

Cognitive Status, Information Structure, and Pronominal Reference to Clausally Introduced Entities

Jeanette K. Gundel
University of Minnesota
gunde003@umn.edu

Michael Hegarty
Louisiana State University
mhear1@lsu.edu

Kaja Borthen
NTNU, Trondheim
kaja.borthen@hf.ntnu.no

Abstract. This paper investigates reference to clausally introduced entities and proposes an explanation for why these are more readily available to immediate subsequent reference with a demonstrative pronoun than with the personal pronoun, *it*. New evidence is provided supporting proposals that such entities are typically activated, but not brought into focus, upon their introduction into a discourse. The study also provides further insight into the role of information structure, lexical semantics, presuppositional contexts, and syntactic structure in bringing an entity into focus of attention.

Key words: reference, givenness hierarchy, presupposition, topic, focus

1 Introduction

Entities introduced by a clause, sequence of clauses, or other non-nominal expressions are accessible to immediate subsequent reference with demonstrative pronouns, but comparatively less accessible to reference with the personal pronoun *it* (Webber 1988, 1991 *inter alia*).¹ Webber (1991) found that of 96 references to clausally introduced entities in written English texts, only 15 used the personal pronoun *it*, while the rest were *this* or *that*. Similarly, in a pilot study investigating reference to entities introduced by non-nominal expressions in both spoken and written data, Hegarty et al (2002) found that 70 out of 95 pronouns were demonstratives. Moreover, substitution of *that/this* with *it* often results in infelicity, as in (1) and (2) or a different interpretation, as in (3) and (4).

- (1) People are excited to know more about where their food comes from and who makes it. That has been one of the most motivating aspects of this experience so far. (B. Dooley ‘Shepherd’s Way Farms’, Twin Cities Natural Food Co-Ops, 6/02)

#It has been one of the most motivating aspects ...so far²

- (2) ... “trying to be happier [may be] as futile as trying to be taller and therefore is counterproductive.”...Do we really believe that Romanian orphan babies left alone in their beds will have the

¹ Examples here will be from English, but similar restrictions can be found in other languages.

² A # indicates unacceptability in the given context, in the absence of special contextual assumptions such as multiple previous mention of the referent. Judgements are those of the authors, corroborated by an informal survey of native speakers and by discussion during conference presentations.

same potential for happiness as those raised by caring parents of ample means? That is precisely what quotes such as those above will be taken to imply. (R. Cook-Deegan, 2001. Hype and hope. *American Scientist* 89.1:62-64.)

It is precisely what quotes such as those above will...

- (3) a. Wheat is versatile and available, so it is made into starches, thickeners, flavorings and even used to dust the rollers at candy factories, and that is a problem when someone really allergic needs some chocolate. (R. Asbell, Gluten-free, Wheat free diets, Twin Cities Natural Food Co-Ops, June 2002)
- b. Wheat is versatile and available, so it is made into starches, thickeners, and even used to dust the rollers at candy factories, and it is a problem when someone really allergic needs some chocolate.
- (4) a. “We believe her, the court does not, and that resolves the matter,” Mr. Montanarelli said today of Ms. Lewinsky’s testimony that ... (NY Times, 5/24/ 00)
- b. “We believe her, the court does not, and it resolves the matter,” Mr. Montanarelli said today of ...

Note that such facts cannot simply be attributed to competition with other potential referents and structures. For example, competition with the court as a possible referent is not the reason *it* cannot be taken as referring to the fact associated with the content of the first two clauses in (4), since the facts remain the same even if the non-clausally evoked ‘competitor’ is removed, as in #*We believe her, and it resolves the matter*.

In previous work (Borthen et al 1997, Gundel et al 1999, and Hegarty et al 2002) we argued that facts regarding the distribution and interpretation of *this/that* vs. *it* referring to entities introduced by a non-nominal expression can be explained within the theory of reference and cognitive status proposed by Gundel, Hedberg and Zacharski (1993 and earlier work). The present paper builds on this previous work, extending it in two primary directions. First, we expand the database of examples from naturally occurring discourse to test and refine hypotheses put forward in the earlier work, extending the analysis to account for reference to entities introduced by clauses occurring in a variety of presuppositional contexts, and by non-finite clauses. Second, we further explore the general question of how linguistic factors promote the salience of discourse entities, and how discourse semantics and syntactic structure interact with

information status in bringing an entity into focus of attention. We begin with a brief summary and further elaboration of the earlier work.

2. The Cognitive Status of Clausally Introduced Entities

Gundel et al (1993) propose that determiners and pronouns constrain possible interpretations of nominal forms by conventionally signaling one of six cognitive statuses that the intended referent is assumed to have in the mind of the addressee. The array of statuses, called the Givenness Hierarchy (GH), is presented in Fig. 1 below

Figure 1 The Givenness Hierarchy (GH) and associated forms in English

in focus > activated > familiar > unique identifiable > referential > type identifiable

it this, that, this N that N the N indefinite this N a N

Statuses on the hierarchy correspond to memory and attention states, ranging from most restrictive, ‘in focus’, to least restrictive, ‘type identifiable’. By being associated with different statuses as part of their conventional meaning, forms thus serve as processing signals that assist the addressee in restricting possible interpretations.³

The statuses are in a unidirectional entailment relation. If something is in focus (center of attention), it is necessarily activated (in working memory); if it is activated, it is familiar (in memory); if it is familiar, then it is uniquely identifiable (the addressee can associate a unique representation); if it is uniquely identifiable, then it must be referential; and if it is referential, it must be type identifiable. The theory thus correctly predicts that a given cognitive status can be appropriately coded by a form which explicitly signals that status, but also by forms whose meanings are entailed by that status. In the latter case (e.g. use of a definite article for a referent that is in focus) the

