

Underspecification of Cognitive Status in Reference Production:
the Grammar-Pragmatics Interface

Jeanette K. Gundel, University of Minnesota
Nancy Hedberg, Simon Fraser University
Ron Zacharski, Mary Washington University

1.0. Introduction. Within the Givenness Hierarchy (GH) framework of Gundel, Hedberg, & Zacharski (1993), referring expressions are assumed to conventionally encode two kinds of information: conceptual information about the speaker’s intended referent and procedural information about the assumed cognitive status of that referent in the mind of the addressee, the latter encoded by various determiners/pronouns. The current work focuses on effects of underspecification of cognitive status, showing that the GH and its predictions, interacting with independently motivated pragmatic/processing factors, makes possible a principled explanation of the distribution and interpretation of different referring forms in spontaneous discourse as well a number of recent experimental results in the psycholinguistic literature.

2. The Givenness Hierarchy. The major premise of the Givenness Hierarchy theory (Gundel et al., 1993) is that different determiners/pronouns encode, as part of their conventional meaning, information assumed by the speaker about the cognitive status of the intended referent in the mind of the addressee. The hierarchy defines an implicational relationship among six cognitive statuses, as shown in (6), along with the forms hypothesized to encode each status in English.

(2) **Givenness Hierarchy (GH)**

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

Forms that encode statuses on the hierarchy thus provide a processing signal that helps guide the addressee in restricting possible referents, by way of procedural information about where and how a mental representation of the referent is to be accessed, as described in (2).

(2)

<i>it</i>	associate representation in focus of attention	(in focus)
<i>this/that/this N</i>	associate representation in working memory	(activated)
<i>that N</i>	associate representation in memory	(familiar)
<i>the N</i>	associate unique representation with NP	(uniquely identifiable)
<i>indefinite this N</i>	associate unique representation	(referential)
<i>a N</i>	associate type representation	(type identifiable)

The rightmost status on the GH, the lowest one, restricts this set the least, and the leftmost status, the highest, is most restrictive. Thus, the indefinite article in English only signals that the addressee is expected to identify the type of thing described. For example, in (3), *a* in *a new scarf* signals that the addressee is to associate an appropriate type representation with the property of being new and a scarf. The pronoun *it* in the second sentence, however, restricts possible referents to those that are currently in the addressee’s focus of attention.

(3) I bought a new scarf. It was on sale.

Thus, the fact that *it* in (3) explicitly encodes the procedural information that the referent is in the addressee's current focus of attention compensates for the lack of conceptual content encoded by this form and thereby facilitates processing by allowing the addressee to associate a fully specified interpretation of the speaker's intended referent as the scarf mentioned prominently in the first sentence.

2.1. The Givenness Hierarchy as an Implicational Scale. The Givenness Hierarchy has often been misunderstood as a scale of degrees of accessibility similar to those proposed by Givón (1983) and Ariel (1990). It is, however, fundamentally different from these other referential hierarchies, both with respect to the role accessibility plays (or does not play) in the theory and in the kinds of assumptions and predictions it makes. Cognitive statuses on the Givenness Hierarchy are part of what Horn (1972) calls an 'implicational scale', a set of items of the same constituent class, ordered in terms of their degree of informativeness, as in (4).

(4) all > most > many > some

Use of a less informative (entailed) form often gives rise to a special kind of pragmatic inference called an 'implicature', following Grice (1975). Thus, someone who hears or reads (5) will typically infer that (6) is not true (even though (6) is logically consistent with (5).

(5) Some first year students came to the orientation.

(6) Not all first year students came to the orientation.

The important point here is that such implicatures arise only in contexts where the information provided by the stronger form (*all* in this case) is relevant. For example, (5) would not necessarily implicate (6) in a context like that in (7).

(7) A: If any first year students came to the party, we'll get reimbursed for the food.

B: Some first year students came to the party.

Such examples show that the quantifier *some* does not conventionally encode the meaning 'not all'; it is simply underspecified for the property 'all'. The Givenness Hierarchy works in the same way. Forms that encode procedural information about statuses on the GH are underspecified for higher (entailing) statuses; they do not exclude them.

As with *some* implicating 'not all', forms lower on the GH are frequently used to implicate that a higher status does not obtain. Thus, pronominal *that*, which explicitly encodes the status 'activated', often implicates that the referent is not in focus. An example is given in (8), where the referent of *that* is taken to be the activated but not in focus closet.

