Definite Descriptions and Cognitive Status in English:

Why Accommodation is Unnecessary

JEANETTE K. GUNDEL, NANCY HEDBERG AND RON ZACHARSKI

ABSTRACT

A commonly held view of English definite articles is that they signal that the referent of an NP is familiar to the addressee. However, it is well known that not all definite article phrases meet this familiarity requirement. To account for such non-familiar uses, Heim (1982) invokes the mechanism of 'accommodation', which enables an addressee to remedy a violation of the familiarity requirement by adding assumptions to the 'common ground'. In this paper we argue that the Givenness Hierarchy framework provides an insightful account of all uses of definite article phrases without requiring an appeal to accommodation. Such an account provides a unified treatment of definite article phrases, including demonstrative phrases and personal pronouns, while at the same time distinguishing among them in a principled way. This proposal is supported by results of a corpus-based examination of the use of definite articles and by an examination of cleft presuppositions.*

1. INTRODUCTION

It is generally recognized that definite article phrases like *the train* share with demonstrative phrases like *that train* or *this train* the property of being in some sense 'definite'; and this same property of definiteness is also shared by personal pronouns like *it* and demonstrative pronouns *that* and *this*. While there is as yet no generally accepted explicit definition of what exactly it means for an expression to be definite, there is a growing consensus that definiteness cannot be defined as a semantic property of the intended referent alone, for example in terms of specificity, referentiality, or uniqueness, as this would not distinguish formally definite NPs from their so-called indefinite counterparts. For example, *a train* in (1) and *the train* in (2) can both refer uniquely to a specific train.

(1) Look out, a train is coming

(2) Look out, the train is coming.

Rather, what characterizes definite NPs must be a property that is grounded in the interactive nature of language use and the speaker's assumptions about the status of the intended referent in the mind of the addressee.

But what exactly is this property that is associated with all definite NPs, and what distinguishes definite article phrases, as in (2), from other definite expressions? The present paper will be concerned with these questions.

In previous work (Gundel, Hedberg & Zacharski 1988, 1993, 2000) we have proposed that definite article phrases like *the train* require that the addressee be able to assign a unique representation to the phrase as soon as he is finished processing it, and that the individuating property must be distinct from what is being predicated of it in the sentence. This property is

* We wish to thank Jacques Durand, Francis Cornish and two anonymous reviewers for their helpful comments and suggestions.

shared with other definite nominals, e.g. demonstrative phrases and personal pronouns. What distinguishes definite article phrases from other definites, we maintain, is that identification doesn't have to be based on previous familiarity with the referent. The addressee is not expected to already have the appropriate representation in memory, though he may. This proposal contrasts with what is still a prevalent view concerning the meaning of the definite article, namely that referents of definite article phrases must be in some sense already familiar to the addressee. This view, which Lyons (1999) refers to as the familiarity hypothesis, goes back at least as far as Christophersen (1939), and has found more recent formal expression in the highly influential file change semantics framework of Heim (1982) and subsequent work by other authors, according to which a definite article is an instruction to the addressee to select an already existing card from the current file.

It is well known, however, that the familiarity hypothesis can't account for all uses of the definite article in English (see Birner & Ward 1994, Fraurud 1990, 1992, Gundel, Hedberg & Zacharski 1990, Hawkins 1978, 1991, Poesio & Vieira 1998 inter alia)[1]. For example, the familiarity hypothesis cannot account straightforwardly for the boldfaced definite article phrases in (3)-(6).

(3) I want to determine **the maximum number of boxcars I can get to Bath** [Heeman & Allen 1995]

(4) Metallica and others in the record industry may complain that unauthorized exchange of music files on the Internet is killing their business, but **the numbers** have yet to bear that out. [Neil Strauss, Teeney-Boppers Shatter a Record. New York Times 5.24.2000]

(5) **The chief state prosecutor of Maryland** decided today to drop criminal wiretap charges against Linda R. Tripp, the Pentagon employee whose secretly recorded conversations with Monica S. Lewinsky exposed an affair that led to President Clinton's impeachment [Don van Natta, Maryland is dropping wiretap case against Tripp. New York Times 5.24.2000]

(6) Laws on taping telephone calls vary from state to state. In Maryland, where Mrs. Tripp lives, it is against the law for one party to tape a phone conversation without **the consent of the other**. [Don van Natta, Maryland is dropping wiretap case against Tripp. New York Times 5.24.2000]

While there is no problem interpreting the phrases *the maximum number of boxcars I can get to Bath* in (3), *the numbers* in (4), *the chief state prosecutor of Maryland* in (5), and *the consent of the other* in (6), there is no reason to expect that anyone reading or hearing these phrases in the given context was already familiar with the intended referent beforehand. This is also evidenced by the fact that in some cases the definite article could be omitted or replaced with an indefinite article without any change in meaning or acceptability. The definite article could be omitted in (6) (*without consent of the other*), for example.

To account for such non-familiar uses of the definite article, Heim (1982) invokes the mechanism of "accommodation" (Lewis 1979). Given this mechanism, which is available to all

[1] Fraurud's work is actually based on a corpus study of Swedish, though she extends her observations to English. See also de Mulder 2000 and references cited therein for French.

cases of presupposition, the familiarity condition acquires the status of actually being fulfilled by virtue of being treated as if it were fulfilled.

In this paper we argue that the Givenness Hierarchy framework proposed in Gundel et al. 1993 can adequately account for all uses of definite article phrases, including ones like those in (3)-(6) without requiring an appeal to accommodation. Such an account allows a unified treatment of definite noun phrases, including demonstrative phrases and personal pronouns, while at the same time allowing us to distinguish between them in a principled way.

We begin with a brief summary of the Givenness Hierarchy framework and its predictions. We then provide results of a corpus investigation that further supports previous findings (e.g. Gundel et al. 1993, and Poesio & Vieira 1998), that close to half of all definite article phrases in naturally occurring English texts, both spoken and written, have non-familiar referents. Finally, we outline a proposal from Hedberg (2000) which extends our approach to cleft sentences, another so-called presuppositional construction that has been analyzed as involving accommodation.

2. THE GIVENNESS HIERARCHY — DEFINITENESS, FAMILIARITY AND UNIQUE IDENTIFIABILITY

2.1 THE GIVENNESS HIERARCHY

Gundel Hedberg and Zacharski (1993) (henceforth GHZ) propose that determiners and pronouns restrict their possible referents by conventionally signaling different cognitive statuses (memory and attention states) assumed of the mind of the addressee with regard to the intended referent.

The full range of cognitive statuses constitutes the Givenness Hierarchy in (7) (along with the relevant English forms), where each status entails all statuses to the right.

(7) **The Givenness Hierarchy**

in			uniquely		type
focus >	activated	> familiar	> identifiable	> referential	> identifiable
{*it*}	{*that* / *this* / *this* N}	{*that* N}	{*the* N}	{indefinite *this* N}	{*a* N}

Each status is overtly signaled by a different form (or forms) that has that status as part of its conventional meaning. By conveying information about the memory and attention states of the referent, the different forms serve as processing signals that assist the addressee in restricting possible referents. The lowest status (type identifiable) is least restrictive and the highest status (in-focus) is most restrictive.

