Applying the Givenness Hierarchy Framework: Methodological Issues

Nancy Hedberg
hedberg@sfu.ca
Simon Fraser University
Burnaby, British Columbia, Canada

1. Introduction.

This paper discusses methodological issues that go into determining what's called the 'information structure' of utterances. In her work (e.g. Gundel 1988, Gundel and Fretheim 2004), Jeanette Gundel proposes that two logically distinct notions of information structure need to be distinguished: 'relational givenness/newness' and 'referential giveness/newness.'

On the one hand, 'relational givenness/newness' describes two complementary parts of a single utterance, where one part can be viewed as 'given' information in relation to the second part, which expresses 'new' information in relation to the first part. In example (1), the question in (1a) established the entity 'John' as given in relation to the predicate 'left' in (1aB), whereas the opposite relation holds in (1bB).

(1) a. A: What did John do?
B: John LEFT.
b. A: Who left?
B: JOHN left.

Gundel and Fretheim 2004 state that this sense of givenness/newness "reflects how the informational content of an event or state of affairs expressed by a sentence is represented and how its truth value is to be assessed." In (1a), the focal information that he left is assessed relative to the topical entity John. While in (1b), the information that it was John is assessed relative to the predicate 'left.'

Terms for this distinction in the literature include complementary notions such as 'presupposition' vs. 'focus' (Chomsky 1971, Jackendoff 1972), 'topic' vs. 'comment' or 'focus' (Gundel 1974, Reinhart 1982, Lambrecht 1994), and 'theme' vs. 'rheme' (Vallduvi 1992). In a recent attempt to define information structure notions in formal semantics, Féry and Krifka 2008 define 'topic' as a constituent that "identifies the entity or set of entities under which the information expressed in the comment constituent should be stored in the CG content," where 'CG' refers to the common ground, the set of propositions and entities viewed as already shared between speaker and addressee.

Note that the capitalization in (1) indicates that the part of the sentence that encodes the relational focus is expressed with primary prosodic stress in English. This differs between (1a) and (1b). Gundel and Fretheim 2004 note that topics can also be stressed in English, e.g. when they are contrastive, although possibly by means of a different pitch accent (fall-rise vs. fall). In Japanese, 'John' would be expressed by the topic marker (*wa*) in (1a), but by the subject marker (*ga*) in (1b).

On the other hand, 'referential givenness/newness' describes a relation between the intended referent of a linguistic expression—typically a nominal expression—and its informational status in the memory/attention states in the hearer's mind. In this sense, a discourse

referent can be said to be 'salient', 'activated', 'familiar', 'specific, 'brand new', etc., as described in Prince 1981, Ariel 1990, Gundel, Hedberg and Zacharski 1993, Chafe 1994, inter alia. Féry and Krifka 2008 seem to be getting at this notion when they define 'givenness' as follows: "A feature X of an expression α is a Givenness feature iff X indicates whether the denotation of α is present in the CG or not, and/or indicates the degree to which it is present in the immediate CG."

I will be focusing on the theory of referential givenness developed in Gundel, Hedberg, and Zacharski 1993 and subsequent work. In that work, we explain how the form by which an entity is referred to is correlated with the cognitive, i.e. memory and attention, status of that entity for the addressee, as assumed by the speaker. My aim is to discuss methodological issues that go into determining which 'cognitive status' a particular referring expression realized in some language has for the addressee in a particular context, as well as into determining how the system of referential expressions in a particular language relates to the Givenness Hierarchy. I will primarily give examples from English, Japanese and Mandarin, with the hope that readers of these proceedings can use this discussion to help them explore the referring expression systems of the Austronesian languages that they are studying.

2. The Givenness Hierarchy.

The Givenness Hierarchy of Gundel, et al. 1993 is a set of six 'cognitive statuses' (memory and attention states) in the mind of the addressee (as assumed by the speaker). These statuses are claimed to constitute meanings of pronominal and determiner forms, and determine necessary and sufficient conditions on the use of each referring form in discourse. The Givenness Hierarchy and the English forms that are claimed to be associated with the different statuses are shown in (2), and the meanings are defined in (3).

