This paper presents necessary and sufficient conditions for the use of demonstrative expressions in English and discusses implications for current discourse processing algorithms. We examine a broad range of texts to show how the distribution of demonstrative forms and functions is genre dependent. This research is part of a larger study of anaphoric expressions, the results of which will be incorporated into a natural language generation system.

1. INTRODUCTION

The primary purpose of this paper is to present a set of necessary and sufficient conditions for the use of demonstrative expressions in English, based on a corpus of naturally occurring discourse from a variety of spoken and written genres. We propose a comprehensive set of constraints on demonstrative use and suggest how they can be incorporated into a computational processing model which integrates the local centering and global focusing aspects of discourse structure. Finally we show how our proposed algorithm for demonstratives can account for stressed pronouns as well.

Existing computational work on demonstratives has been based on examples from only three genres: experimentally elicited apartment descriptions (Linde 1979), technical dialogues (Reichman-Adar 1984), and scientific textbooks (Sidner 1983). Testing computational algorithms against multiple genres of natural discourse is important, especially given the universal scope of current frameworks (cf. Groz and Sidner 1986). We have chosen to systematically examine texts from a broad range of genres, which vary in spoken versus written medium, number of participants, degree of pre-planning, and formality of situation. These genres include informal conversation, partially spontaneous televised discussion, newspaper articles, and planning and technical documents.

2. REVIEW OF THE LITERATURE

Demonstratives have generally been considered to be one type of deictic expression. Most of the linguistic literature on deixis has concentrated on classifying deictic expressions according to various parameters such as the linguistic or extralinguistic status of the referent. For example, Buhler (1982) distinguishes between demonstratio ad oculos, anaphora, and deixis at phantasma, whose intended referents are respectively found in the extralinguistic context, the linguistic context, or in memory or imagination. Similarly, Fillmore (1975) distinguishes gestural, anaphoric and symbolic uses of deictic expressions; and Lakoff (1974) distinguishes spatio-temporal, discourse, and emotional deixis. Such distinctions have been useful as a basis for descriptive classification of various uses of demonstrative expressions, but we do not feel that they are relevant for an explanatory account of demonstrative function and so shall not elaborate further on them here. For the same reason we also will not discuss attempts to show that discourse deixis and/or anaphora is derived from or less prototypical than spatio-temporal deixis (cf., Lyons 1977; Fillmore 1982). Our account is more compatible with the view that deictic expressions signal a change in focus of attention while anaphoric expressions signal focus continuation (cf., Ehlich 1982; Bosch 1983). It is this view which most closely reflects the assumptions underlying computationally explicit models of focus-constrained discourse processing (e.g., Linde 1979; Reichman 1985; Groz and Sidner 1986).

We turn now to a presentation of specific claims about demonstratives which have been made in the literature.

Focus shift.

a. That but not it may be used to refer to an item no longer in focus (Linde 1979). That marks the end of a discourse section (Reichman 1984, 1985). This but not that signals focus movement, except when this occurs in the scope of a quantifier or when its head noun is not identical to the head noun of an existing immediate focus (Sidner 1983).

b. That but not it may be used to refer to a preceding statement taken as a fragment (Linde 1979).

c. Stressed that must change focus (Isard 1975).

Contrast.

d. That may be used to indicate contrast, even for items currently in focus (Linde 1979). When this and that are used together contrastively ('copresent' use), this specifies the primary, continuing focus and that specifies a secondary, temporary focus (Sidner 1983).

Proximity.

d. This is used to denote objects relatively close to the speaker and that for objects relatively far from the speaker or relatively close to the hearer. (Fillmore 1975; Halliday and Hasan 1976; Lakoff 1974).

e. Both this and that may corefer with elements in the preceding linguistic context, but only this may corefer with elements in the following linguistic context. (Fillmore 1975, 1982; Halliday and Hasan 1976; Lakoff 1974).

f. Both this and that can be used to comment upon a speaker's own prior remarks, with this often signifying greater speaker involvement or continued discussion. Only that may be used to comment upon the remarks of another speaker (Halliday and Hasan 1976; Lakoff 1974; Sidner 1983).

g. The interval denoted by a determiner this used with a time expression includes the speech time. (Fillmore 1975). That tends to be associated with a past time referent and this with one in the present or future. (Fillmore 1975; Halliday and Hasan 1976).