The theory predicts that a pronoun or determiner will be used (and used appropriately) only if the status that it conventionally codes within the language is satisfied in the given context in which it is used. In addition to this absolute prediction, the theory also predicts that the same status can be coded by different forms. This follows from the fact that the statuses are in a unidirectional entailment relation. If something is in focus, then it is by definition also activated. If it is activated, then it is necessarily familiar (already represented in memory), and so on. A given status could therefore be appropriately coded by a form which explicitly signals that status, because it has that status as its meaning, and it can also be coded by forms whose meanings are

entailed by that status. In (8), for example, the phrase *these men* explicitly signals that the referent is activated, since this is the meaning of the proximal demonstrative determiner *this/these* in English.

(8) As far as the likelihood of Indian women dating American men, my observation is that Indian girls born in America or raised in America from a very young age do not care to date or marry Indian men, for the same reasons mentioned above. They see the double standards Indian men hold on to, even those men that are born and raised here. In a vast majority of cases, **these men** inherit hangups from their Indian parents you see.

The form *these men* is appropriate here since the referent was just mentioned in the immediately preceding sentence and therefore could be expected to be activated (in working memory) for the addressee. In fact, Indian men can be expected to be not only activated, but in-focus, as this referent was mentioned three times in the previous two sentences and can thus be assumed to be in the addressee's focus of attention. So an unstressed pronoun *they* could have been used as well. And since anything in focus is not only activated, but also familiar, uniquely identifiable, referential and type identifiable, other forms would have been equally appropriate here, including *those men*, which requires only an existing representation somewhere in memory (familiar), *the men*, which requires only the ability to associate a unique representation on the basis of what is encoded in the nominal expression (uniquely identifiable), and even *Indian men* or *Indian men that are born and raised here*, which requires only the ability to understand what the words mean, i.e. to identify the type (type identifiable).

The distribution of forms across statuses that meet necessary conditions for their use is not random, however. For example, the indefinite article is rarely used to code statuses higher than referential; and most in-focus referents are not coded with demonstrative determiners or pronouns, even though they could be. Within the GHZ approach, such facts are not the direct consequence of conventional meanings of the forms in question. Rather, they follow from interaction of the Givenness Hierarchy (and the form-status correlations in the language) with general pragmatic principles that govern language processing, for example the assumption that the speaker/writer is being "as informative as required" (see Grice 1975, Gundel & Mulkern 1998, Sperber & Wilson 1986/95).

Consider, for example, the observation made in Cornish (1999) that demonstrative determiners, unlike the definite article, often serve to make the referent more salient for the addressee, as in (9) (his 2.19a, see also Cornish this volume)

(9) [It is dusk and John and Mary are returning from a shopping trip. As John is parking the car, Mary exclaims:] Good God! Look at that incredibly bright light.

Cornish points out that a definite article could not effectively call the addressee's attention to the light in (9).

Within the GHZ framework, the function of making a referent more salient by calling the addressee's attention to it is not part of the conventional meaning associated with the demonstrative determiner (or with any other form for that matter). The demonstrative does not necessarily have that function in (8) for example, where a definite article would have been equally appropriate. Rather, it follows from interaction of the meaning of demonstratives (i.e. the fact that they explicitly signal the statuses 'activated' and 'familiar') with the Maxim of Quantity: make your contribution as informative as required (Q1); don't make your contribution more informative than necessary (Q2). GHZ argue that for full definite NPs, signaling identifiability with the definite article is usually sufficient information about cognitive status to allow the addressee to assign a unique representation to the referent, given the descriptive content of the

noun and possible modifiers. Using a demonstrative, which is more informative about cognitive status since it restricts possible referents to those that are at least familiar, is usually unnecessary unless the speaker or writer wants to achieve some special effect, as in (9). Since the distal demonstrative explicitly signals familiarity, it invites the hearer to search for a familiar referent. Moreover, its use in (9) implicates by the first part of the Quantity Maxim that the referent does not have a higher status, i.e. is not currently activated or in-focus. By so doing, it signals a shift in focus and raises the salience of the referent by calling the addressee's attention to it.

According to the GHZ account, then, the definite article is underspecified for statuses higher than uniquely identifiable and the demonstrative determiners are underspecified for statuses higher than familiar (in the case of the distal demonstrative determiner) or activated (in the case of the proximal demonstratives). This predicts correctly that any of these forms could be used to code referents that are activated or even in-focus, but none of them would have to be; and interaction of these facts with the two parts of the Quantity Maxim explains why the definite article is used to code in-focus referents more frequently than demonstratives are. But like other pragmatically motivated facts, these are tendencies that arise in appropriate contexts, rather than invariant meanings associated with the forms in question.

The one to many pairing of referring forms with cognitive statuses distinguishes the GHZ approach from other referential status hierarchies and taxonomies (e.g. Ariel 1990, Chafe 1994, Lambrecht 1994, Prince 1981, 1992) in which different statuses are wholly or partly mutually exclusive. It thus allows the theory to make a complex and wide-ranging set of predictions, both absolute and probabilistic, about the distribution and interpretation of referring forms in discourse. In addition, it makes possible an explicit definition of the conceptual notion of 'definiteness'. Forms that have been pretheoretically classified as 'definite' are all associated with the status 'uniquely identifiable', either directly, as in the case of the definite article, or by implication, as with the forms that overtly signal higher statuses that entail 'uniquely identifiable', such as personal pronouns or demonstratives. But what exactly does it mean to be uniquely identifiable?

2.2 Definites and Unique Identifiability

As we are using the term, an intended referent is uniquely identifiable if the addressee can assign a unique representation to it, where the individuating property is distinct from what is predicated of the referent in the rest of the sentence. This should not be confused with the Russellian notion of uniqueness as a property of the referent or the concept of identifiability as the ability to recognize a particular referent independently of the description itself. As we use the terms here, uniqueness and identifiability are not properties of the actual referent in the world or in the universe of discourse. Rather, they are properties of the addressee's **representation** of the referent. A representation is unique just in case it contains sufficient information to distinguish it from other representations in memory. This sense of uniquely identifiable is thus also neutral between the referential and attributive uses discussed in Donnellan (1966).

An individual can assign a unique representation by retrieving an existing representation from (long-term or working) memory, since anything that is familiar is also uniquely identifiable, given an appropriate description. In (10), for example, intended referents of the phrases *the prosecutor* in (b), *the decision* in (c), and *the case* in (d) would all be uniquely identifiable because they are familiar. In each of these cases, the addressee can be expected to assign a unique representation because such a representation already exists in his memory, as a result of processing the previous text.

(10) (a) The chief state prosecutor of Maryland decided today to drop criminal wiretap charges against Linda R. Tripp, the Pentagon employee whose secretly recorded conversations with Monica S. Lewinsky exposed an affair that led to President Clinton's impeachment.