(2) The Givenness Hierarchy (with English forms used for illustration):

in focus >	**activated** >	**familiar** >	**uniquely identifiable** >	**referential** >	**type identifiable**
it[1]	*IT*/*this/that/ this* NP[2]	*that* NP	*the* NP	indefinite *this* NP	*a* NP

(3)

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

[1] 'It' here stands for all unstressed personal pronouns. Capitalized personal pronouns under 'activated' stand for all stressed personal pronouns.

[2] The "DP hypothesis" is assumed here, where by nominal phrases are headed by a determiner which surfaces as a bare determiner in "pronominal" uses and which takes an NP (or Classifier Phrase) as complement when occurring adnominally.

For example, utterance of (4a) would be felicitous only if it were reasonable for the speaker to assume the addressee already had his/her attention focused on the referent in question; while (4b) is felicitous if the addressee can associate a memory representation with the dog, even if the dog has not been mentioned in the current discourse. On the other hand, (4c) could be used to introduce the dog to the addressee for the first time because a unique representation can be associated with the DP due to its conceptual content. And (4d) can be used as long as the addressee knows what a dog is.

(4) a. I couldn't sleep last night. **It** kept me awake.
b. I couldn't sleep last night. **That dog next door** kept me awake.
c. I couldn't sleep last night. **The dog next door** kept me awake.
d. I couldn't sleep last night. **A dog** kept me awake.

The cognitive statuses are defined in such a way that they stand in a unidirectional entailment relation and thus form a hierarchy. Any DP referent that is in the addressee's focus of attention is also represented in working memory, is represented in memory generally, can be associated with a unique token representation expressed by the DP, can be associated with a unique token representation in general, and can be associated with a representation of a type of entity. However, a referent can be familiar but not activated, for example, or referential but not uniquely identifiable, because the entailment relation only goes in one direction.

The status with which a form is associated is a necessary condition for felicitous use of that form, but because higher statuses entail lower statuses, when a higher status obtains, the necessary conditions for all lower forms also hold. A lower form can thus be chosen if specific information about a higher status obtaining does not need to be expressed to ensure that reference succeeds. For example, in (5), the particular goldfish is in the focus of attention and thus could be referred to by a personal pronoun, but a definite article is chosen by the speaker instead. Other forms that explicitly signal a status lower than in focus (*this fish*, *that fish*) would also have been felicitous.[3]

(5) The man wins this time, and the fish that he selects is a big goldfish, which is, at the point when he selects it, hidden in a rocky formation in the tank, and it's impossible for the man conducting the game to get at **the fish** with the net.
{it, this fish, that fish}

However, if a higher status does not obtain, then a higher form cannot be chosen. Thus, in (6) the water in the bowl is uniquely identifiable via associative inference from the information that the fish is swimming around in a large glass bowl, but the water has not been explicitly introduced and hence is not familiar to the addressee (and hence not activated or in focus either).

(6) The fish is swimming around in a large glass bowl on the table right next to the birdcage. And the scene jumps back and forth between the bird, the fish and the cat, who's outside roaming around the streets. The bird and fish seem to be playing, turning themselves

[3] The data here was originally collected for a study by Fuller and Gundel 1987. Speakers viewed a silent film called "The Golden Fish" and, immediately after viewing the film, described it to another native speaker of their language. In this paper, I use examples from the goldfish stories told in English, Japanese and Mandarin.

upside down and doing almost a kind of dance, the bird spinning around on his perch and the fish swimming upside down and jumping out of **the water**.
{#it, #this water, #that water}

Gundel et al. 1993 associate forms in several languages other than English with the Givenness Hierarchy: Mandarin, Japanese, Spanish and Russian. Gundel, Bassene, Gordon, Humnick and Khalfaoui 2010 examine four additional languages: Eegimaa (a Niger-Congo language), Kumyk (a Turkic language), Ojibwe (an Algonquian language), and Tunisian Arabic. Other languages have also been explored in this framework, e.g. Irish (Mulkern 2007), Norwegian (Borthen 2003), and Persian and Turkish (Hedberg, Görgülü and Mameni 2009).

In order to form the basis for discussion in the present paper, the pairing of cognitive statuses on the Givenness hierarchy with linguistic (i.e. various determiners and pronouns) postulated in Gundel et al. 1993 for Mandarin and Japanese are given in (7) and (8).