Special effects.

a. Demonstratives often have subtle emotional effects such as solidarity, distancing, sympathy, anger, irony, etc. (Lakoff 1974, Fillmore 1982; Halliday and Hasan 1976).

b. Unstressed determiner this has an indefinite specific use in colloquial narrative speech, (Lakoff 1974); shared attention and experience (Halliday and Hasan 1976); topicality (Prince 1981a).

c. Unstressed determiner that phrases have a first-mention use for expected shared referents. (Wald 1983).

3. CONSTRAINTS ON DEMONSTRATIVE USE

The basic concepts which we take to be crucial for any adequate description of demonstratives are ones which are central to a theory of reference in general. These are concepts such as topic, focus and various types of givenness. Since these concepts concern the cognitive status of a referent, definitions have often been
for computational implementation. On the other hand, because of the complexity of these concepts, attempts to furnish more precise operational definitions, e.g. on the basis of surface linear order or grammatical relations, have failed to capture their cognitive basis. The result has been a terminological and conceptual conflation in the literature which has led computational linguists to create new constructs such as focus and center, in some cases without relating them to similar linguistic concepts (Hajicová 1987 is a notable exception). Our aim here is to characterize as precisely as possible what the relevant concepts for pronominal fillers are and how they relate to one another. In a later section we will attempt to show how they relate to more operational constructs proposed in the computational literature.

Definition of givenness. As has often been pointed out, the basic notion which determines appropriateness of a given referential expression is the status of the referent of that expression vis-à-vis a cognitive state of the addressee (cf. Chafe 1976). In the weakest case, the speaker expects the addressee to understand what type of entity is being described, but not to uniquely identify the entity in question. Such entities are typically referenced with indefinite noun phrases. The most basic distinction in demonstrative function is between the indefinite use of the determiner this, as exemplified in (1) and all other uses of both determiner and pronominal demonstratives, which are definite. As has been pointed out by a number of researchers, indefinite this occurs only in casual, impromptu discourse. This observation was confirmed by our own study, which found instances of indefinite this only in the informal conversations.

(1) I couldn't sleep last night. My neighbors have this dog that kept me awake.

We will be primarily concerned here with definite noun phrases, where the speaker expects the addressee to uniquely identify the entity in question vis-à-vis the description in question. The referents of such expressions have often been characterized as 'given'. However, as pointed out by Gundel (1978a, 1978b) and Prince (1981b), givenness is not a unitary concept. There are different senses of givenness each of which is relevant to the form of referring expressions, but in different ways.

Identifiability. In the weakest sense of givenness, the speaker expects the addressee to uniquely identify the referent, but the basis for the identification is irrelevant. Not only can it be linguistic or extralinguistic, based on entities in immediate or long-term memory, but it need not be based on any previous shared experience at all (cf. Hawkins 1978). The basis for the identification may be encoded in the form itself, as in (2).

(2) I couldn't sleep last night. The neighbor's dog kept me awake.

This type of givenness, which we refer to as identifiability, is both necessary and sufficient for the use of definite articles.

Shared familiarity. Most identifiable entities are identifiable because of some shared experience between speaker and addressee; again this may be linguistic or extralinguistic, based on long-term or long-term memory associated with shared cultural or personal experience. It is this status which we claim is necessary for the use of demonstrative expressions (with the exception of indefinite this). Thus, (3), unlike (2), is licit only if the addressee has prior knowledge of the dog in question.

(3) I couldn't sleep last night. That dog next door kept me awake.

Activation. Entities which are familiar on the basis of presence in the immediate discourse context (either linguistic or extralinguistic) are not only shared but 'in awareness.' This sense of givenness, which we refer to as 'activated', (cf. Chafe 1976, Gundel 1978, Hajicová and Vrbová 1982) is necessary for pronominal reference, including pronominal demonstratives. Thus, that in (4) could only refer to the backing of a dog if this had been activated by the immediate discourse context:

(4) I couldn't sleep last night. That kept me awake.