(b) **The prosecutor**, Stephen F. Montanarelli, said it would be impossible to successfully pursue a trial, which was to begin on July 10, because a judge this month suppressed testimony by Ms. Lewinsky that, Mr. Montanarelli said, was crucial to obtaining a conviction.

(c) **The decision** concludes the only criminal case against a major figure in **the most serious Oval Office scandal since Watergate**. Mrs. Tripp had faced **the possibility of a 10-year prison sentence and a $10,000 fine**.

(d) "I had no choice" but to drop **the case**, Mr. Montanarelli said in an interview today, [Don van Natta, Maryland is dropping wiretap case against Tripp. New York Times 5.24.2000]

A referent may also be uniquely identifiable because it is familiar from general experience, cultural knowledge, or presence in the spatiotemporal context even if it has not been previously mentioned in the text (what Prince 1992 calls 'hearer old, but not discourse-old'). This would be the case with the phrase *the orange juice factory* in (11), for example, where the factory has not been mentioned previously in the conversation, but both speech participants know about it because it is on the map, which they have in front of them.

(11) s: hello can I help you
u: yeah I want t- I want to determine the maximum number of boxcars of oranges that I can get to Bath by seven a.m. tomorrow morning so hm so I guess all **the boxcars** will have to go through oran- through Corning because that's where **the orange juice orange factory** is [Heeman & Allen 1995]

Thus, in order for familiarity to serve as the basis for identifiability, it is irrelevant how an entity came to be familiar, whether through textual introduction, cultural or personal experience, presence in the immediate extralinguistic context, and so on. In fact, we know of no language that codes such distinctions. The relevant factor is the cognitive status that a particular entity has for the addressee, not how it acquired that status.

In (10), the referents of the phrases *the chief state prosecutor of Maryland, the Pentagon employee whose secretly recorded conversations with Monica S. Lewinsky exposed an affair that led to President Clinton's impeachment*, and *the only criminal case against a major figure in the most serious Oval Office scandal since Watergate* may be already familiar to the addressee from extralinguistic experience as well, but it would not be **necessary** for these referents to be familiar in order for them to be uniquely identifiable. The subject of (10a), for example, would be uniquely identifiable even if it had been *the vice-magistrate of Transylvania*. And the reader surely cannot be expected to have an already existing representation of *the possibility of a 10-year prison sentence and a $10,000 fine* in (10c). But all these examples are uniquely identifiable because there is enough descriptive content encoded in the phrase itself to allow the reader to construct a new unique representation.

Thus, while previous familiarity is a sufficient condition for unique identifiability as we have defined this term, it is not a necessary one. The status 'uniquely identifiable' can be met either by already having an existing representation in memory (i.e. through familiarity) or by constructing a new unique representation. The latter is possible if the addressee is able to construct a unique representation based on information encoded in the phrase alone, as in the examples from (10) discussed above. It is also possible if the addressee can easily access what

Clark and Haviland (1977) call a 'bridging inference' to a recently activated entity, as in so-called inferrable or associative anaphors like those in (12)- (14).

(12) Metallica and others in the record industry may complain that unauthorized exchange of music files on the Internet is killing their business, but **the numbers** have yet to bear that out. [Neil Strauss, Teeney-Boppers Shatter a Record. New York Times 5.24.2000]

(13) I've just been to a wedding. **The bride** wore blue. [Lyons 1999, p7]

(14) [Nurse entering operating theatre]
I wonder who **the anesthetist** is today. [Lyons 1999, p7]

The boldfaced phrases in (12)-(14) do not refer to a particular entity that the addressee can already be expected to have represented in memory; so they are not familiar in our sense of the term. However, they are all uniquely identifiable since the addressee is able to construct a unique representation by way of a bridging inference to a recently activated entity. This bridging inference is often made possible by some familiar script or frame, e.g. in (13) the knowledge that weddings normally include a bride or in (14) the knowledge that an anesthetist is normally present during an operation. But the particular referents of the inferrable phrases would not already be in memory (at least not necessarily so) at the point when the phrase that refers to them is encountered.[2]

Finally, as noted above, identifiability does not require ability to recognize the intended referent independently of the description. Thus, the observation in Lyons (1999) that the definite phrase in an example like (14) is not identifiable because the addressee can't be expected to recognize the individual referred to if he passes her on the street the next day cannot be taken as evidence that the referent is not uniquely identifiable as we define it. Nor does the description itself have to pick out a unique referent. For example, the phrase *the door* in (15) would be uniquely identifiable in the sense that the hearer can easily construct a unique representation of an open door in the present situation, even though the descriptive content in the phrase does not itself pick out a unique door since it doesn't exclude other doors in the immediate context.

(15) [In a room with three doors, one of which is open]
Close **the door**, please. [Lyons 1999, p16]

Following Heim (1982), non-familiar definites have sometimes been accounted for by invoking a process of accommodation, whereby a new representation (or 'card' to use Heim's metaphor) is created by way of linking to a familiar entity. Although the term 'accommodation' suggests a kind of repair, Heim herself builds accommodation into her theory of definites, presumably a model of competence, and thus equivocates on the question of whether non-familiar definites constitute a kind of intentional performance error.[3] In any case, those who subscribe to

[2] Note that the Givenness Hierarchy framework, or any other theory of definites that we know of, does not itself address the question of how the language processor determines the means for assigning a unique representation, e.g. whether by constructing a bridging inference or by searching for an existing representation in long term or working memory. Such questions, we maintain, are outside the domain of a theory of definites or other referring forms, and can only be adequately addressed within a general pragmatic theory of language understanding. For specific proposals that address these questions within the framework of Relevance Theory (Sperber and Wilson 1986/95), see Gundel (1996), Gundel and Mulkern (1998), and Matsui (2000)

[3] But see Cornish (1996) and Gundel, Hedberg and Zacharski (2000) for a more clearly performance-based notion of accommodation, invoked to account for 'antecedentless' pronouns whose referents cannot be assumed to be in focus for the addressee.

this account take familiarity to be the primary meaning of the definite article, with non-familiar definites being exceptions to the norm whose interpretation requires some special mechanism. By contrast, the Givenness Hierarchy framework provides a unified account of both familiar and non-familiar uses of the definite article, which captures similarities and differences between the two uses without treating one or the other as more 'marked'.

An account that considers the essential meaning of the definite article to be unique identifiability rather than familiarity is also supported by investigations of the relative frequency of non-familiar definite article phrases in naturally-occurring discourse, as we shall see in the next section.