³ Note that the statuses themselves, like other ‘meanings’, are independent of language. Thus, a mental representation associated with a non-linguistic stimulus, e.g. a house or a tune, can also be in focus or not, activated or not, familiar or not, and so on. The cognitive status of some entity should thus also be determinable independent of the form that is used to refer to it or of whether it has been referred to at all. For example, something may be familiar because it was previously mentioned in the discourse or because it is known from general experience, and it may be in focus because it was recently introduced in a syntactically prominent position or because the addressee is intently looking at it.

form is simply underspecified for cognitive status of the intended referent.⁴

Of relevance here are the statuses ‘activated’ and ‘in focus’. An entity is activated for an individual if that individual has a representation of it in working memory; and it is in focus if it is activated and, moreover, at the center of that individual’s attention. As seen in Figure 1, Gundel et al propose that unstressed personal pronouns, including *it*, are appropriately used only when the referent can be assumed to be in focus for the addressee prior to processing of the referring form. Demonstrative pronouns *that/this*, however, are unspecified for the status ‘in focus’, and can be appropriately used when the referent can be assumed to be at least activated, whether or not it is also in focus. This proposal is supported by the fact that acceptability of *it* is associated with such arguably salience promoting factors as introduction in a syntactically prominent position and multiple prior mention of the referent, whereas demonstratives *this* and *that* often signal a focus shift. It also permits an explanation of facts like those in (1)-(4) if entities (indirectly) introduced by a whole clause, or sequence of clauses, are activated, but are less likely to be brought into focus than entities introduced by syntactically prominent nominal constituents (Borthen et al 1997, Gundel et al 1999, Hegarty et al 2002). For example, it is reasonable to assume that in (1), at the conclusion of the first sentence, the fact that people are excited to know more about where their food comes from and who makes it is activated, but not necessarily in focus. The focus of attention after the sentence is processed is more likely to be on food, which has been mentioned twice in the preceding two sentences. Similarly, in (3), the fact that wheat is used to dust the rollers at candy factories is rendered activated by the second clause, and can thus be referred to with *that* in (3a). But there is no reason to assume this fact to be in focus, as would be the case for the wheat, which has been mentioned twice in subject position, and is thus the most natural interpretation of the pronoun *it* in (3b).

A fact or proposition introduced by an NP/DP, as in (5) and (6), is thus also more likely to be brought into focus than one introduced by the whole clause.⁵

⁴ The use of underspecified forms has limits, however, due to interaction of the GH with general pragmatic principles involved in language production and understanding (Grice 1975, Sperber and Wilson 1986/95). Gundel et al (1993) argue that the implicational nature of the GH gives rise to ‘scalar implicatures’ which further restrict the distribution and interpretation of referring forms (see also Gundel and Mulkern 1998). For example, in English, the indefinite article is associated with non-familiarity because its use typically implicates by the first part of the Quantity Maxim that conditions for using a more restrictive form are not met. Another result of interaction of the GH with the Quantity Maxim is that most in-focus referents are not coded with demonstratives, even though they could be; and demonstratives often implicate a focus shift.

- (5) I read about an interesting fact yesterday. It shocked me.
- (6) Life lurking in the softly seeping springs of the Martian highlands would be one of the defining discoveries of the 21st century. **It** would also be a body blow to the Gaian notion that life is necessarily a planet-wide phenomenon (New Scientist, July 8, 2000)

As we noted in Hegarty et al (2002), a factor that seems to have an effect on whether or not a clausally introduced entity can be referenced with *it* is the degree of world immanence of the entity and, correlatively, its manner and degree of individuation (Asher 1993). Events and states, which have causal, spatial and temporal properties, have high world immanence; “purely abstract objects” such as propositions have very low world immanence, and their individuation properties depend more on the means we use to describe them than on independent properties of objects in the world. Facts and situations are in between. Events, whose individuation properties are largely independent of the means we use to describe them, have referential properties similar to those of concrete objects and other referents denoted by nominal constituents, as seen in (7) and (8), where either *it* or *this/that* can refer to the event described in the first clause.⁶

- (7) a. John insulted the ambassador. That happened at noon.
 b. John insulted the ambassador. It happened at noon.
- (8) a. What do you think he would do if Taylor’s baby died?
 Do you think he’d just go on with his life like nothing mattered? Like it didn’t even happen? (“The Bold and the Beautiful”, Jan. 30, 2001, CBS)
- b. What do you think he would do if Taylor’s baby died? Do you think he’d just go on with his life like nothing mattered? Like that didn’t even happen?

Such facts are explained if the individuating properties that events share with referents of nominal constituents make it more likely that

⁵ Hegarty (forthcoming) argues that propositions, facts and situations introduced by nominals are of semantic type *e*, and therefore have the semantic ontology of ordinary, discrete individuals, by virtue of which they can compete with ordinary individuals for the attention of participants in a discourse. He argues by the same criteria that propositions, facts and situations introduced by clauses have a type-raised denotation, less cognitively accessible, and therefore at a disadvantage in competing for the attention of participants in a discourse.