Activation is also a necessary condition for determiner this. Demonstrative this, both pronominal and definite determiner, has the additional condition that it not only be activated but speaker-activated, either linguistically or extralinguistically by virtue of its inclusion in the speaker's context space, as in (5):

(5) A: Have you seen the neighbors' dog?
   B: Yes, and that dog kept me awake all night.

In focus. Finally, the most highly activated entities are not only in the speaker's and hearer's awareness but are the center of attention at a particular point in the discourse. We refer to this status as 'in focus.' Entities in focus always include at least the topic of the sentence as well as any higher level discourse topics which may not be overtly represented in the sentence itself. Under certain conditions, they may also include other elements such as the reference of the linguistic focus. Thus a shift in focus always entails a shift in topic but not vice versa. The status 'in focus' is a necessary condition for anaphora and also for zero anaphora (cf. Gundel 1978b).

The relationship between the various (definite) referential statuses and the forms that correlate with them is shown in (6). These statuses are in a unidirectional implication relation such that any status on the hierarchy implies all statuses higher on the hierarchy but not vice-versa:

(6) GIVENNESS HIERARCHY:

in focus \Rightarrow activated \Rightarrow shared \Rightarrow identifiable

4. CRITIQUE OF CLAIMS IN THE LITERATURE.

Focus Shift. Observations regarding the focus shift function of demonstratives follow naturally from the theory outlined above. Given two additional, but uncontroversial assumptions -- that pronominal it is necessarily un stresses and that the overwhelming majority of definite noun phrases are not only uniquely identifiable, but shared. As noted above, the primary distinction between demonstrative pronouns and unstressed personal pronouns is that the referents of the latter must not only be activated, but in focus. A pronominal whose referent is not currently in focus is necessarily stressed (cf. Hirschberg and Pierrehumbert 1987). Since it is necessarily un stressed, the only third person singular pronominal available for focus shift are the demonstratives that and this. This accounts for the distribution of that versus it noted by Lüttke (1979) as well as the uses of that noted by Reichman. Use of that in referring to previous statements (also noted by both these authors) is just a special case of focus shift, since the focus of attention at the point after a statement is made is typically not the statement itself. However, our data does not support Sidner's (1983) claim that this but not that is used for focus movement. Examples illustrating the focus shift function of pronominal this and that include the following:

(7) K 1: And...so what he DID was...came in, set up the tree...
   2: and then he made wassail, with rum in it?
   A: And...made it in coffee cans and heated it on the stove in the graduate lounge.

A 4: Oh, gee.

K 5: And this was the solstice tree.

In (7), the topic and hence the focus of attention in K5 is the tree, which is activated by its mention in K1 thus licensing the
use of a pronominal. However, since it is not in focus in K2, K3, and A4, the reference to it in K5 constitutes a focus shift and thus requires a stressed demonstrative form. Note that since the tree was speaker activated, either *this* or *that* could be used.

(8) N 1: I like the poor dog who was buried six times in one day!

   K 2: Oh. That must, that must be a story that comes from the Second Minnesota history,
   3: because that didn't appear in the, in the di-
   4: ary, so it must have come from somewhere,

   In (8), the topic and hence focus of K2 is the story. Since the story is activated but not in focus in N1 (the focus of attention in N1 is stories that N likes), reference to the story in K2 constitutes a focus shift and thus requires a stressed demonstrative. Note that the story continues to be topic (in focus) in K3 and K4 and that in K3, an unstressed pronominal *it* could have been used instead of *that*.

(9) N 1: "Bob loves Mary",
   2: and someone else wrote "Mary loves Jim"
   3: and I wrote "Jim loves Bob"! (laughter)
   4: It was three different handwritings, three differ-
   ent people.

   K 5: Yeah, that's good.