(7) *Mandarin*.

in focus >	**activated >**	**familiar >**	**uniquely identifiable >**	**referential >**	**type identifiable**
Ø *ta* 'he/ she/it'	*TA* *zhe* 'PROXIMAL' *nei* 'DISTAL' *zhei* NP		*nei* NP		*yi* NP 'a NP' Ø NP

(8) *Japanese*

in focus >	**activated >**	**familiar >**	**uniquely identifiable >**	**referential >**	**type identifiable**
Ø	*kare* 'he' *kore* 'PROXIMAL' *sore* 'MEDIAL' *ano* 'DISTAL' *kono* NP *sono* NP	*ano* NP			Ø NP

In arriving at the pairings of forms in a language with the Givenness Hierarchy, both corpus analysis and native speaker intuitions are utilized. Corpus examples have been used, for example to establish that Mandarin and Japanese allow bare nominals, indicated as 'Ø NP' in (7) and (8), for both definite and indefinite DPs.

The Mandarin examples in (9) show that bare NPs can be definite or indefinite, which correlates with their status on the hierarchy as 'type identifiable,' a status entailed by all higher statuses and thus the only necessary condition on the use of a bare NP. In (9a), the reference to 'wine' is indefinite, and is at most type-identifiable in the Givenness Hierarchy framework because no particular wine is intended. This meaning is expressed as a bare NP. Reference to 'rock' in (9b) is also made with a bare NP and is classified as at most referential since by the time the whole sentence is processed, a unique representation can be associated with 'rock that

the goldfish hid behind'. Finally, definite reference to the goldfish in the second sentence of (9b) is also made with a bare NP, and this would be classified as activated, and possibly in focus, as it was mentioned in the previous sentence. The use of bare NPs to refer to entities with statuses higher than type identifiable is consistent with the fact that all higher statuses also satisfy the condition of being type identifiable, as lower statuses are entailed by all higher ones.

(9) a. shang jie qu mai **jiu**.
mount street go buy wine
'[He] went out to buy *some wine.'*

b. ta feichang xiang dedao yi-zhi hong-de jinyu
he very want get one-CLS red-PRT goldfish
'He very much wanted to get a red goldfish.'

danshi **hong-de jinyu** cang zai **shitou** houmian
but red-PRT goldfish hide in rock behind

'But *the red goldfish* was hiding behind a rock.'

In the Japanese examples in (10), the black cat is first introduced with a bare NP in (10a) that would be classified as at most referential, since the addressee can associate a unique representation after the whole sentence is processed (indefinite reference). In (10b), the already activated goldfish and bird (definite references) are also expressed as bare nominals.

(10) a. ichiban hajime ni detekita bamen ga
first beginning at appeared scene SM

ano ee mannaka hen ni **kuroi neko** ga ugoiteite
well eh center about at black cat SM be.moving

'The scene that first appeared has *a black cat* moving in the center'.

b. **kingyo** to **kotori** ga uta o utattari
goldfish and bird SM song OM sing

sorekara **kingyo** ga kurukuru mawattari
and.then goldfish SM round.and.round turn

'*The goldfish* and *the bird* sing songs. And *the goldfish* turns round and round.'

For some distinctions, interviews with native speakers have been relied upon in order to elicit judgments about critical examples. For example, the Mandarin sentence (11) is reported by consultants to be felicitous in the context where the addressee is not familiar with the dog in question.

(11) Zuotian wanshang wo shui bu zhao. **Gebi de nei tiao gou** jiao de lihai.
yesterday evening I sleep not achieve next.door PRT that CLS dog bark PRT terribly
'I couldn't sleep last night. *The dog next door* kept me awake.'

This judgment motivates placing the distal demonstrative *nei* in Mandarin under the 'uniquely identifiable' position on the hierarchy, the position for definite articles in English. Such a placement supports the claim by Li and Thompson 1981 (among others) that the distal demonstrative determiner in Mandarin is in the process of grammaticalizing into a definite article.

A comparable elicited example of a distal demonstrative in Japanese (which has a three-way deictic distance distinction) supports the idea that the Japanese distal demonstrative determiner *ano* is not grammaticalizing into a definite article. Thus (12) was reported by consultants to only allow a familiar reading of the demonstrative expression (i.e. one where the addressee is expected to already know that the speaker's neighbor has a dog):

(12) Kinoo wa hitobanjuu nemurenakatta. **Tonari no ie no ano inu** no sei
yesterday TOP all.night could.not.sleep. Neighbor GM house GM that dog GM reason
da.
is
'I couldn't sleep last night. *That dog next door* is the reason.'