⁶ Since *that* merely requires activation, and anything in focus is also activated, in focus entities can be referenced with either *that* or *it*.

they will be brought into focus immediately subsequent to their introduction with a full clause, since the event described by a sentence is directly (and necessarily) introduced by virtue of processing that sentence. The addressee, in processing the first sentence in (7), for example, posits a relation ‘insult’ between John and the ambassador, and this relation involves an event of John insulting the ambassador. In the terms of Discourse Representation Theory, with an underlying event semantics for active verbs, the introduction of **insult** (u, v, e), into a DRS, for discourse entities u, v satisfying **John**(u) and **ambassador** (v), requires a discourse entity e for the event in which John broke the vase. Example (9) below might seem contrary to this observation, since *that/it* in the second sentence of this example refers to an event introduced in the previous sentence, yet it is not felicitously referenced with *it*.

- (9) Gods usually come about after people settle down, start tilling crops, and develop hierarchical social structures. None of that happened until almost 10,000 years after these figurines were made. (Science, v. 263, 1994, p. 923)

Cf. # None of it happened until almost 10,000 years after these figurines were made.

However, *that* in (9) refers to a complex event whose parts were introduced in the previous sentence, while the complex event itself was not. The complex event is thus activated by the first sentence, but it is not brought into focus since additional processing is required to construct the complex event from its component parts, hence the unacceptability of *it* here. (See Hegarty forthcoming for further discussion)

Situations are less accessible to reference with *it*, as in (10).

- (10) a. John insulted the ambassador. That/this was intolerable.
b. John insulted the ambassador. ??It was intolerable.

The first sentence in (10), as in (7) above, directly introduces the event of John’s insulting the ambassador, thus activating that event and most likely also bringing it into focus. However, the pronominal subject of the second sentence in (10), unlike that in (7) refers to the situation resulting from the event, not the event itself. The predicate *intolerable* in (10) precludes an interpretation on which its subject refers to the event of John insulting the ambassador, since an event is unchangeable once it has occurred, and thus cannot fail to be tolerated. The situation of John insulting the ambassador, in contrast, includes its ramifications, and those at least, are subject to amelioration or change, making it sensible to say that the situation is

intolerable to the embassy.⁷ The contrast between (10) and (7) can thus be explained by the proposal that the pronoun *it* requires its referent to be in focus, since situations, which are less world immanent than events, and less susceptible to individuation by spatiotemporal extent, are (unlike events) not directly introduced into the discourse when a sentence is uttered (unless of course they are introduced as complements of predicates that take situations as their arguments). Since construction of a representation corresponding to the situation requires additional processing on the part of the addressee, the situation is also less likely to be brought into focus simply by virtue of the first sentence being processed. The examples in (1)-(4), show that facts and propositions pattern with situations, and not with events, in their availability for subsequent pronominal reference.

We have proposed that in order for an utterance to bring some entity into focus it is necessary (though not sufficient) that the entity be directly expressed as part of the conventional semantic content of the utterance. If this is true, we would expect that speech acts (i.e. acts performed by an utterance, which are not part of the semantic content, or even inferable from it) are never brought into focus. Although the fact that a particular speech act has been performed can be expected to be activated for the addressee immediately following its performance, the addressee's focus of attention can be expected to be on the semantic content of the utterance used in the act, not on the act itself. Thus, if *it*, unlike *this/that*, requires its referent to be in focus, speech acts should be accessible to subsequent reference with *that* or *this*, but inaccessible to subsequent reference with *it*. This prediction is borne out by the examples in (11)–(13)

- (11) Thorne: So you fired her?
 Eric: We're going to do a lot more than just fire her.
 Thorne: What does that mean? ("The Bold and the Beautiful",
 Jan. 30, 2001, CBS)

What does it mean?

- (12) A. John snores.
 B. That's rude.

⁷ The distinction between facts, situations, events, etc. is not always clear cut, but can usually be determined from the semantics of the predicate (Asher 1993, Hegarty forthcoming). Thus, the same clause or phrase may refer to a different type of entity depending on what predicate it occurs with. For example, the phrase *John's insulting the ambassador* refers to an event in *John's insulting the ambassador happened at noon* (cf. *The event happened at noon, but ??The situation happened at noon*), but it refers to a situation in *John's insulting the ambassador was intolerable* (cf. *The situation was intolerable, but ??The event was intolerable*).

B'. It's rude.

- (13) A. I just ate three pieces of cake.
 B. Can you repeat that.
 B'. ? Can you repeat it.

In (11), the demonstrative *that* is interpreted as referring to Eric's statement, "We're going to do more than just fire her". This interpretation is impossible if *that* is replaced with *it*, and the resulting sentence is thus unacceptable in this context.⁸ In (12), the demonstrative *that* in (B) is ambiguous between an interpretation where it refers to the act of John snoring and one where it refers to A's illocutionary act of informing B of this fact. In contrast, (12B') can only have the former interpretation. Similarly, *that* in (13B) is ambiguous between an interpretation where the speaker is being asked to repeat the statement in A and one where she is being asked to repeat the act of eating three pieces of cake. But (13B'), if it is acceptable at all, can only be a request to repeat the act of eating the cake; it cannot be interpreted as a request to repeat the act of saying that she ate three pieces of cake.

2.1. The Role of Presupposition and Prior Beliefs

We have argued that the reason clausally introduced entities such as facts and situations are relatively inaccessible to subsequent reference with the personal pronoun *it* is that the clause typically does not bring these entities into focus of attention. Our research has shown, however, that various semantic and pragmatic factors which play a role in boosting the salience of discourse entities can also affect the distribution of *it* in referring to clausally introduced entities. These include covert arguments, presuppositions, and even inquisitive looks, all of which can cause an entity to be "reprocessed", and thus brought into focus, even when it is overtly mentioned only once. Consider (14) and (15), for example (from Borthen et al 1997). In (14), the speaker, upon clausally introducing the fact that linguists earn less than psychologists, can assume that this fact is rendered activated, but not in-focus, for the hearer, leading to a preference for *that* over *it* in the follow-up reference to this fact.