(9) illustrates the use of *that* to refer to a previous statement or utterance. Since *that* refers to the topic of K5, the story activated in N1-4 (but not in focus), use of *that* constitutes a focus shift.

The focus shift function of determiner *that* (cf. Reichman 1984) can be explained as a consequence of Grice's maxim of quantity, specifically *don't be more informative than necessary* (cf. Grice, 1975). Since the overwhelming majority of definite noun phrases refer to entities that are shared, use of a demonstrative as opposed to the less restrictive definite article in most instances carries little additional information. Thus a demonstrative determiner is generally used only when the signaling of shared familiarity has some additional communicative function. This is the case when there is a shift in focus, as in (10) and (11), since a focus shift always entails a shift in topic and topics are necessarily shared (cf. Gundel 1985).

In the particular case of (10), there is not enough information encoded in the noun phrase itself to allow the addressee to uniquely identify the referent. Thus a demonstrative (as opposed to a definite article) is required in order to link the referent up with entities shared as a result of immediate discourse context. In (11b) on the other hand, the demonstrative functions simply to signal a focus shift and therefore a definite article would be equally appropriate.

(10) a. John, this speech was a magnificent triumph for the President. He showed he could stay awake for twelve whole minutes. He showed that he could speak every word of his teleprompter, even the long ones. But the speech doesn't have any chance of putting the scandal behind him, because the scandal is not about mistakes, as he said, and it's not about mismanagement, as the Tower Commis-

   b. And those issues remain. [McLaughlin 3/6/87].

(11) a. These incredibly small magnetic bubbles are the vanguard of a new generation of ultradense mem-

   b. These systems are extremely rugged;

   c. they are resistant to radiation and are nonvolatil-

Contrast. The contrastive function of demonstratives, like the focus shift function, is related to the fact that contrast is marked by stress and pronominal *it* cannot carry stress. Moreover, contrast may be just a special case of focus shift, since a contrastive noun phrase always brings into focus other entities with which it is being contrasted, as illustrated in (12):

(12) a. In both cycle steal mode and DPC, the attachment feature responds with a condition code.

   b. For commands that do not require interrupts (that is, commands executed under DPC), the condition code provides current device status information.

   c. For commands that require execution in cycle steal mode followed by an interrupt request, this first condition code provides information concerning acceptance of the command by the attachment feature.

   d. Upon interrupt servicing by the processor, the attachment feature provides a second condition code and an interrupt word.

Thus, the referent of this first condition code in (12c) is already in focus since it is also the topic of (12b). However, since the use of this noun phrase brings into focus contrasting condition codes (cf. *a second condition code* in (12d)) it is not only contrastive, but constitutes an implicit focus shift as well.

Proximity. The speaker-activated condition on *this* predicts correctly that both *this* and *that* are based upon a speaker's own prior remark as in examples (7) and (10) respectively, but only *that* can be used to comment upon the remarks of another speaker, as illustrated by example (9).

The same condition also explains why *this* is used for extralinguistic objects relatively close to the speaker and *that* for those relatively further away as in (13), and why the interval denoted by determiner *this* includes speech time, while *that* tends to be associated with some time prior to speech time as in (14). This is so because speaker-activated means not only linguistically activated but activated within the speaker's context space.

(13) N: *This* tastes like water. *This* tastes like THAT! Wait a minute--the stuff that's $1.99 for two quarts tastes a- the same as one that's $2.07 for a fifth.

(14) K: There he was that hairy bound from Bud-

   dapest/Never leaving us alone./Never have I ever

   known a ruder pest.

Special Effects. As seen in the diagram in (6), each of the referential statuses is also correlated as a necessary condition for a different type of definite reference. Since the statuses are implicationally related, reference with a particular form will generally imply appropriateness of reference with all forms associated with a status higher on the hierarchy, but not vice-versa. Thus, pronominal *that* in (13) may be replaced with *that wine and these systems* in (11b) can be replaced with *the systems*. However, not all cases of the *N* are replaceable with *that N*, as illustrated in (2) and (3); and not all cases of *that N* are replaceable with *that*, as illustrated in (3) and (4).