Finally, Gundel et al. 2010 (p. 1780) used another method of eliciting cognitive status judgments. They presented two alternatives of an extended, multi-sentence Ojibwe text designed to determine if a full noun phrase or only a zero pronoun would be required for a certain referent. The English translation of the text is shown in (13) Through consultation with native speakers, it was determined that a full noun phrase is needed in this context in Ojibwe.

(13) The woman saw a dog yesterday, but she had to work, so she ignored it. In fact, she had to work the whole day, and again all day the day after. Right then was when she heard that **{the dog/Ø}** had been run over.'

3. The Coding Protocol.

In order to specify more precisely exactly what cognitive status a given referent in a natural language text exhibits, a coding protocol has been developed by the originators of the Givenness Hierarchy Framework (Gundel et al. 2006). This protocol specifies guidelines for determining cognitive status. These guidelines are not definitional but do help establish cognitive status. Most of the examples here will be given in English, Japanese and Mandarin. The criteria are to be checked in higher- to lower-status order so that the highest cognitive status that a referent exhibits can be determined as soon as a criterion is met.

3.1. In Focus. The Coding Protocol gives three criteria for establishing that a referent is 'in focus.' (i) it is the interpretation of the main clause subject or the syntactic topic in the immediately preceding sentence. (ii) It was established by a preceding reference in the same sentence, by a referent focused in a preceding existential or cleft sentence, or by the event expressed in the previous sentence. (iii) It can be associated with a higher-level topic that is part of the interpretation of the preceding sentence (whether it is overtly mentioned there or not).

The subject criterion is illustrated in (14), where subsequent references to the man, who was introduced in subject position, are all in focus.

(14) And a man came to play the game and **Ø** was also eyeing that strikingly different goldfish. And then **he** won and **he** tried to get the fish, but they couldn't get the fish out of the little rock formation.

An example from Japanese in (15) of the topic condition shows an in-focus zero pronoun referring back to the referent of a *wa*-marked syntactic topic phrase that is embedded in an existential complex nominal subject.

(15) de otokonoko wa yorokonde uchi e modotteiki iku tte iu bamen ga ate
and boy TM happily home to go.back go QU say scene SM be
'and there's a scene where the boy goes home happily.'

sorede uunto **Ø** sono kingyo o ookii garasu no hachi no naka ni irete
then well that goldfish OM big glass GM fishbowl GM inside in put.and
'then, well, [*he*] puts that goldfish into a big glass fishbowl.'

(16a) shows an example of an unstressed personal pronoun whose referent was mentioned earlier in the same sentence. (16b) shows an antecedent in existential focus, and (16c) shows an antecedent in cleft focus. In (16d), the unstressed personal pronoun finds its referent in the event expressed by the preceding sentence.

(16) a. You can wear my scarf if you can find **it**.
b. There was a mouse on the table. **It** was very large.
c. It was the dog that Bill was afraid of. **He** was very large.
d. John fell off his bike. **It** happened yesterday.

Finally, in-focus status can be associated with "a higher-level topic", as in (17) from Japanese, where the boy, which is a higher-level discourse topic mentioned as subject of the first sentence is referred to again with a zero pronoun as the subject of the third sentence after an intervening sentence that didn't mention him overtly but perhaps evokes him covertly.

(17) de Ø ato hanbun o nanka jibun no kotori ni
and remaining half OM something self GM bird to

yatte ita mitai desu ne
was giving seem be PRT

'And it seems that (he) was giving his bird the remaining half.'

sorede soo shitara nanka sono kotori no kago ga oite aru
then and then something that bird GM cage SM put be

yoko no tana ni eeto konna isshoo bin ka nanka
side GM shelf on well like.this one bottle QP something

to sorekara kooiu kwootaa to sorekara hitotsu penny ga atte
and then like.this quarter and then one penny SM be

'Then on the side of the place where the cage is set, there is a shelf on which there are a bottle or something and a quarter and one penny."

Ø sono ishoo bin to okane o motte
that one bottle and money OM take

soto e dete iku ee
outside to take out well

'(**He)** takes up the bottle and money and goes out.'