3. FREQUENCY OF NON-FAMILIAR DEFINITES

3.1 PREVIOUS STUDIES

In a study that investigated transcripts of casual conversations, narrated film descriptions, and magazine articles (Gundel et al. 1993) we found that 108 out of 280 definite article phrases (39%) were non-familiar in the sense that the addressee could not be expected to have an existing representation in long term or working memory. Poesio and Vieira (1998) surveyed 1,040 definite article phrases taken from the Wall Street Journal corpus contained in the ACL/DCI CD-ROM, and concluded, based on their own annotations of the data, that 48% were not related to an antecedent previously introduced in the text. This count included some phrases that we would consider familiar, for example redescriptions of previously mentioned referents and NPs whose referents were new to the discourse but familiar from previous experience. But it also included some phrases that we would consider non-familiar. Specifically, Poesio and Vieira's "associative" class (18.55%) included inferrables, which we could classify as non-familiar, because the addressee would not be expected to have a preexisting representation of the particular referent, even though the association might be based on a familiar frame or script. Although the two studies classify the data somewhat differently, both found that close to half the uses of definite article phrases were non-familiar in some sense.[4]

In our current study we sought to broaden the empirical base of our 1993 work by exploring the extent to which definite article phrases refer to entities that are non-familiar in both written (planned) and conversational (unplanned) English discourse.

3.2 METHODOLOGY

We analyzed 321 definite article phrases taken from three genres: 138 examples from three Canadian Parliament transcripts included in the Hansard Corpus available on the World Wide Web, 105 examples from three New York Times articles, and 78 examples from four spoken dialogues taken from the TRAINS corpus (Heeman & Allen 1995). All definite article phrases in these texts were analyzed, with the exception of proper names, appositives, or phrases included in direct quotes.

Three coders (each of the authors) coded each transcript separately. We coded each definite article phrase for highest cognitive status (type identifiable, referential, uniquely

[4] Corpus studies of other languages yield similar findings. For example, Fraurud (1990) analyzed 745 definite article phrases in 11 Swedish nonfiction written texts and found that 61% of them were first mention, i.e. new to the discourse. Gundel et al (1993) analyzed 132 definite article phrases in a variety of spoken and written texts in Spanish and found that 42 of these (32%) had referents that were uniquely identifiable, but not familiar to the addressee (either from previous discourse or from general experience).

identifiable, familiar, activated or in-focus).[5] Most of the divergences in coding came from difficulties in distinguishing in-focus referents from those that are activated, but not in-focus, and activated referents from those that are familiar, but not activated. Since these problems aren't directly relevant to the present study, as both activated and in-focus referents are also familiar, we only calculated agreement statistics with respect to whether the noun phrase was originally coded as familiar (in-focus, activated, or familiar) or non-familiar (at most uniquely identifiable). The results are shown in Table 1.

	Total number of definite article phrases	*Number of phrases coded identically by the coders*	*% Agreement*
Parliament	138	106	77%
Newspapers	105	86	82%
TRAINS	78	69	88%
Total	321	261	81%

Table 1. Percentage of original coder agreement as to whether NP is familiar or non-familiar.

3.3 SUMMARY OF RESULTS

We found that 140 out of 321 definite article NPs (roughly 44%) were non-familiar in the sense that the hearer/reader could not be expected to have an existing representation in memory (hearer-old in Prince's 1992 terms). The distribution across genres is shown in Table 2:

	Total number of definite article phrases	*Number of non-familiar definite article phrases*	*% non-familiar*
Parliament	138	85	61.6%
Newspapers	105	44	38.1%
TRAINS	76	20	19.7%
Total	319	140	43.9%

Table 2. Percentage of non-familiar definite article phrases.

It is clear from the figures in Table 2 that there is a large difference between genres in how often a definite noun phrase referent is familiar or not. The genre with the lowest percentage of non-familiar definites (19.7%) is the TRAINS corpus. Here, both speaker and addressee were focussed on a map which they had in front of them and various entities which could be placed on the map at various times in the problem space. Since the primary purpose of the dialogue was to

[5] For the Newspaper articles and one of the Parliament transcripts, we discussed each case where the original coding diverged and came to an agreement on how to code the example for the final count. For the other two Parliament transcripts and the TRAINS transcripts, we took the majority viewpoint on the original coding as final.

discuss the entities in various places on the map, it is not surprising that most of the entities discussed were familiar.

Newspaper articles were intermediate in the percentage of definite NPs with non-familiar referents (38.1%). Here, a certain number of referents were introduced and then were discussed in narrative style. It is reasonable to assume that in narratives, a relatively large number of the introduced referents are talked about again later in the text.

The Parliament transcripts were highest in the percentage of definite NPs with non-familiar referents (61.6%). Parliament debates are an example of expository rather than narrative texts. The speakers produce arguments, which often entail the introduction of new referents rather than the re-evocation of old referents.

The overall average of novel referents in the three genres in the current study (43.9%) compares quite closely to the average reported in GHZ 1993 (39%), which also included a diverse set of genres, ranging from spontaneous conversation to formal, written prose. We can thus be confident that the proportion of non-familiar definites in overall discourse is quite high.

3.4 EXAMPLES OF DIFFERENT CODINGS

In this section we give examples of each of the possible codings. These examples are taken from the set of examples where the 3 coders had unanimous agreement as to status.

3.4.1 NON-FAMILIAR

UNIQUELY IDENTIFIABLE BUT NOT FAMILIAR

As mentioned above, roughly 44% of the definite article NPs we examined were at most uniquely identifiable, i.e. uniquely identifiable but not familiar. At most uniquely identifiable entities, all of which require the addressee to construct a new unique representation, fall into two broad classes – examples where a unique representation can be constructed by way of a 'bridging inference' to a recently introduced entity and examples where a new representation can be constructed solely on the basis of descriptive content encoded in the phrase itself. The first type is illustrated in (16)-(18):

(16) According to Soundscan, a media-sales monitoring company in Hartsdale, N.Y., "Oops!" sold 1.3 million copies in its first week to reach **the No. 1 position**. [Neil Strauss, Teeney-Boppers Shatter a Record. New York Times 5.24.2000]

(17) [article on the hijacking of a plane. No previous mention of the cockpit]
The hijacker, wearing a blue ski mask and carrying a handgun and a grenade, ordered the pilot of Flight PR812 to return to Davao City, 600 miles from Manila, but the pilot said he did not have enough fuel, Manila airport general manager Antonio Gana said. At one point, the hijacker fired a shot inside **the cockpit**, perhaps accidentally, one of the three pilots aboard said. [Associated Press. Hijacker Leaps to Safety After Robbing Passengers. 5.25.2000]

(18) The plane was depressurized so that the hijacker could jump with a parachute while it circled 13 miles from Manila at 6,000 feet, PAL spokesman Rolando Estabillo said. A strong gust of wind swept into the plane when **the door** was opened, a passenger said. [Associated Press. Hijacker Leaps to Safety After Robbing Passengers. 5.25.2000]

In (16), the speaker can use the expression *the No. 1 position* without assuming that the addressee already has some mental representation of the referent, since the addressee can easily construct a unique representation of the intended referent via a `bridging inference` that is accessible from the immediate context. The same thing is true in inferrable/associative uses such as *the cockpit* in (17) and *the door* in (18). As noted earlier, such uses of the definite article phrases are not considered to be familiar because the speaker's intended referent is not already represented in the addressee's memory. Even though the knowledge required for the appropriate bridging inference can be assumed to be familiar, and even highly accessible to the addressee, there is no reason to expect that he already has a mental representation of the intended referent at the point immediately before the noun phrase is encountered.