- (14) a. I hear linguists earn less than psychologists, and that's terrible.
 b ?? I hear linguists earn less than psychologists, and it's terrible.

⁸ Note that this is true regardless of whether primary stress in the sentence falls on *that* or on *mean*.

In (15), in contrast, the follow-up reference is made by another speaker, which results in somewhat more complicated inferences regarding the cognitive status of the fact at issue.

- (15) A: I just read that linguists earn less than psychologists.
 B: **That's** terrible!
 B': It's terrible!

At the completion of A's utterance, B can assume that the fact that linguists earn less than psychologists is at least activated for A. In response B, B's use of *that* overtly signals the assumption that this fact has been activated. This is not inconsistent with it also being in focus, as anything in focus is by definition also activated. But, as Gundel, et al (1993) argue, use of the demonstrative form, which overtly signals only activation, often implicates, by the first part of the Quantity maxim, that the referent is not in focus. This invites A to infer that the fact is news to B. In response B', B's use of *it* signals the assumption that the fact is in focus, or ought to be, as would be consistent with it already having been accepted background information for discourse in the relevant social circle; this invites A to infer that his initial utterance brought the fact into focus for B, since B already knew the fact, and its mention was therefore sufficient to bring it into focus. A similar example is provided in (16).

- (16) Marah: ...she won't let me and Shayne stay here alone, but
 can we stay with friends who will gladly have us?
 No, we're stuck with my dad and Olivia.

Tony: That's lousy. (Guiding Light 2/28/01. CBS)

Cf. Tony: It's lousy.

Tony's response in (16), where *that* refers to the fact that Marah is stuck with staying with her dad and Olivia, is neutral with respect to whether or not he was previously aware of this fact; but the most natural interpretation is that he wasn't. If *that* is replaced with *it*, on the other hand, the interpretation is that the fact is not news to Tony.

Consider also (17) (adapted from Kamio and Thomas 1999).

- (17) A: Janice fired her secretary yesterday.
 B1: Yes. Everyone in the office is aware of that.
 B2: Really? The people in the office weren't aware of that.
 B3: Yes. The people in the office are aware of it.
 B4: *Really? The people in the office weren't aware of it.

In (17), the fact that Janice fired her secretary yesterday is rendered activated by A's assertion, thus making it possible to refer to this fact with *that* in B1 or B2. When B's reply opens with *Yes*, signaling that the fact in question was already familiar to B, B can thereupon assume that A's mention of the fact has made it mutually manifest that the fact is in focus for both A and B, thus also permitting reference with *it*, as in B3. But when B's reply begins with an indirect admission that this fact is new to B, replacement of the demonstrative with the personal pronoun *it* is correspondingly infelicitous, as in B4.

In (18) below, the proposition that B has a dental appointment at 2 is clausally introduced by A. If the mere utterance of a sentence does not bring the expressed proposition into focus, this would explain why (18B') sounds unnatural, given that *it* requires the referent to be in focus, whereas *that* merely requires activation.

- (18) A: You have a dental appointment at 2.
 B: That's true. B': ??It's true. B'': It's true, then.

But (18B'') is noticeably more acceptable than (18B'). Gundel et al (1999) suggest a relevance-theoretic explanation for this fact (Sperber and Wilson 1986/1995), noting that *then* in B'' functions as an interpretive particle which conveys the meaning that the content of the sentence it is appended to follows by way of inference from something that was just said. The response in (18B'') means essentially, "Given your assertion that I have a dental appointment at 2, then I can take it as confirmed that I have a dental appointment at 2" (Hegarty et al (2002)). Since this can yield contextual effects for A only if A's utterance confirmed the truth of a proposition that B had been questioning, the fact that B had a dental appointment at 2 was not activated for the first time by A; rather, A's utterance brought into focus a fact that was already mutually manifest to both A and B beforehand, thereby licensing the use of *it* in B''

The examples in (14) through (18) show that clausally introduced propositions, facts or situations are more accessible to reference with *it* if they have already been entertained (i.e. mentally represented) by the addressee or if they are otherwise 'reprocessed' subsequent to their introduction, both of these factors being ones that can also be expected to raise the salience of the entity in question. Similar factors can be invoked to explain the differential behavior of entities introduced by clausal complements of bridge verbs as compared to factive verbs. As argued in Hegarty (2001) and Hegarty et al (2002), complements of bridge verbs, when they represent the information-structural focus of the sentence, are typically accessible to subsequent reference with a demonstrative, but not with the pronoun *it*⁹. This is illustrated by the constructed examples in (19)¹⁰,

⁹ By information structure, we mean a bifurcation of material in an utterance

as well as the naturally occurring examples in (20) and (21).

(19) What does Alex think?

A: Alex believes [_F the company destroyed the FILE].

B: That's false; the file was submitted to the judge.

B': # It's false; the file was submitted to the judge.

(20) Noone is quite sure of the purpose of the little sculptures, about 200 of which are known. Some scientists have speculated that they are fertility figures. But anthropologist Patricia White of West Virginia University says that many are not pregnant and none has a baby, casting doubt on that idea. White says they may have a "mythological" role - i.e. a role in explaining "how things came to be the way they are." That seems likely for at least one of the newly rediscovered figures. (Science, vol. 263, 1994, p. 923)

cf. # It seems likely for at least one of the newly discovered....