3.2. Activated. The Coding Protocol gives three conditions under which a referent can be coded as activated: (i) "It is part of the interpretation of one of the immediately preceding two sentences." (ii) "It is something in the immediate spatio-temporal context that is activated by means of a simultaneous gesture or eye gaze." (iii) "It is a proposition, fact, or speech act associated with the eventuality (event or state) denoted by the immediately preceding sentence(s)."

The first condition can be illustrated by the use of 'sono ishoo bin to okane o' [that bottle and money] in the last line of the Japanese example in (17)—those two referents had been introduced in the immediately preceding sentence. The second condition is illustrated with the example in (18), and the third condition is illustrated by the examples in (19) and (20):

(18) [Looking at the wrench] Please hand me **that (wrench** (over there)))

(19) A: John fell off his bike.
B: **That**'s not true.

(20) A: John fell off his bike.
B: Can you say **that** again?

The use of activated pronominal forms can be analyzed as sometimes implicating that in-focus status does not obtain, as in (21).

(21) Anyway, 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 that other side of **that** is another little hallway leading to a window…. [Personal letter, Gundel et al. 1993, ex. 56, p. 298].

Use of (stressed) 'that' here indicates that an activated but not in-focus referent (the closet) was intended here instead of the in-focus potential referent (the kitchen). Gundel et al. 1993 propose that this inference is a conversational implicature, which is generated by the fact that the statuses on the Givenness Hierarchy stand in a unidirectional entailment relation, which means that they constitute a "Horn Scale" (Horn 1972). The Givenness Hierarchy can thus be expected to support quantity implicatures that are based on the Maxims of Cooperative Conversation of Grice 1975. The hearer can be understood as reasoning that if the higher cognitive status of in-focus had been

intended, the speaker would have used a higher form ('it') so as to be as informative about cognitive status as possible, as dictated by the first part of the Quantity Maxim ("Make your contribution as informative as required"). Because the speaker did not use that expression, she must have intended that it would be infelicitous to do so. Hence the form selected refers to the merely activated entity.

Another example of a possible focus-shift implicature arises in the Japanese goldfish story, In (22) a personal pronoun refers to the boy. Then there are three clauses with no reference to the boy, but in the fifth clause he is referred to again with the personal pronoun 'kare' and the topic marker 'wa'.

(22) de kare wa soko e sono omise no took e itte iku
and he TM there to that shop GM place to go go
'He goes to that shop.'

to soosuruto nanka koo suisoo no naka ni takusan no kingyo ga ite
and then something er fishbowl GM inside in many GM goldfish SM be
'And then there are a lot of goldfish in a fishbowl.'

sorede sono naka ni eeto hotondo ga kuroi kingyo nanda
then that among PREP well almost SM black goldfish be
'And then there are a lot of goldfish.'

kedo sono naka ni ippiki dake akai kirei kingyo ga ite
but that inside in one only red pretty goldfish SM be
'But among the goldfish there is only one pretty red goldfish.'

de **kare** wa ano sono akai kingyo o hoshiii to omou n desu ne
and he TM well that red goldfish OM want QU think PRT be EM
'And *he* wants that red goldfish.'

The form 'kare' is associated with the status 'activated'. Here again we can postulate that the boy is a higher-level discourse topic that is covertly evoked and hence activated in this passage. On this analysis, the form 'kare' is used to refer to the boy to indicate a focus shift. A zero pronoun might have misleadingly been interpreted as the in-focus red goldfish. This analysis of the boy as a discourse topic is supported by the fact that reference to the boy was explicitly marked by the topic marker 'wa' in the first clause in (22).

Languages typically have several forms that encode activated status, including all demonstrative pronouns. The Givenness Hierarchy does not distinguish between these different forms, although Gundel et al. 1993 did propose a parameter of 'speaker activation' to account for the distinction between proximal and distal determiners. Often proximal determiners are only used to refer to entities in "the speaker's context space", but this constraint is not universal. Thus, Mandarin 'zhe' apparently does not obey the constraint, as the elicited example in (23) shows:

(23) A: Wo juede wode xiaoshuo bi luxun xie-de hao.
I think my novel than Luxun write-PRT good

'I think my novels are better than Luxun's.'

B: **Zhei-ge/?nei-ge** wo chengren.
this-CLS/that-CLS I admit.
'I agree with *?this/that*.'

3.3. Familiar. The Coding Protocol gives two conditions under which a referent can be said to be familiar: (i) "it was mentioned at any time previously in the discourse"; and (ii), "it can be assumed to be known by the hearer through cultural/encyclopedic knowledge or shared personal experience with the speaker."