The maxim of quantity would dictate that speakers will use the strongest possible form, i.e. will signal the most information as is relevant regarding the givenness status of the entity in question. The same maxim predicts that speakers would not use a stronger form than necessary in a given situation, i.e. they will not signal more information than is appropriate. Violation in either direction will often result in a special effect or implicature, as in (15), where the use of *this* as opposed to the equally appropriate *the convey* an effect of solidarity:

(15) *Gov. D (from videotape):* I've got the energy to run this marathon, the strength to run this country, the experience to manage our government, and the values to lead our people. [McLaughlin, 3/20/87]

We already noted that use of a demonstrative determiner often has a special effect, such as signalling a topic shift, because virtually all definite NPs are also shared, and thus demonstrative determiners do not normally convey much more information than would be conveyed by the definite article. Similarly, in cases where the referent is not activated, determiner *that* acts as an overt signal to the addressee to search long term memory for the referent, as in (16):

(16)
(16) A: You can get those little sugar finger jubes that tell you when to water it. They're only $10 or something.

5. GENRE DIFFERENCES

The distribution of demonstrative types differs according to genre, as seen in Table 1. (See Hockenbury, forthcoming, for more on these genres.) Thus, for example, the relative percentage of pronominal that appears to be greater in multi-participant oral communication than in written communication. The casual conversations and electronic specification documents represent the respective extremes on the scale.

| GENRE                  | that | this | Total |
|------------------------|------|------|-------|
| Spoken Genres          |      |      |       |
| Casual Conversations   | 12   | 14   | 26    |
| Television Discussion  | 20   | 22   | 42    |
| Written Genres         |      |      |       |
| Newspaper Articles     | 3    | 4    | 7     |
| Planning Documents     | 3    | 5    | 8     |
| Specification Document | 5    | 10   | 15    |

Table 1. Demonstratives in five genres

5.1 Casual conversations. The casual conversations were taped for purposes unrelated to linguistic analysis; participants are family members reunited for holiday gatherings.

| Discourse topic | that | this | Total |
|-----------------|------|------|-------|
| Same Speaker    | 60   | 75   | 135   |
| Other Speaker   | 30   | 70   | 100   |
| Extralinguistic | 6    | 6    | 12    |
| Reminders       | 10   | 20   | 30    |
| Indefinite      | 10   | 15   | 25    |
| Time            | 10   | 15   | 25    |
| Forward         | 5    | 5    | 10    |
| Total           | 130  | 220  | 350   |

Table 2. Demonstratives in casual conversations

Other Speaker’s Remarks. A large proportion of pronominal that in the casual conversations is due to its use for reference to an immediately preceding contribution of another conversational participant. For an example see (9) above.

(17) N: Is this your car?

(18) K: What is that, Beethoven what, on that t-shirt.

Remind that. Since determiner that does not require activation of the referent, but only that it be shared, it can be used to refer to entities that are not present in the immediate discourse context, as in (19).

(19) K: I realized something that seems significant to me about George, that in the, in the fall, he, as everyone else, he wears...

As noted above, that N is sometimes used as an explicit signal to the hearer to search memory for the referent. One would expect such signals to be most common in interactions between individuals with shared personal experiences, so it is not surprising that they would occur frequently in casual conversations between family members. Typically such phrases include a relative clause specifying additional information to aid in the search and are often embedded in a request for confirmation that the referent has indeed been located. Since the referent is shared but not activated, it frequently occurs in left-dislocated constructions which have the function of introducing or reintroducing a topic into the discourse. (cf. Kerman and Schieffelin, 1976)

5.2 Television Discussions. The televised discussion was a videotaped episode of The McLaughlin Group (initial transcript obtained from the Federal News Service). This genre is similar to the casual conversations in being spoken and multi-participant, but differs in degree of formality and spontaneity and in awareness of an audience. Four journalists participate in a structured discussion about current affairs under the control of a moderator. As shown in Table 3, most of demonstratives are pronominal that, as expected in interactive discourse.