(21) In 1991, psychologists Diane Halpern of California State University and Stanley Coren of the University of British Columbia reported that left-handers, perhaps because of accidents and weakened immune systems, die on average 9 years earlier than right-handers. Then, several studies said that was bosh. (Science, vol. 263, 1994, p. 1567)

but cf. # Then, several studies said it was bosh.

Similar facts are illustrated by examples like (22), where the infinitival complement denotes a possible but unrealized event or state of affairs, futurate with respect to the reference time of the main clause. (Such an event or state is sometimes called 'irrealis'.)

(22) A: I want [(for) the governor to meet with us].

B: That would almost certainly get things moving.

B': ? It would almost certainly get things moving.

If the speaker of (22A) is introducing the idea of meeting with the governor, the follow-up in (22B) is natural, and the follow-up in (22B') less natural. Note however that the follow-up with (22B')

into what has been called focus versus ground, comment versus topic, or rheme versus theme. This notion is not to be identified with contrastive focus or with the more general distinction between new versus old information. Information structural focus is also distinct from the cognitive status 'in focus'. See Vallduví (1990) and Gundel (1999a) for more detailed discussion of related terminological and conceptual issues. We will indicate information structural focus by the subscript 'F'.

¹⁰ This constructed example was tested in a small survey of English speakers.

works if spoken with a certain rising intonation peaking on the first syllable of *moving*, signaling that the consequences of a meeting with the governor have been thought through, and that (22B') is a considered judgement. The situation here is thus similar to that in (15) above. In both cases, possibility of reference with *it* is correlated with an interpretation that implies prior knowledge and/or processing, which would also be expected to boost salience, making it more likely that the entity in question was now in focus.

On the other hand, entities introduced by complements of factive verbs are equally accessible to reference with a demonstrative or a personal pronoun, as seen in (23).

- (23) A. Alex verified that the company destroyed the file.
 B. That's false; the file was submitted to the judge.
 B'. It's false; the file was submitted to the judge.

Thus, the contrast in (19) between subsequent reference with *it* versus *that* is not exhibited in (23). This is because the semantics of the factive verb enforces the condition that the entity expressed by the complement clause be already familiar (or at least capable of being accommodated as familiar) to the addressee, so that its mention in A's utterance renders this entity in-focus.¹¹

Using a situation variable in the semantics, in the context of Discourse Representation Theory (Kamp and Ryle 1993), the interpretation of the factive ascription in (23A) can be expressed by the Discourse Representation Structure (DRS) shown in (24) below.

(24)

| |
|--|
| u, v, z, s |
| Alex (u) |
| Company (v) |
| File (z) |
| destroy (v, z, s) (w_o) |
| verify ($u, \lambda w[$ destroy (v, z, s)(w)]) |

In contrast, a belief ascription such as that in (19A), using a bridge verb, is interpreted semantically as just a relation between Alex and the proposition expressed by the complement clause. A DRS for (19A) is presented in (25).¹²

¹¹ As shown in Hegarty et al (2002), this pattern is also obtained in interrogatives, which share presuppositional effects with factives, as well as with complements to certain non-factive (and non-bridge) predicates which share with factives the property of being felicitous when the proposition or situation expressed by the complement clause is already accepted as given.

¹² The ascription made by A in (19) might actually express a proposition already familiar to the hearer. The property distinguishing bridge verbs is not

(25)

| |
|--|
| u, v, z |
| Alex (u) |
| Company (v) |
| File (z) |
| believe ($u, \lambda w[\exists s[\text{destroy} (v, z, s)(w)]]$) |

Presuppositional contexts generally, and not just factive ones, are capable of boosting the salience of a clausally introduced entity. In (26) below, the complementizer *while* or *although* introducing the adverbial subordinate clause is a presupposition trigger (Levinson 1983), and the fact expressed by the finite clause complement of *tell*, within the scope of *while* or *although*, is subject to immediate subsequent reference with *it*.

(26) While/although we had told Max that Susan is coming to town, it wasn't on his mind when he called today.

The same result is obtained in (27), where the futurate unrealized eventuality mentioned within the scope of *while/although* is subject to immediate subsequent reference with *it*.

(27) While/although Jim told him not to raise the issue of benefits at the first interview, Max did it anyway.

? Max did that anyway.

Interestingly, subsequent reference with *it* in (27) is even preferred to reference with *that*. Contrast this result with a parallel example, absent the presupposition trigger, as in (28).

(28) a. We told Max that Susan is coming to town, but that wasn't on his mind when he called.

b. ??We told Max that Susan is coming to town, but it wasn't on his mind when he called.

In (29) below, the same effect is obtained in the scope of another presupposition trigger, *continue*.

(29) Bill continued to assert/claim that Susan had undermined him before the board. We kept telling him that it's / that's not true,

that the content of their complement **must** be unfamiliar, but only that it **can** be.

but he wouldn't listen.

A DRS for (29) is given in (30) below, where n is a reference time beyond which Bill's assertions that Susan had undermined him count as continued assertions, and t^{n-} is the time interval before n , and t^{n+} the time interval after n .

(30)

| |
|--|
| $u, v, z, n, t^{n-}, t^{n+}, s_o, p$ |
| Susan (u) |
| Bill (v) |
| the-board (z) |
| $p = \lambda w \exists s [[\text{undermine}(u, v, s) \ \& \ \text{before}(s, z)](w)]$ |
| $s_o \subseteq t^{n-}$ |
| assert (v, p, s_o) |
| $(\exists S : S \geq 2) [(\forall s \in S) [(s \subseteq t^{n+}) \ \& \ \text{assert}(v, p, s)]]$ |

The penultimate line of the DRS (30) records the presupposition that Bill asserted the proposition in question on occasions before the (continued) assertions of it being reported by the use of (30).