For cases where the addressee is able to construct a unique representation based solely on the descriptive content encoded in the noun phrase itself, illustrated in (19) and (20), the only kind of familiarity that needs to be assumed is familiarity with the type of thing described by the expression – a condition that is necessary for appropriate use of ALL noun phrases, including indefinites. (*The exact opposite* in (19) is an example of the type in (17) and (18), where identification is possible via a bridging inference to something that has been recently activated.)

(19) To use **the terminology of one radio consultant**, pop music currently is in the doldrums, where blandness is king. But coming up next, according to this theory, is **the exact opposite**, when alienated music fans embrace something loud and ugly, which this time around may be such an extreme reaction that fans will be pining for **the safe, predictable days when Britney Spears couldn't get no satisfaction**. [Neil Strauss, Teeney-Boppers Shatter a Record. New York Times 5.24.2000]

(20) Montreal has something to offer **the religious tourist** and **the night life tourist**, **the art connoisseur**, **the sports fan**, **the intrepid walker** and **the avid consumer**. [Debates of the House of Commons of Canada 5.7.97]

In (19), *the safe, predictable days when Britney Spears couldn't get no satisfaction* is an example of the non-familiar use which Hawkins' (1978) calls the "establishing relative clause" use of definite descriptions.[6] Prince (1981) refers to such phrases as "containing inferrables." They all share the property of being uniquely identifiable, based on the descriptive content encoded in the phrase, without being already familiar. The examples in (20) represent a kind of generic use of the definite article. Since generics refer to a class of objects or individuals that the addressee can be expected to identify, in the sense that he can assign a unique representation that distinguishes the class, they would inherently satisfy the unique identifiability condition. This is why many languages (French for example) consistently mark generics with the definite article. Generics are often familiar as well, but this is not always the case. The generics in (20) were classified as at most uniquely identifiable (i.e. non-familiar) because the addressee could not be expected to have an existing representation of the class of intrepid walkers or religious tourists in memory.

Definite article phrases whose intended referents are at most uniquely identifiable can also include superlatives, as in (21) and (22). For some superlatives, all the information necessary to construct a unique representation is encoded in the noun phrase itself. But since the examples

[6] Additionally in (19), *the terminology of one radio consultant* is an example of Hawkins' (1978) non-familiar "associative clause" use. Here the phrase contains both "the trigger and the associate of a pragmatically permissible associative anaphoric sequence" (Hawkins 1978: 139).

in (21) and (22) are elliptical, their identification requires a bridging inference to the immediate discourse context, as with the examples in (16)-(18).[7]

(21) A: so that'll be two hours to Corning uh to pick up the oranges one hour to load plus two hours to get back um and then one hour to make the orange juice so then we would have so that would be after six hours we'd have the tanker of orange juice ready at Elmira

B: is there a quicker way

A: um no that looks about **the fastest** [Heeman & Allen 1995]

(22) B: there's no other quicker way right

A: uh, I don't think so let's see what –so the one that takes **the longest** is to make the orange juice and there doesn't seem to be a quicker way to do that because we have to go get the oranges and back again. Right, no, that looks like the fastest. [Heeman & Allen 1995]

3.4.2 FAMILIAR

Roughly 56% of the definite article noun phrases in our corpus had familiar referents. Since the category 'familiar' includes referents that are activated and in-focus as well as those that are at most familiar, this category can be broken down into the following three types.

FAMILIAR BUT NOT ACTIVATED

This type includes referents that the speaker can be expected to have represented in memory, but which are not in working memory because they have not been recently mentioned and are also not present in the immediate spatiotemporal context. In some cases the referent can be assumed to be familiar from personal experience or general cultural knowledge, even though it has not been previously mentioned in the discourse, as in (23)-(25). Hawkins (1978) calls such uses "larger situation uses."

(23) If one takes a step back and looks at the rest of this week's music-group news, the situation looks bad for ugly, unpredictable rock 'n' roll: one of the most popular American rock bands of **the 90's**, the Smashing Pumpkins, announced that it would break up after its current tour, while Noel Gallagher of the biggest British rock band of the 90's, Oasis, walked out on the group mid-tour. [Neil Strauss, Teeney-Boppers Shatter a Record. New York Times 5.24.2000]

(24) Nevertheless, I would like to explain the advantages of this bill, if not to convince members opposite, then at least to let **the public** know about the sometimes mysterious ways in which the party in power operates. [Debates of the House of Commons of Canada 5.7.97]

[7] Many superlatives are examples of the non-familiar use that Hawkins (1978) calls 'unexplanatory' modifiers.

(25) We have only to remember the 1987 season, which was very good for us because it followed on **the terrorist attack on the *Achille Lauro* in the Mediterranean**. [Debates of the House of Commons of Canada 5.7.97]

In other cases, the referent is familiar due to previous mention in the discourse ('discourse old' in the sense of Prince 1991). However, the mention was not recent enough for the speaker to assume that the referent is still in the addressee's working memory. Examples of this type are given in (26) and (27). Hawkins (1978) classifies such examples as "anaphoric uses" of definite descriptions; he doesn't make a distinction between 'working' and 'long-term' memory.

(26) Central to the case was a Lewinsky-Tripp conversation that Mrs. Tripp taped on Dec. 22, 1997. This was the last talk between the two women that Mrs. Tripp recorded, and it occurred, prosecutors said, just after her lawyer had informed her that secret tape-recording was illegal in Maryland.
[4 intervening sentences]
"There are no other witnesses to the conversation whom the state can call to testify," the prosecutor said, "and Tripp cannot be compelled to testify."
In their conversation of Dec. 22, **the two women** discussed what they should say about Ms. Lewinsky's sexual relationship with the president … [Don van Natta, Maryland is dropping wiretap case against Tripp. New York Times 5.24.2000]

(27) okay and we'll go back with the oranges in the tanker car, make the orange juice and we can load it in the tanker car
[30 intervening utterances with no mention of the orange juice]
so the one that takes the longest is to make **the orange juice** [Heeman & Allen 1995]

ACTIVATED BUT NOT IN-FOCUS

Many familiar examples can be assumed to be activated (i.e. in working memory) because they have been recently mentioned or are present in the immediate spatiotemporal context. Such uses also correspond to Hawkins' (1978) "anaphoric uses" or "immediate situation uses." Some activated entities are at most activated, since they cannot be assumed to be in the current focus of attention, as in (28) –(30), for example.