| Discourse topic | that | this | Total |
|-----------------|------|------|-------|
| Same Speaker    | 20   | 5    | 25    |
| Other Speaker   | 20   | 5    | 25    |
| Discourse Topic | 6    | 11   | 17    |
| Extralinguistic | 1    | 1    | 2     |
| Reminders       | 2    | 2    | 4     |
| Indefinite      |      | 3    | 3     |
| Time-Person-Place |    | 11   | 11    |
| Forward         | 5    | 5    | 10    |
| Other           | 2    | 2    | 4     |
| Total           | 48   | 42   | 90    |

Table 3. Demonstratives in televised discussions

Discourse topic this. The televised discussions are distinguished from the casual conversations by the frequent use of this to refer to non-speaker activated entities. Such uses contradict our claim that this must be speaker-activated. To account for such examples, we suggest that a distinction be made between inclusive and exclusive speaker space. In the case of local discourse segments, speaker space often excludes the addressee, but with higher-level discourse topics, which are often shared, speaker space includes the addressee. In such cases, which are characteristic of highly structured interactions, this may be used for something which was not initially activated by the speaker. In (20), reference is made to the explicit discourse topic of the segment, Gorbachev’s decision not to cancel the summit:

(20) K: I think it’s the Ninth, isn’t it.
5.3 Planned, non-interactive genres. Tables 4-6 show the distribution of demonstratives in newspaper stories (New York Times "Week in Review" section, 6-11/87; a University of Minnesota administrative planning document; and an electronic specification document supplied by Control Data Corporation.

A characteristic of these non-interactive, written genres is a relatively high percentage of demonstratives. This may be partly attributable to the fact that everything that is activated is speaker activated. In addition, there are many time expressions with this in the newspaper articles and metadiscourse references such as this document in the planning document. Unlike the discourse-topic use in the televised discussions, referents of determiner this in the written genres are typically activated in the immediately preceding sentence and constitute a focus shift.

| Table 4. Demonstratives in newspaper stories |
|---------------------------------------------|
| n-0 | n-1 | n-2+ | Time | Metadiscourse | Total |
|-----|-----|------|------|--------------|-------|
| 9   | 9   | --   | --   | --           | 18    |
| 13  | 2   | 9    | 6    | 11           | 39    |
| 20  | 11  | 4    | 11   | 8            | 44    |
| 39  | 11  | 8    | 44   | 8            | 83    |

| Table 5. Demonstratives in planning document |
|---------------------------------------------|
| n-0 | n-1 | n-2+ | Total |
|-----|-----|------|-------|
| 3   | 5   | 3    | 9     |
| 4   | 78  | 8    | 82    |
| 3   | 92  | 3    | 102   |

In Focus. Elements in focus are those which are most highly activated. These always include at least the topic of the sentence, any high-level topics (including those not overtly represented in the sentence), and under certain as yet poorly understood conditions, the referent of the linguistic focus. In the current centering paradigm, the topic of the sentence is equated to the backward looking center (cf. Joshi and Weinstein 1981).

6. IMPLICATIONS FOR ALGORITHMS

The previous sections proposed constraints on demonstratives and discussed their use in different genres. This section outlines components of a natural language system that would capture the relevant notions of shared familiarity, activation and focus and explores possibilities for incorporating these into current discourse-processing algorithms.\(^1\)

1. Shared familiarity. At minimum, a computational model of shared familiarity requires maintenance of a user discourse history in which a record is maintained of all entities referenced in conversations with a particular user. Thus, for any entity in its knowledge base, the natural language system knows whether that entity has been discussed before (shared familiarity) or not (familiarity unknown). Only in the former case can a definite demonstrative expression be used. The recognition of discourse units (e.g. speech acts of Allen 1983) and the relations between them, e.g. conversational moves (Rothman 1985), rhetorical predicates (McKeown 1985, Mann and Thompson 1986) are also important in demonstrative resolution. These require a sophisticated model which keeps a record of beliefs and intentions of discourse participants. While such a model could incorporate into existing discourse structure frameworks (e.g. Groz and Sider 1986), no specific proposals to account for shared familiarity have yet been advanced (but see Sparck Jones 1986).