In (31), the cleft construction triggers presuppositionality for the content of the *wh*-clause. The fact mentioned by this clause is correspondingly available for immediate reference with *it*.

- (31) It was Harry who told John that Susan is coming to town.
Harry had found it out, in turn, from Bill.

In these examples, the presuppositionality trigger confers givenness on the clausally expressed entity, to a degree sufficient to render it in focus and therefore accessible to reference with *it*.

2.2. Information Structure vs. Cognitive Status

As the examples in (19)-(22) show, mention within the information structural focus of a sentence does not in itself bring an entity into focus of attention, if the content of the clause is not already assumed to be familiar or otherwise presupposed. On the other hand, as Hegarty et al (2002) point out, information structure does appear to play a role in whether or not a clausally introduced entity is brought into focus of attention when that entity is mentioned in the ground (theme; topic), regardless of whether or not it has been previously introduced in the discourse. Thus, for example, *it* and *that* are equally good in (32), even though the entity in question is mentioned in the clausal complement of a bridge verb.

- (32) A: Alex [_F INSISTS/BELIEVES] that the company destroyed the file.
 B: But that's/it's false; the file was submitted to the judge.

Similarly in (33) [secondary stress on *murdered*]:

- (33) a. Alex is hopeless.
 b. He [_F INSISTS] that Tom was murdered, for example,
 c. --even though there's not a shred of evidence for that/ it.

Presentation of a clausally introduced entity within the topic of an utterance is thus another way to promote salience, and bring the entity into focus, even if it is, in fact, new to the discourse. Since an entity associated with the topic is already at least familiar to the addressee prior to the utterance (see Gundel 1988 *inter alia*), its mention suffices to bring it into the focus of attention, if it does not already have that status, making it available to reference using *it*.

Similar reasoning applies to cases where we have a bridge verb complement which is an information structural focus, but is already previously activated in the discourse. In such cases, the entity introduced by the complement clause is brought into focus and is therefore accessible to reference with *it*, as seen in (34).

- (34) A1: I believe that the company destroyed the file, but not everybody does.
 B1: What does Alex believe?
 A2: Alex believes [_F that the company destroyed the file].
 B2: But it's/that's false; the file was submitted to the judge.

It can also be shown that the effects in presuppositional contexts exemplified in (26)-(29) are independent of information structure. Example (35) below embeds example (29) in a context in which the complement clause to *assert* is in the information-theoretic focus or rheme of the utterance. Yet the possibilities for subsequent pronominal reference are identical to those in (29).

- (35) A: At that time, everyone had their own obsession, which they maintained despite all counterevidence. What was Bill's?
 B: Bill continued to assert [_F that Susan had undermined him before the board]. We kept telling him that it's not true, but he wouldn't listen.

This undercuts an alternative explanation of (26)-(29). It might be supposed that these examples, presented in isolation, unconsciously evoke an interpretation in which the clausally introduced entity

subject to subsequent pronominal reference is, in fact, the topic of the utterance. As in examples (33) and (34) above, this would facilitate promotion of this entity to ‘in focus’ cognitive status, permitting the subsequent reference with *it*. The example in (35), by imposing a contrary information structure, in which the clause at issue is in the informational focus or rheme of the utterance, removes this factor from the picture, showing that lexical properties of the presupposition trigger are responsible for the cognitive status conferred on the clausally expressed entity.

Relational givenness/newness (topic-focus or information structure) and referential givenness/newness (cognitive or information status) are not necessarily coextensive (see Gundel 1980, 1988, 1999, Vallduví 1990, Lambrecht 1994). For example, the information structural focus in (34A2) above represents a proposition that is new in relation to the topic (what Alex believes), but it is referentially given (at least activated) since it was mentioned in A1. The facts described above show that it is referential givenness (i.e. cognitive status of a discourse entity) not relational givenness (topic-focus structure) that determines whether a clausally introduced entity will be brought into focus of attention. Bifurcation into topic and focus thus appears to have an effect on the cognitive status of an entity only when an entity is mentioned (even introduced) within the topic, because this necessarily signals a higher cognitive status for the entity.

2.3. The Role of Syntactic structure.

We saw above that prior information status plays an important role in determining whether a clausally introduced entity is brought into focus and thus rendered accessible to pronominal reference with *it*. In this section, we will show that purely syntactic factors can also serve to bring an entity into focus of attention.

Gundel et al (1993) note that the entities in focus at a given point in the discourse include those activated entities that are likely to be continued as topics of subsequent sentences, and that membership in this set is largely (though not wholly) determined by syntactic structure (cf. the Centering Algorithms of Grosz, Joshi and Weinstein 1983, 1995, Brennan, Friedman and Pollard 1987). For example, as already noted, entities introduced by nominals in a syntactically prominent matrix clause position are more likely to be brought into focus than referents of non-nominal constituents, including whole clauses. Assumptions about the role of syntax in promoting the salience of discourse entities are also supported by psycholinguistic experiments which have shown that entities are most likely to be brought into focus if they are introduced in matrix subject position (e.g. Bock and Warren 1985, Gordon et al 1993, Hudson D’Zmura and Tannenhaus 1998) or in focus position of a cleft sentence (Arnold

1998, Almor 1999.)¹³

The importance of subject position in promoting the salience of discourse entities, thus making them available to subsequent reference with *it*, is illustrated in the following examples (from Gundel et al 1993).

