syntax and semantics 12: discourse and syntax*. New York: Academic Press.

Givón, T. (ed.) (1983). *Topic continuity in discourse: a quantitative cross-language study*, Amsterdam: John Benjamins.

Greenberg, J. H. (1978). How does a language acquire gender markers? In Greenberg, J. H., Ferguson, C. & Moravcsik, E. (eds), *Universals of human language*, Stanford: Stanford University Press.

[15] For an analysis of presuppositions, see Ariel (1985) and (forthcoming).

Grosz, B. J. (1981). Focusing and description in natural language dialogues. In Joshi, A. *et al.* (eds), 84–105.
Gundel, J. K. (1980). Zero NP-anaphora in Russian: a case of topic prominence. In Kreiman, J. & Ojeda, A. E. (eds). 139–146.
Halliday, M. A. K. & Hasan, R. (1976). *Cohesion in English*. London: Longman.
Harris, M. B. (1980). The marking of definiteness: a diachronic perspective. In Traugott, E. C., Labrum, R. & Shepherd S. (eds). *Papers from the 4th International Conference on Historical Linguistics*. Amsterdam: John Benjamins. 75–86.
Hawkins, J. A. (1974). *Definiteness and indefiniteness*. Ph.D. Thesis, Cambridge.
Hawkins, J. A. (1984). A note on referent identifiability and copresence. *JoP* **8**. 649–660.
Hinds, J. (1978). Anaphora in Japanese conversation. In Hinds, J. (ed.), *Anaphora in discourse*. Alberta: Linguistic Research.
Joshi, A. K., Webber, B. L. & Sag, I. A. (eds). (1981). *Elements of discourse understanding*. Cambridge: Cambridge University Press.
Kempson, R. M. (1975). *Presupposition and the delimitation of semantics*. Cambridge: Cambridge University Press.
Kintsch, W. (1970). *Learning and conceptual processes*. New York: John Wiley.
Kreiman, J. & Ojeda, A. E. (eds) (1980). *Papers from the parasession on pronouns and anaphora*. Chicago: Chicago Linguistic Society.
Levinson, S. C. (1987). Pragmatics and the grammar of anaphora: a partial pragmatic reduction of binding and control phenomena. *JL* **23**. 379–434.
Li, C. N. & Thompson, S. A. (1979). Third-person pronouns and zero-anaphora in Chinese discourse. In Givón, T. (ed.). 311–335.
Liberman, M. (1973). Alternatives. In Corum, C., Smith-Stark, T. C. & Weiser, A. (eds). *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society. 346–355.
Linde, C. (1979). Focus of attention and the choice of pronouns in discourse. In Givón, T. (ed.) 337–354.
Loftus, E. F. (1972). Nouns, adjectives and semantic memory. *J. Experimental Psychology* **96**. 213–215.
Loftus, E. F. (1974). Reconstructing memory. The incredible eye-witness. *Psychology Today* **8**. 116–119.
Loftus, E. F. (1975). Leading questions and the eye-witness report. *Cognitive Psychology* **7**. 560–572.
Longacre, R. (1979). The paragraph as a grammatical unit. In Givón, T. (ed). 115–134.
Lyons, J. (1977). *Semantics*, vols 1 and 2. Cambridge: Cambridge University Press.
Marslen-Wilson, W., Levy, E. & Komisarjevsky Tyler, L. (1982). Producing interpretable discourse: the establishment and maintenance of reference. In Jarvella, R. J. & Klein, W. (eds), *Speech, place and action*. Chichester: Wiley.
McCawley, J. D. (1979). Presupposition and discourse structure. In Oh, C. K. & Dinneen, D. (eds). 371–388.
Oh, C. K. & Dinneen, D. (eds) (1979). *Syntax and semantics*, vol. 11. *Presupposition*. New York: Academic Press.
Prince, E. F. (1978a). On the function of existential presupposition in discourse. In Farkas, D. *et al.* (eds), *Papers from the 14th Regional Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society. 362–376.
Prince, E. F. (1981a). Toward a taxonomy of given-new information. In Cole, P. (ed.), *Radical pragmatics*. New York: Academic Press. 223–255.
Prince, E. F. (1981b). On the inferencing of indefinite *this* NP's. In Joshi, A. K. *et al.* (eds), 231–250.
Reinhart, T. (1981). Pragmatics and linguistics: an analysis of sentence topics. *Philosophia* **27**. Special issue on pragmatic theory. 53–94.
Reinhart, T. (1983). *Anaphora and semantic interpretation*. Worcester: Croom Helm.
Reinhart, T. (1986). Center and periphery in the grammar of anaphora. In Lust, B. (ed.), *Studies in the acquisition of anaphora*. Dordrecht: Reidel. 123–150.
Russell, B. (1919). *Introduction to mathematical philosophy*. London: George Allen & Unwin. 167–180.
Sanford, A. J. & Garrod, S. C. (1981). *Understanding written language*. Chichester: Wiley.

Sanford, A. J. & Garrod, S. C. (1985). On the real-time character of interpretation during reading. *Language and cognitive processes* 1. 43–59.
Schiffman, R. J. (1984). The two nominal anaphors *it* and *that*. In Drogo, J. *et al.* (eds), *Papers from the 20th Regional Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society. 344–357.
Smith, N. V. (this volume). Principles, parameters and pragmatics. *JL.* **24**: 189–201
Solan, L. (1983). *Pronominal reference: child language and the theory of grammar*. Dordrecht: Reidel.
Sperber, D. & Wilson, D. (1982). Mutual knowledge and relevance in theories of comprehension. In Smith, N. V. (ed.), *Mutual knowledge*. London: Academic Press. 61–87.
Sperber, D. & Wilson, D. (1986). *Relevance*, Oxford: Blackwell.
Stenning, K. (1978). Anaphora as an approach to pragmatics. In Halle, M., Bresnan, J. & Miller, G. A. (eds). *Linguistic theory and psychological reality*. Cambridge MA: M.I.T. Press. 162–200.
Strawson, P. F. (1956). On referring. In Zabeeh, F. *et al.* (eds). 159–192.
Strawson, P. F. (1964). Identifying reference and truth values. In Zabeeh, F. *et al.* (eds). 193–216.
Yule, G. (1981). New, current and displaced entity reference. *Lingua* **55**. 41–52.
Zabeeh, F., Klemke, E. D. & Jacobson, A. (eds) (1974). *Readings in semantics*. Urbana: University of Illinois Press.

SOURCES

Berlovich, Y. (ed.) (1984). *Sipurey Nashim*. Israel: Tarmil.
Haaretz. A daily paper.
Paley, G. (1956/1986). *The little disturbances of man*. London: Virago (reprint).
Trouble on the set: an analysis of female characters on 1985 television programs. National Commission on Working Women (U.S.).
Tucker, J. E. (1985). *Women in nineteenth-century Egypt*. Cambridge: Cambridge University Press.
Young, E. H. (1934/1985). *The curate's wife*. London: Virago (reprint).
Yediot Achronot. A daily paper.