Activation. An adequate model of activation must isolate that subset of shared entities which is activated at any given point in the discourse. This subset includes entities referenced in the immediately preceding sentence, entities present in the immediate spatio-temporal context of the discourse, and beliefs and intentions relevant to the current discourse segment. Many current discourse algorithms which function at the local level of discourse structure can be used to model activation due to the immediately preceding sentence. For example, Hajicˇov (1987) points out that elements in McKeown's (1985) potential focus list can be equated with activated elements. Similarly, in the current centering paradigm (Groz, Joshi and Weinstein 1986, Brennan, Friedman and Pollard 1987) elements in the set of forward looking centers can also be considered activated. At the global level, Sider and Groz (1986) describe a model of discourse structure which indicates currently activated beliefs and intentions at any given point in a discourse.

In Focus. Elements in focus are those which are most highly activated. These always include at least the topic of the sentence, any high-level topics (including those not overtly represented in the sentence), and under certain as yet poorly understood conditions, the referent of the linguistic focus. In the current centering paradigm, the topic of the sentence is equated to the backward looking center (cf. Joshi and Weinstein 1981). In this

1. Kryk (1987) presents an algorithm for demonstrative interpretation in Polish and English based on Sider's 1983 algorithms. Kryk's algorithm, however, does not appear to be explicit enough to serve as a basis for implementation. Moreover, some of the algorithm's predictions, for example that the demonstrative that is never used for non-co-present extralinguistic reference whereas non-co-present this is used only for extralinguistic reference, are inaccurate.

2. Although our notion of activation is similar to Hajicˇov (1987), it differs in that the consider the linguistic focus to be more highly activated than the topic. We feel that our assumptions about degrees of activation are more compatible with the fact, noted also by Hajicˇov, that topic maintains its activation longer than does the referents of the linguistic focus.
paradigm unstressed pronouns are viewed as preferring an existing backwards-looking center as antecedent. We have suggested that use of a demonstrative expression or stressed pronoun signals that the option to shift the backwards-looking center has been selected over the otherwise-preferred option of continuing an existing backwards-looking center. Brennan, Friedman and Pollard (1987), in fact, mention in passing that stressed pronouns in oral discourse could be used to unambiguously signal one such type of center shift. (cf. Sgall (1984), who remarks that demonstratives can be used to unambiguously specify reference to the linguistic focus of the immediately preceding sentence.)

Genre differences. A final application of our constraints is in the area of discourse genre variation. We have shown that both the form and function of demonstrative expressions varies between different spoken and written discourse genres. Our final suggestion is that these and other genre differences should be further explored, so that it will eventually be possible to design maximally-efficient discourse-processing algorithms which differentially exploit such genre distinctions.

7. REFERENCES

Allen, James. 1983. Recognizing intentions from natural language utterances. Computational models of discourse, ed. by M. Brady and R. C. Berwick, 267-330. Cambridge, MA: MIT Press.

Bosch, Peter. 1983. Agreement and anaphora. London: Academic Press.

Brennan, Susan E., Marilyn W. Friedman and Carl Pollard. 1987. A centering approach to pronouns. Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics, Stanford University, 155-162.

Buhler, Karl. 1982. The deictic field of language and deictic words. Speech, place, and action, ed. by R. J. Jarvella and W. Klein, 9-30. Chichester, NY: John Wiley & Sons.

Chafe, Wallace. 1976. Givens, contrastiveness, subjects, topics, and point of view. Subject and topic, ed. by C. N. Li, 25-55. New York: Academic Press.

Ehlich, Konrad. 1982. Anaphora and deixis: same, similar or different. Speech, place, and action, ed. by R. J. Jarvella and W. Klein, 315-338. Chichester, NY: John Wiley & Sons.

Fillmore, Charles J. 1975. Santa Cruz lectures on deixis. Indiana University Linguistics Club. Bloomington, IN: Indiana University Linguistics Club.