(28) The Canada-New England run, with its 420,415 cruise days, accounts for only 1.2 percent of the total cruise market, which will reach 50 million cruise days in two years. The route that takes in the St. Lawrence occupies only a very small part of **the market**, and ranks twelfth among cruise routes. [Debates of the House of Commons of Canada 5.7.97]

(29) [discussion about boxcars]
u: and there's two available at Mira
s: that's right
u: there's already two at Bath
s: yes there are two at Bath
u: okay how long would it take from Dansville to Bath going the Corning Bath way
s: uh, okay um let's see we'll have to get an engine to Dansville um to pick up **the boxcars**. [Heeman & Allen 1995]

(30) The route that takes in the St. Lawrence occupies only a very small part of the market, and ranks twelfth among cruise routes. Even this position is threatened by the sustained and organized efforts made by southeastern Asia, Australia, New Zealand, and the Far East. A look at the evolution of traffic on the St. Lawrence since 1980 reveals regular growth, with two particularly good years. These statistical anomalies are directly related to the fact that **the route** is considered particularly safe [Debates of the House of Commons of Canada 5.7.97]

IN-FOCUS

A relatively small percentage (10%) of definite article NPs in our data have referents that can be assumed to be in the addressee's focus of attention. Examples are given in (31)-(32).

(31) I am to take one tanker of orange juice and then deliver **the one tanker of orange juice** plus four boxcars of oranges to Bath as soon as possible. [Heeman & Allen 1995]

(32) A quick survey of the decision makers in the cruise lines indicates that there are two drawbacks: the climate, and the fact that casinos cannot be open. These are what might be called the two irritants to the development of the St. Lawrence route. It is hard to do anything about **the climate**, but I hope that if the other irritant were removed by this bill, the deck would be stacked in our favour. [Debates of the House of Commons of Canada 5.7.97]

In all these examples, the definite article NP is perfectly appropriate, since anything that is in focus, or at most activated or familiar, is also uniquely identifiable. In some cases, there is no special need to use a full definite article NP rather than a form that more narrowly restricts the cognitive status of the referent. Thus, in (31) a pronoun (*it*, *this* or *that*) or a demonstrative determiner (e.g. *this tanker of orange juice* or *that tanker*) would have been perfectly felicitous in this context.[8] In other examples, such as (32), however, a full noun phrase is needed in order to distinguish the intended referent from other entities that can be assumed to be in the addressee's center of attention.

To sum up so far, we have proposed that the definite article in English only requires the addressee to be able to assign a unique representation, a criterion which is always met by familiar referents, but may be met by some non-familiar ones as well. Our theory therefore does not require any special mechanism of accommodation to account for the relatively large percentage of non-familiar definites that occur in both planned and unplanned discourse.

3.5 CODING DISAGREEMENTS AND POSSIBLE REASONS FOR THESE DISAGREEMENTS

As mentioned in the methodology section above (§3.2) there was some disagreement among the coders in determining the cognitive status of definite article phrases. In the course of discussing how to resolve coding differences, we uncovered various problems with coding and possible reasons for disagreement. Sometimes, discussion revealed that the diverging coder had simply made a mistake and would readily agree with the other two coders. More often, however, there was a principled reason for the disagreement.

[8] The pronoun *it* might have been a bit awkward here followed by *plus.*

One principled source of disagreement came up in coding the second instance of *the cockpit* in example (33):

(33) "At one point, the hijacker fired a shot inside **the cockpit**, perhaps accidentally," one of the three pilots aboard said.... [14 sentences later] "Those aboard the plane did not get a good look at the hijacker because when he stood up, he told everyone to hide their faces in their laps and not look at him, then he walked to **the cockpit**," passengers said in radio reports. [Associated Press. Hijacker Leaps to Safety after Robbing Passengers. 5.25.2000]

Two of us coded this NP as familiar because the referent had been mentioned before and one of us coded it as uniquely identifiable because she had forgotten that it had been mentioned before. We coded this example as familiar, since that was the original majority opinion. In such a case, as in all other examples of definite article usage for familiar referents, the phrase would simply be underspecified for familiarity, probably because such marking would be unnecessary (see Gundel et al 1993, p. 300). However, upon reflection, we suggest that sometimes speakers assume that referents are merely uniquely identifiable, even if they might be familiar. The reason is that constructing a new representation in such cases may require less processing effort than searching for an existing one in memory. When there is nothing to be gained by recognizing that an intended referent is already in memory, the speaker may as well treat it as new.[9] This would also explain why explicitly signaling familiarity by using the expression *that cockpit* would have been infelicitous here.

The related example in (34) was pointed out by Poesio and Vieira:

(34) For the Parks and millions of other young Koreans, the long-cherished dream of home ownership has become a cruel illusion. For **the government**, it has become a highly volatile political issue. [Poesio & Vieira 1998, p198]

Poesio and Vieira write: "This definite description could be classified as larger situation [Hawkins' 1978 terminology] because it refers to the government of Korea, and presumably the fact that Korea has a government is shared knowledge; but it could also be classified as being associative on the predicate *Koreans*."

In our terms, the larger situation use would be classified as familiar, whereas the associative use would be an inferrable and thus classified as at most uniquely identifiable. As in the case of (33), this would seem to be a case where the writer codes something as at least uniquely identifiable and it is irrelevant whether hearers will identify it by constructing a new representation or by retrieving an existing one from memory.

A second principled source of disagreement, related to the first, arose from the difficulty in determining whether an entire group of multiple addressees can be expected to already have a representation of the referent in long-term memory. Take, for example, the phrase *the skyline* shown in (35):

[9] Fraurud (1992:82) makes a similar point. She writes "The presence of a coreferent NP somewhere in the preceding discourse does not, however, necessarily imply that the interpretation of the [definite NP] must rely on this previous mention..." Gundel and Mulkern (1998) extend this point to definite article NPs with full descriptions (e.g. "the Chinese restaurant on Poland Street" as an answer to 'Where should we eat tonight?"), whose referents may be familiar from general experience. Taking a relevance-theoretic approach to language interpretation, they note that the extra processing effort required to search memory for a familiar referent in such cases would often be gratuitous, since an optimally relevant interpretation could be arrived at on the basis of descriptive content encoded in the phrase itself.

(35) Quebec City, according to a number of surveys, is the port of call preferred by passengers on this route. With its harbour a stone's throw from its historic and much visited centre, with the Château Frontenac that dominates **the skyline** and its unique location." [Debates of the House of Commons of Canada 5.7.97]

Two of us coded this as at most uniquely identifiable and one of us coded it as familiar because we initially disagreed over whether all Canadian Members of Parliament could be expected to already have a representation of the Quebec City skyline in long-term memory. We settled on the coding of this NP as uniquely identifiable because it's not clear if the referent is necessarily familiar to ALL the addressees and, as in the examples (33) and (34), it also doesn't matter whether or not it is. This was also the majority opinion.

4. CLEFT PRESUPPOSITIONS

Definite article phrases have often been analyzed as involving an existential presupposition (Strawson 1950). For example, in uttering *The present King of France is bald*, the speaker presupposes that there exists a present King of France. According to Stalnaker (1974) such "pragmatic presuppositions" indicate that the referent is present in the "common background," and thus that the referent is familiar to both speaker and addressee. Lewis (1979) proposes that referents of definite descriptions that are not in fact present in the common ground can sometimes be "accommodated" and thus treated as if they were in fact in the common ground. As discussed earlier, the GHZ account obviates the need for the mechanism of accommodation with definite descriptions since it treats the condition on the use of the definite article as unique identifiability rather than familiarity. The condition is that the addressee be able to assign a representation to the referent by the time the definite description is processed. It would be interesting to explore whether other well-known cases of presupposition can be given an account parallel to that of definite descriptions, thereby obviating the need to invoke accommodation. We suggest that cleft sentences are instances of this type.