(8) ...going on back from the kitchen then is a little hallway leading to a window, and across from the kitchen is a big walk-through closet. On the other side of that is another little hallway leading to a window. [Personal letter]

Similarly, the English indefinite article *a*, explicitly signals only that the referent is type identifiable, and this in turn typically implicates that the referent is not uniquely identifiable and therefore also not familiar and so on, as shown in (9), where both the first and second occurrence of *a student* implicate that the referent is not uniquely identifiable to the addressee.

- (9) I met a student before class. A student also came by after class.
[Adapted from Hawkins 1991]

However, non-familiarity (or non-unique-identifiability) is not part of the conventional meaning of the indefinite article. For example, the non-familiarity implicature of the second occurrence of *a student* in (9) can be cancelled without contradiction, as in (10).

- (10) I met a student before class. A student also came by after class. In fact, it was the same student I had seen before.

And in cases like (11), the non-familiarity implicature simply does not arise, as specification of the higher status is irrelevant.

- (11) I met with my student, Bill Martin, this morning. Since I already had one meeting with a student today, I'm going home.

Use of some scalar items, such as the definite article *the* in English, rarely, if ever, leads to the implicature that information encoded by an entailing form does not obtain, because information provided by the 'more informative' form is not relevant. Given the explicit information that the referent is uniquely identifiable, along with the encoded conceptual content, information about higher statuses is usually unnecessary for identifying the speaker's intended referent, especially given independent pragmatic principles that favor the interpretation that involves the least processing effort. Phrases with the definite article *the* thus frequently underspecify the cognitive status of the intended referent, as in (12) and (13).

- (12) Mr. Clinton appeared to step on Mr. Bush's dog, Millie, momentarily, then bent down to pet the famous Springer Spaniel.

- (13) I have a dog and a cat. The dog has been with me for ten years.

In some cases, like (12), using a full noun phrase which provides more conceptual information, but underspecifies procedural information, allows the speaker/writer to add additional descriptive information. In others, like (13), it makes it possible to include conceptual content crucial for disambiguating between two potential referents which have the same cognitive status. In both examples, using the determiner *this* or *that*, although licit, would provide more information about cognitive status than necessary. The Givenness Hierarchy thus constrains the forms that can be used to refer to an entity in a given context, but cognitive status is not the sole determinant of what form is used. Actual selection of forms follows from interaction of the Givenness Hierarchy with more general cognitive/pragmatic principles and processes that govern language use (Grice, 1975; Gundel & Mulkern, 1998; Sperber & Wilson, 1986/95; Gundel, 2010, Gundel, Hedberg, and Zacharski, to appear).

2.2. Why the Givenness Hierarchy is not an accessibility/salience scale. As noted in the previous section, each status on the Givenness Hierarchy entails all lower statuses, and forms are simply underspecified for higher statuses; they do not exclude them. Moreover, cognitive statuses on the hierarchy encode procedural information about manner of accessibility (how and where a referent is to be accessed); they do not encode information about degree of accessibility. These features distinguish the Givenness Hierarchy from the Accessibility Hierarchy of Ariel (1990), which directly posits that forms that are higher on the hierarchy are associated with a greater degree of referent accessibility than forms that are lower on the hierarchy. Since the Givenness Hierarchy does not posit that different forms encode (or are otherwise necessarily associated with) different degrees of accessibility, it does not predict that referents attaining higher statuses are necessarily more accessible than referents attaining a lower status.

Example (14), from Gundel, Hedberg, & Zacharski (2001), illustrates a case where the referent of a phrase headed by the definite article might be more accessible, requiring less processing effort, than the same referent when the phrase is headed by a demonstrative determiner.

- (14) ‘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. [Hijacker Leaps to Safety After Robbing Passengers, Oliver Teves, Associated Press, 25 May 2000]

To arrive at the intended interpretation of the second occurrence of *the cockpit* in (14), the reader will have to process the conceptual and procedural meaning encoded in the phrase (which would have to be done in processing any expression). The cockpit is already familiar because it had been mentioned 14 sentences earlier, but it would no longer be activated. There is therefore no reason to think that it is easier for the reader to access the existing representation of the cockpit from memory than to simply create a new unique representation by way of “bridging” to the activated airplane (Clark & Haviland, 1977, Prince, 1981). Although this has yet to be tested experimentally, the new (at most uniquely identifiable) representation is thus arguably more accessible than the (familiar) representation in memory, since less effort would be required to arrive at it. However, the cognitive status that applies—uniquely identifiable—is lower on the hierarchy than the status of the representation in memory—familiar.