Fillmore, Charles J. 1982. Towards a descriptive framework for spatial deixis. Speech, place, and action, ed. by R. J. Jarvella and W. Klein, 31-59. Chichester, NY: John Wiley & Sons.

Grice, H. P. 1975. Logic and conversation. Speech Acts, Syntax and Semantics 3, ed. by P. Cole and J. Morgan, 41-58. New York: Academic Press.

Grosz, Barbara J. and Candace L. Sidner. 1986. Attention, intentions, and the structure of discourse. Computational Linguistics 12:175-204.

Hajicová, Eva and Jarka Vrbová. 1982. On the role of the hierarchy of activation in the process of natural language understanding. Proceedings of the Ninth International Conference on Computational Linguistics, Prague, 107-113.

Holliday, M.A.K. and Ruqaiya Hasan. 1976. Cohesion in English. London: Longman.

Hawkins, John. 1978. Definiteness and indefiniteness: a study in reference and grammaticality prediction. Atlantic Highlands, NJ: Humanities Press.

Hedberg, Nancy. Forthcoming. The discourse function of cleft sentences. University of Minnesota dissertation.

Hirschberg, Julia and Janet Pierrehumbert. 1986. The intonational structuring of discourse. Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics, Columbia University, 136-144.

Isard, Stephen. 1975. Changing the context. Formal semantics of natural language, ed. by E. L. Keenan, 287-296. Cambridge: Cambridge University Press.

Joshi, A.K., and Weinstein, S. 1981. Control of inference: role of some aspects of discourse structure - centering. Proceedings of the 1981 Meeting of the International Joint Conference on Artificial Intelligence, Vancouver, 385-87.

Keenan, E. Ochs and B. Schieffelin. 1976. Topic as a discourse notion: a study of topics in the conversations of children and adults. Subject and Topic, ed. by C. N. Li, 337-358. New York: Academic Press.

Kryk, Barbara. 1987. On deixis in English and Polish: the role of demonstrative pronouns. Frankfurt am Main: Verlag Peter Lang GmbH.

Lakoff, Robin. 1974. Remarks on this and that. Papers from the Tenth Regional Meeting of the Chicago Linguistics Society, ed. by M. LeCay, R. Fox, and A. Bruck, 345-356. Chicago, IL: Chicago Linguistics Society.

Linde, Charlotte. 1979. Focus of attention and the choice of pronouns in discourse. Discourse and syntax, ed. by T. Givon (Syntax and semantics, 12), 337-354. New York: Academic Press.

Lyons, John. 1977. Semantics, Volume 2. Cambridge: Cambridge University Press.

Mann, William C. and Sandra A. Thompson. 1986. Rhetorical structure theory: description and construction of text structures. Reprint ISI/KS-86-174. Information Sciences Institute, Marina del Rey, California.

McKeown, Kathleen, R. 1985. Text generation. New York: Cambridge University Press.

Prince, Ellen. 1981a. On the inferring of indefinite-this NPs. Elements of discourse understanding, ed. by A. Joshi, B. Webber, and J. Sag, 231-250. Cambridge: Cambridge University Press.

Prince, Ellen. 1981b. Towards a taxonomy of given and new information. Radical pragmatics, ed. by P. Cole, 223-256. New York: Academic Press.

Reichman-Adar, Rachel. 1984. Technical discourse: the present progressive tense, the deictic "that," and pronounization. Discourse Processes 7:337-369.

Sgall, Petr. 1984. A note on text and reference. Contributions to functional syntax, semantics and language comprehension, ed. by P. Sgall, 139-144. Amsterdam: John Benjamins.

Sidner, Candace, L. 1983. Focusing in the comprehension of definite anaphora. Computational models of discourse, ed. by M. Brady and R. C. Berwick, 267-330. Cambridge, MA: MIT Press.

Spark, John. 1986. Issues in user modelling for expert systems. Artificial intelligence and its applications. New York: John Wiley & Sons.

Wald, Benji. 1983. Referents and topic across discourse units: observations from current vernacular English. Discourse perspectives on syntax, ed. by F. Klein-Andreu, 91-116. New York: Academic Press.