- (36) a. My neighbor's Bull Mastiff bit a girl on a bike.
 b. It's/ That's the same dog that bit Mary Ben -
- (37) a. Sears delivered new siding to my neighbors with
 the Bull Mastiff.
 b. # It's/ That's the same dog that bit Mary Ben.

Since the Bull Mastiff is introduced in matrix subject position (and is most likely also the topic) in (36a), it is reasonable to assume that it has been brought into focus, and can therefore be appropriately referred to with either *that* or *it* in (36b). The pronoun *it* is possible because the intended referent is in focus. The pronoun *that* is possible because anything in focus is also activated. But in (37), where the Bull Mastiff has been introduced in a more peripheral position, we would expect it to be activated but not brought into focus. Therefore, only reference with *that* is possible.

Similar effects of subject position can be shown for clausally introduced entities. As we saw above, accessibility to pronominal reference with *it* for such entities depends on whether they can be interpreted as previously familiar, for example because of the semantic properties of the predicate or because the clause is within the topic (theme, ground). However, as seen in (38) and (39), introduction in subject position is sufficient to bring an entity into focus, independently of semantic or information structural factors.

- (38) A: What surprised you?
 B: [_F That Max got (so) drunk at his fundraiser] surprised me. And it really hurt his image too.
- (39) A: What would surprise you?
 B: [_F For Max to get drunk at his fundraiser] would surprise me. And it would really hurt his image too.

Entities introduced by the subject clause are accessible for reference with *it* in both (38) and (39). In (38), this could be attributed to the use of the past tense in the subject of *surprise*, with a past tense matrix

¹³ Our finding that introduction in the information focus is in itself not sufficient for bringing an entity into focus of attention (see also Gundel 1999) suggests that it is the syntactic prominence of the clefted constituent that is responsible for boosting the salience of the introduced entities, and not, as these researchers assume, its status as an information focus.

verb, i.e. *surprise* seems to be factive here. But with an infinitival clause, as in (39), *surprise* is definitely not factive. Nevertheless, the material introduced in the subject clause is still accessible to reference with *it*. Similarly, in (40) below, the unrealized event introduced with a subject infinitive clause seems to be available for immediately subsequent reference with *it*.

- (40) a. [For the governor to meet with us] would be very helpful.
 b. It would almost certainly get things moving.

Thus, in contrast to complements of non-factive predicates in examples like (19)-(22), entities introduced by clausal arguments of non-factive predicates in subject position are accessible to immediately subsequent reference with *it*. This is exactly what we would expect if the structural prominence of subject position is sufficient to bring an entity into focus regardless of its information status. Note further that these effects are preserved when the clausal argument of *surprise* is extraposed, as in (41)-(42).

- (41) A: What would surprise you?
 B: It would surprise me [_F for Max to get drunk at his fundraiser]. And it would really hurt his image too.
- (42) A: What surprised you?
 B: It surprised me [_F that Max got (so) drunk at his fundraiser]. And it really hurt his image too.

While it might be argued that the position of grammatical subject in (40) inclines us to interpret its content as topical, it is clear that no such explanation can be provided for the facts in (39), (41) and (42), where the context forces an interpretation where the clause is part of the information structural focus.

3. Conclusion

The fact that clausally introduced entities, immediately after their introduction into a discourse, are less likely to be accessible to reference with the personal pronoun *it* than with a demonstrative pronoun can be explained on the basis of the general hypothesis that *it* requires its referent to be in focus whereas demonstrative pronouns *this/that* only require activation. This account is further supported by the fact that reference with *it* is possible in contexts that arguably contribute to the salience of the entity, for example when it already has a degree of givenness/familiarity sufficient to promote it to ‘in focus’ status immediately after its introduction.

This is the case, for example, in complements of factive predicates as

well as other presuppositional contexts, and in the position of information-structural topic (theme, ground). Introduction within the information-structural focus (rheme, comment) however, has no effect in itself in bringing an entity into focus of attention. Such an effect is realized only if the information focus is also syntactically prominent, for example if it is in subject position or in focus position of a cleft sentence. Examination of the contexts in which clausally introduced entities can be subsequently referred to with *it* vs. *this/that* thus yields new insights into the factors that affect the salience of discourse entities, as well as providing further evidence for the claim that cognitive status is a crucial factor in determining acceptability and interpretation of different referring forms.

The overall form of our argument here is inference to the best explanation. If we can explain the distribution of different pronominal forms in referring to clausally introduced entities by invoking independently established features of these forms in signaling different cognitive statuses, then this explanation is to be preferred over one which stipulates an alternative set of correlations for entities introduced with non-nominal expressions. Moreover, the attributions of cognitive status, both in the original Gundel et al work and in our work here, are independently determined and corroborated by boundary conditions of plausibility and other known guides to the salience and cognitive status of discourse entities. These include, for example, the assumption that entities introduced in syntactically prominent positions are more likely to be brought into focus of attention than ones introduced in a less prominent position, that previous familiarity is likely to promote salience, that directly introduced entities are more likely to be brought into focus than ones that must be inferred, or that full sentences do not bring into focus of attention the acts performed in uttering the sentence. This is not to say that it wouldn't be possible to account for the facts in question, at least partially, by simply positing a correlation between certain forms (*it* vs. *this/that* in this case) and certain linguistic contexts (e.g. complements of different types of predicates). Depending on one's goals, such an account might even be preferable, as it would directly align the facts about referring forms and contexts without appealing to cognitive status, and specifically to attention states such as 'activated' and 'in focus', which, given current technology, are difficult to determine by direct empirical means. But it would fail to explain why the form-context correlations are as they are and not otherwise, and would provide little insight into how referring forms are actually processed and interpreted. It would also preclude a distinction between facts that are due to (knowledge of) the language system and more general, non-linguistic factors governing human information processing. Moreover, as we have shown, there is no single structural context that can be directly correlated with the use of *it* vs. *this/that* in referring to clausally introduced entities, and the