3. Explaining experimental results. In recent years, several researchers have argued that forms like demonstrative pronouns (Brown-Schmidt, Byron, & Tanenhaus, 2005; Kaiser & Trueswell, 2008) and indefinite article phrases (Masharov, 2008) do not simply encode lower degrees of salience than personal pronouns and distal demonstrative/definite article phrases, respectively. These authors propose that their results support a form-specific, multiple-factor account of constraints on the use of these forms. However, Gundel, Hedberg, and Zacharski (to appear) argue that the Givenness Hierarchy, because it is an implicational scale of manner of accessibility rather than a hierarchy of degrees of accessibility/salience, is not only consistent with the experimental results reported in these works, but contributes to a principled explanation for them, without requiring formulation of form-specific constraints beyond the cognitive status

information and person/number/gender constraints encoded in the referring expressions.

For example, Masharov (2008) investigated the behavior of English determiners, *a*, *the*, and *that*, through a series of referent selection, scene verification and eye-tracking experiments. In one experiment, participants selected a referent in response to an auditory instruction by clicking on an item on a screen at the same time as their eye movements were recorded. An example input was, “Click on the heart above the lemon. Now click on the broom. Now click on a lemon.”

The overall proportion of selection of a previously selected item as opposed to a previously unmentioned item was highest for ‘*that* N’, then for ‘*the* N’, and lowest for ‘*a* N’. But there was still a strong preference for ‘*a*’ to pick out an already mentioned entity, contrary to what would be predicted by theories that posit that ‘*a*’ is a signal to pick out an unfamiliar referent.

We argue that the GH account is consistent with, and in fact explains Masharov’s results. Given grammatically encoded meanings of the forms together with unidirectional entailment of cognitive statuses, all three forms should be able to refer to previously mentioned entities. Interaction of the grammar with pragmatic principles predicts both the relative preference of ‘*that*’ over ‘*the*’ over ‘*a*’ and the fact that there is still a strong preference for ‘*a*’ in referring to previously mentioned entities, since additional information about cognitive status is not relevant in the experimental context. Both linguistically encoded procedural meaning and extralinguistic cognitive processes at the grammar-pragmatics interface influence production and interpretation of referring forms, but in different ways.

References

- Ariel, M. (1990). *Accessing noun-phrase antecedents*. London: Routledge, Croom Helm.
- Brown-Schmidt, S., Byron, D. K., & Tanenhaus, M. K. (2005). Beyond salience: Interpretation of personal and demonstrative pronouns. *Journal of Memory and Language*, 53, 292-313.
- Clark, H. H., & Haviland, 1977. Comprehension and the given-new contract. In R. Freedle (Ed.), *Discourse Production and Comprehension* (pp 1-40). Norwood, NJ: Ablex.
- Givón, T. (1983). *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam: John Benjamins.
- Grice, H. P. (1975). Logic and conversation. In P. Cole, & J. L. Morgan (Eds.), *Syntax and Semantics 3: Speech Acts* (pp 41-58). New York: Academic Press.
- Gundel, J. K. 2010. Reference and accessibility from a Givenness Hierarchy perspective. *International Review of Pragmatics*, 2, 148-168.
- Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69, 274-307.
- Gundel, J. K., Hedberg, N., & Zacharski, R. To appear. Underspecification of Cognitive Status in Reference Production: Some Empirical Predictions. *Topics in Cognitive Science*.
- Gundel, J. K., & Mulkern A. (1998). Quantity implicatures in reference understanding. *Pragmatics and Cognition*, 6, 21-45.
- Hawkins, J. A. (1991). (In)definite articles: Implicatures and (un)grammaticality predictors. *Journal of Linguistics*, 27, 405-442.
- Horn, L. R. (1972). *On the Semantic Properties of Logical Operators in English* (Doctoral dissertation). University of California (reprinted: Bloomington: Indiana University Linguistics Club, 1976).
- Kaiser, E., & Trueswell, J. C. (2008). Interpreting pronouns and demonstratives in Finnish:

Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes*, 23, 709-748.

Masharov, M. (2008). *Reference Resolution and Discourse Salience* (Unpublished doctoral dissertation). University of Rochester, Rochester, New York.

Prince, E. F. (1981). Toward a taxonomy of given-new information. In P. Cole (Ed.) *Radical Pragmatics* (pp 223-56). New York: Academic Press.

Sperber, D., & Wilson, D. (1986/95). *Relevance: Communication and Cognition*. Oxford: Wiley-Blackwell.