We thus close by turning to a brief discussion of cleft sentences in English, as exemplified in (36):

(36) By and large, **it's the woman who buys the house**. Most men, in my experience, let the wife decide... [Terkel, Working, p. 429, Realty Broker]

As mentioned, cleft sentences, like definite descriptions, have generally been analyzed as involving a familiarity presupposition, which sometimes needs to be accommodated. However, Hedberg (2000) argues that the cleft pronoun and the cleft clause in a cleft sentence function together pragmatically as a discontinuous definite description, and that unique identifiability rather than familiarity is thus the condition that the cleft clause must meet. Once again, the condition is that the addressee is able to assign a representation to the content of the cleft clause by the time that the cleft clause is processed. Clefts can be found which fulfill all four of the definite statuses on the Givenness Hierarchy. For example, the cleft clause of the truncated cleft in (37) is necessarily in-focus:

(37) After the preliminaries were completed and the Coroner had explained to the jury what they were there for and how serious a matter it was, he proceeded to call witnesses. My heart beat fast, for I had thought that as the discoverer of the body I would <u>be the first to be called</u>; but to my surprise, **<u>it</u> was Marcel**. He stepped forward, neat, dark, debonair...
[Mary Fitt, Death and the Pleasant Voices, p. 156]

Here, the elided material in the truncated cleft *it was Marcel* was brought into focus by the phrase *be the first to be called*, which occurs at the end of the immediately preceding clause.

The cleft clause of a *this*-clause as in (38) is necessarily activated, but it need not be in focus.

(38) I wasn't surprised by the massacre in China. [pause] **This is not Iowa we're talking about**—this is a different society [Eric Severid, CSPAN, 12/31/89]

The fact that something is being talked about is activated in (38), but the focus of attention is on the content of the talk itself.

Hedberg argues that the form *this* in (38) acts analogously to the determiner *this* in a *this*-NP and the content of the cleft clause acts analogously to the nominal content of the NP in a *this*-NP.

The cleft clause of a *that*-clause is necessarily familiar, but it need not be activated, as in (39):

(39) N: That's the reason I don't want to go to Miami!

B: Yeah. **Wasn't that somewhere in Southern Florida where they thought those people got AIDS from bug bites**—getting bit a hundred times a night or something, because the place was so roach infested? [breakfast conversation, 2/89]

Here the fact that some people were believed to have contracted AIDS from bug bites is assumed to be familiar to the addressee, but it is not activated since it hadn't been mentioned recently (or at all) in this discourse.

Finally, while the cleft clause of an *it*-cleft is always uniquely identifiable in the sense that the addressee is expected to associate a unique representation of an entity that meets the description of the clause, it does not have to be familiar. Thus, in (40), the addressee is expected to be able to construct a representation corresponding to the fact that a congressman was prompted to strike out on his own against the AIDS epidemic, but he is not expected to have any previous familiarity with this fact.

(40) The federal government is dealing with AIDS as if the virus was a problem that didn't travel along interstate highways and was none of its business**. It's this lethal national inertia in the face of the most devastating epidemic of the late 20th century that finally prompted one congressman to strike out on his own.** [Ellen Goodman, op-ed column, 5/35/87]

This is an example of what Ellen Prince (1978) calls an "informative presupposition" *it*-cleft, where the material in the cleft clause is treated as if it were new to the reader. Prince characterizes this type of *it*-cleft as serving to mark the information in the cleft clause as a "known fact, unknown only to the readership."

The examples in (38)-(40) show that the cleft pronoun functions like determiner *this*, *that* and *the*, respectively, on the Givenness Hierarchy, and the material in the cleft clause functions like the nominal content of a definite noun phrase. Hedberg (2000) follows Postal 1966 and Wirth 1978 in analyzing *it* as the 'intransitive' form of *the*, i.e. *it* is the form that occurs when the descriptive content does not immediately follow.

As the Givenness Hierarchy predicts, the *it*-cleft clause (if it is part of a discontinuous definite description) can contain material of any of the four definite statuses. The cleft clause material is generally characterized in the literature as 'presupposed', and presupposition is generally taken to mean information that is assumed to be part of the common ground of the

discourse, and thus already familiar to the addressee. As Delin 1989 (p. 121) points out, however, Prince's "informative presupposition" *it*-clefts cannot strictly be taken to contain presupposed information in this sense. Instead, she suggests that the cleft clause content in such clefts is accommodated by the hearer. She concludes that "clefts mark the fact that antecedents have to be found or created for particular information they convey."

Hedberg's account makes no appeal to accommodation, however. She claims simply that the content of the cleft clause has to be uniquely identifiable, whether familiar or not. As we have seen, a large number of definite determiner phrases, 44% of our naturally occurring examples, are only uniquely identifiable, not familiar. Given the parallels between cleft constructions and definite descriptions outlined above, it is of interest to compare cleft clauses with definite article phrases to see to what extent cleft clauses mark information that is uniquely identifiable but not familiar. Delin analyzed a corpus of 50 spoken *it*-clefts and found that 36 or 72% of them contained new or inferrable information in the cleft clause. Hedberg 1990 analyzed 701 naturally occurring *it*-clefts in a variety of spoken and written genres, including casual conversations, television discussions, mystery novels, newspaper op-ed pieces, and historical narratives, and concluded that 49% contained new or inferrable information. The distribution is shown in Table 3.

	cleft clause material	*spoken*	*mysteries*	*op-ed pieces*	*historical narratives*	*TOTAL*
familiar	activated	24	66	7	2	99
	reactivated	22	58	21	8	109
	familiar	20	47	43	41	151
SUBTOTAL						359 (50.8%)
non-familiar	cause	19	57	28	8	112
	consequence	7	64	8	12	91
	superlative	7	16	8	8	39
	informative	11	58	25	11	105
SUBTOTAL						347 (49.2%)
TOTAL		110	366	140	90	701 (100%)

Table 3. Distribution of cleft clause material in spoken and written texts of diverse genres

Analogously to the case of definite article phrases, then, the large proportion of cleft clauses that are non-familiar (49-72%) supports the argument against an analysis that treats them as exceptional cases to be handled by the mechanism of accommodation. An analysis of the cleft pronoun + cleft clause as a uniquely identifiable definite description captures the similarity between definite article phrases and cleft sentences. Furthermore, there is no need to appeal to the notion of accommodation for either definite article phrases or cleft sentences on this account.