relevant factors are sometimes not linguistic at all. This situation is exactly what we would expect, given our finding that information structure affects the distribution and interpretation of *it* vs. *this/that* only to the extent that it is correlated with factors that affect cognitive status. Cognitive statuses (i.e. memory and attention states) are properties of mental representations, regardless of whether or not these have been linguistically evoked. As such they can only be partially captured by direct reference to linguistic contexts.

References

- Almor, A. 1999. "Noun-phrase anaphora and focus: the informational load hypothesis." *Psychological Review* 106.4:748-763.
- Arnold, J. 1998. Reference form and discourse patterns. Ph.D. Dissertation, Stanford University.
- Asher, N.. 1993. Reference to abstract entities in discourse. Dordrecht: Kluwer Academic Publishers.
- Bock, J. K. and R. Warren. 1985. "Conceptual accessibility and syntactic structure in sentence formulation." *Cognition* 21:47-67.
- Borthen, K, T. Fretheim and J. K. Gundel. 1997. "What brings a higher-order entity into focus of attention?" In R. Mitkov and B. Boguraev, eds., *Operational factors in practical robust anaphora resolution*, pp. 88-93.
- Brennan, S. E., M. W. Friedman and C. J. Pollard. 1987. "A centering approach to pronouns." *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics*, Stanford, 155-162.
- Gordon, P. C., B. J. Grosz, and L.A. Gilliom. 1993. "Pronouns, names, and the centering of attention in discourse." *Cognitive Science* 17:311-347.
- Grice, H. P. 1975. "Logic and Conversation." In P. Cole and J.L. Morgan, eds., *Syntax and Semantics 3. Speech Acts*. 41-58.
- Grosz, B. J., A. K. Joshi, and S. Weinstein. 1983. "Providing a unified account of definite noun phrases in discourse." In *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics*, Cambridge, Mass., 44-50.
- Grosz, B. J., A. K. Joshi, and S. Weinstein. 1995. "Towards a computational theory of discourse interpretation." *Computational Linguistics* 21:203-25.
- Gundel, J. K. 1980. "Zero NP-anaphora in Russian: a case of topic-prominence." *Proceedings from the Parasession on Anaphora*. Chicago Linguistic Society, 139-146.
- Gundel, J. K. 1988. "Universals of topic-comment structure." In M. Hammond et al, eds. *Studies in syntactic typology*.

- Amsterdam: John Benjamins, 209-239.
- Gundel, J. K. 1999. "On different kinds of focus." In P. Bosch and R. van der Sandt, eds., *Focus. Linguistic, cognitive and computational perspectives*. Cambridge Univ. Press, 293-305.
- Gundel, J.K., K. Borthen and T. Fretheim. 1999. "The role of context in pronominal reference to higher order entities in English and Norwegian." In P. Bouquet et al., eds., *Modeling and Using Context..Proceedings from the Second International and Interdisciplinary Conference, CONTEXT' 99, Lecture Notes in Artificial Intelligence 1688*. Springer, 475-78.
- Gundel, J. K., N. Hedberg and R. Zacharski. 1993. "Cognitive status and the form of referring expressions in discourse." *Language* 69:274-307.
- Gundel, J. K. and A. Mulkern. 1998. "Quantity implicatures in reference understanding". *Pragmatics and cognition 6: Special issue on the concept of reference in the cognitive sciences*, 21-45.
- Hegarty, M. 2001. "Reference to abstract entities within clausal complements." *Proceedings of the 26th Annual Meeting of the Berkeley Linguistics Society*, 121-132.
- Hegarty, M. Forthcoming. "Semantic types of abstract entities." *Lingua*.
- Hegarty, M., J. K. Gundel and K. Borthen. 2002. "Information structure and the accessibility of clausally introduced referents." *Theoretical Linguistics* 27(2-3).
- Hudson D'Zmura, S. and M. K. Tanenhaus. 1998. "Assigning antecedents to ambiguous pronouns. The role of the center of attention as the default assignment." In M.A. Walker, A. K. Joshi and E. F. Prince, eds., *Centering theory in discourse*. Clarendon Press, Oxford. 199-226.
- Kamio, A-, and M. Thomas. 1999. "Some referential properties of English it and that." In A. Kamio and K.-I. Takami, eds., *Function and Structure*. Amsterdam:John Benjamins.
- Kamp, H. and U. Reyle. 1993. *From discourse to logic*. Dordrecht: Kluwer. Academic Publishers
- Lambrecht, K. 1994. *Information structure and sentence form. topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.
- Levinson, S. C. 1983. *Pragmatics*. Cambridge University Press.
- Sperber, D. and D. Wilson. 1986/95. *Relevance: Communication and cognition*. London: Blackwell. 2nd Edition.
- Vallduví, E. 1990. *The Informational Component*. Doctoral dissertation, University of Pennsylvania, Philadelphia.
- Webber, B L. 1988. "Discourse deixis and discourse processing." Technical report, University of Pennsylvania.
- Webber, B. L. 1991. "Structure and ostension in the interpretation of discourse deixis." *Language and*

Cognitive Processes 6.2:107-135.