5. CONCLUSION

This paper has provided further empirical support for the view that 'familiarity' is not the basic meaning of the definite article in English. We have shown that both familiar and non-familiar uses of the definite article can be adequately accounted for within the Givennness Hierarchy framework of Gundel, et al. (1993). On this account, the definite article is unspecified for familiarity. The relatively high frequency of definite article phrases referring to familiar entities is consistent with the grammar (specifically with the lexical meaning of unique identifiability associated with the definite article), but is not determined by it. No special mechanism is therefore required to explain non-familiar uses. We have further shown that this account can be naturally extended to cleft sentences where the cleft clause represents material that is not previously familiar to the addressee, assuming that the cleft pronoun and cleft clause constitute a discontinuous definite description, as proposed in Hedberg 2000.

Our account of definite article phrases should be testable experimentally. A familiarity account that relies on accommodation would predict that non-familiar definite article phrases should take longer to process than familiar definite article phrases since they would require an extra step. On our account, there should be no such difference in processing time. Such an empirical test of the two accounts awaits future research.

REFERENCES

Ariel, M. (1988). Referring and accessibility. *Journal of Linguistics* 24.67-87.

Ariel, M. (1990). *Accessing noun-phrase antecedents*. London: Routledge

Birner, B. & G. Ward (1994). Uniqueness, familiarity, and the definite article in English. *Proceedings of the Annual Meeting of the Berkeley Linguistic Society* 93-102.

Chafe, W. L. (1994). *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago: Chicago University Press.

Clark, H. H. & S. E. Haviland (1977). Comprehension and the given-new contract. In Freedle, R (ed.) *Discourse Processes: Advances in Research and Theory, Volume 1: Discourse Production and Comprehension*. Norwood, NJ: Ablex. 1-40

Cornish, F. (1996) Antecedentless' anaphors: deixis, anaphora, or what? Some evidence from English and French, *Journal of Linguistics*, 32, 19-41.

Cornish, F. (1999). *Anaphora, Discourse, and Understanding: Evidence from English and French*. Oxford: Clarendon Press.

Christophersen, P. (1939). *The Articles: a Study of their Theory and Use in English*. Oxford: Oxford University Press.

de Mulder, R. (2000). Démonstratifs et accessibilité. *Verbum* 22: 103-125.

Delin, J. L. (1989). *Cleft constructions in discourse*. University of Edinburgh dissertation.

Donnellan, K. S. (1966). Reference and definite descriptions. *Philosophical Review* 75: 281-364.

Fraurud, K. (1990). Definiteness and the processing of NPs in natural discourse. *Journal of Semantics* 7: 395-433.

Fraurud, K. (1992). Processing Noun Phrases in Natural Discourse. Stockholm University dissertation.

Grice, H.P. (1975). Logic and conversation. In Cole, P. & J. Morgan (eds.) *Speech Acts.* New York: Academic Press. 41-58.

Gundel, J. K. (1996). Relevance theory meets the givenness hierarchy: an account of inferrables. In Fretheim, T. and J. Gundel (eds). *Reference and Referent Accessibility*. John Benjamins & Co, 141-153.

Gundel, J. K.; N. Hedberg; and R. Zacharski (1988). The generation and interpretation of demonstrative expressions. *Proceedings of the XIIth International Conference on Computational Linguistics.* John Von Neumann Society for the Computing Sciences, Budapest, 216-221.

Gundel, J. K.; N. Hedberg; & R. Zacharski (1989). Givenness, implicature, and demonstrative expressions in English discourse. *Chicago Linguistic Society*. 25/2: 89–103.

Gundel, J. K.; N. Hedberg; & R. Zacharski (1990). Givenness, implicature, and the form of referring expressions in discourse. *Berkeley Linguistics Society.* 16: 442–453.

Gundel, J. K.; N. Hedberg; & R. Zacharski (1993). Cognitive status and the form of referring expressions in discourse. *Language* 69: 274-307.

Gundel, J. K.; N. Hedberg; & R. Zacharski (2000). Statut Cognitif et Forme des Anaphoriques Indirects. *Verbum* 22: 59-78.

Gundel, J. K., & A. Mulkern (1998). Quantity implicatures in reference understanding. *Pragmatics and Cognition: Special Issue on the Concept of Reference in the Cognitive Sciences*. 6: 21-45.

Hawkins, J. A. (1978). *Definiteness and Indefiniteness: A Study in Reference and Grammaticality Prediction*. London: Croom Helm.

Hawkins, J. A. (1991). On (in)definite articles. *Journal of Linguistics* 27: 405-42.

Hedberg, N. (1990). *Discourse Pragmatics and Cleft Sentences in English*. University of Minnesota dissertation.

Hedberg, N. (2000). The referential status of clefts. *Language* 76: 891-920.

Heeman, P. A. & J. F. Allen (1995). The TRAINS-93 dialogues. TRAINS technical note TN 94-2, University of Rochester Department of Computer Science, Rochester, NY.

Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. MIT dissertation.

Lambrecht, K. (1994). *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press.
Lewis, D. (1979). Score-keeping in a language game. *Journal of Philosophical Logic* 8: 339-359.
Lyons, C. (1999). *Definiteness*. Cambridge: Cambridge University Press.
Matsui, T. (2000). *Bridging and Relevance.* John Benjamins.
Poesio, M. & R. Vieira (1998). A corpus-based investigation of definite description use. *Computational Linguistics* 24: 183-216.
Postal, P. M. (1966). On so-called pronouns in English. In Dinneen, F (ed.) *The Nineteenth Monograph on Language and Linguistics*. Washington, D.C.: Georgetown University Press. 291-224.
Prince, E. (1978). A comparison of *wh*-clefts and *it*-clefts in discourse. *Language* 54: 883-906.
Prince, E. (1981). Toward a taxonomy of given-new information. In Cole, P (ed.) *Radical Pragmatics*. New York: Academic Press. 223-256.
Prince, E. (1992). The ZPG letter: Subjects, definiteness, and information status. In Thompson, S & W. Mann (eds.) *Discourse Description: Diverse Analyses of a Fund-Raising Text*. New York: John Benjamins. 295-325.
Sperber, D & D. Wilson (1986). *Relevance: Communication and Cognition*. Cambridge, MA: Harvard University Press. (second edition published in 1995)
Stalnaker, R. C. (1974). Pragmatic presuppositions. In Munitz, M. & P. Unger (eds.) *Semantics and Philosophy*. New York: New York University. 197-214.
Strawson, P. F. 1950. On referring. *Mind* 61.320-44.
Wirth, J. 1978. The derivation of cleft sentences in English. *Glossa* 12.58-91.

Jeanette K. Gundel
gunde003@umn.edu
University of Minnesota and NTNU, Trondheim
215 Nolte Center, 305 Pillsbury Ave. S.E.
University of Minnesota
Minneapolis, MN
USA

Nancy Hedberg
hedberg@sfu.ca
Department of Linguistics
Simon Fraser University
Burnaby, BC V5A 1S6
CANADA

Ron Zacharski
raz@crl.nmsu.edu
Computing Research Laboratory
New Mexico State University
Box 30001 Dept. 3CRL
Las Cruces, New Mexico 88003
USA