The first condition can be illustrated with the reference to the boy (the student) in Mandarin where he had not been mentioned for 34 clauses, but yet the proximal demonstrative determiner is used to refer to him in (24):

(24) suoyi qiaqiao zhei-ge shihou **zhei-ge** **xuesheng** huidao-le jia.
so by.chance this-CLS time this-CLS student return-ASP home
'So (it) happened (just at) this time that *the student* returned home.'

Such usage perhaps motivates placing the proximal demonstrative determiner in Mandarin under the 'familiar' category for Mandarin instead of under the category of 'activated' as was postulated for Mandarin in Gundel et al. 1993. However, perhaps the usage in (26) should be attributed to the fact that the student is an ongoing discourse topic and thus is globally activated in the context of the story. The Coding Protocol, however, would dictate coding the proximal phrase in (26) as 'familiar' because the boy had not been mentioned in the preceding two sentences, and this was how it was coded in Gundel et al. 1993.

At the same narrative point in the Japanese goldfish story as the Mandarin demonstrative reference to the boy (the student) in (24), however, a bare NP was used, as shown in (25), to refer to the boy, who had not been mentioned for 32 clauses; and the bare NP was also coded as 'familiar'.

(25) de sono ato de kondo wa **otokonoko** ga shoogakko ga owatte
and that after PREP this.time TM boy SM elem.school SM end

kaettekuru tochuu no bamen ga detekite.
come.back way GM scene SM appear

'After that, this time a scene appears where *the boy* comes back after school is over.'

The example from the Coding Protocol of the first subcase of the second criterion for familiarity, whereby the referent is known based on cultural/encyclopedia knowledge is shown in (26). Here, the 90's had not been mentioned in the discourse.

(26) 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**.

Examples of the second subcase of the second criterion for familiarity, whereby a referent can be assumed to be known to the hearer through shared personal experience with the speaker may be illustrated by examples that Himmelmann 1996 refers to as "recognitional uses" of demonstrative systems. He proposes that such usage is a universal feature of deictic systems where "the intended referent is identified through specific knowledge shared by the speaker and addressee, rather than through context." We refer to such uses as 'reminder' uses.[4]

Jarbou 2010 reports that the utterances in (27) and (28) exemplify recognitional deixis in Spoken Jordanian Arabic. Note that the proximal demonstrative determiner apparently has this function in SJA although the proximal demonstrative determiner in English is restricted to activated uses.

(27) btiðəkər **hazaak elfonduq elli nzilnə fee əwəl marə?**
remember-you.2SG.M that.SG.M hotel which stayed-we in first time?
'Do you remember *that hotel where we stayed in our first visit*?'

(28) **hatha elmakan elli rohna 3alai yoom elJum3a** 3njad kan hilu
this-SG.M place which went-we to day Friday really was beautiful
'*This place where we went on Friday* was really beautiful.'

It can be concluded from data reported in Gundel et al. 1993 and Jarbou 2010 that languages differ as to which demonstrative in a deictic paradigm allows for recognitional deixis, uses of which would be coded as familiar but not activated on the Givenness Hierarchy. Spoken Jordanian Arabic and Russian allow the proximal demonstrative in a two-way deictic system to do this, but English and Mandarin do not. As for three-way deictic systems, Spanish allows the medial and distal demonstrative to encode a referent that is familiar but not activated, while Japanese allows only the distal demonstrative to encode the recognitional function.

3.4. Uniquely Identifiable. The Coding Protocol gives two conditions for coding a referent as uniquely identifiable: (i) "the referent form contains adequate descriptive/conceptual content to create a unique referent," and (ii) "a unique referent can be created via a 'bridging inference' by associating with an already activated referent." Both conditions are illustrated in (29), where (29a) is repeated from (4c).

(29) a. I couldn't sleep last night. **The dog next door** kept me awake.
b. I went to a wedding yesterday. **The bride/#that bride/#this bride** wore pink.

In (29a) the addressee does not need to be previously familiar with the dog in question, but rather can construct a unique representation from the linguistic material given in the DP. A unique representation can also be constructed for associative anaphora, for uses called 'inferrable' by Prince 1981. (29b) shows that higher forms are not possible with such uses. Here a unique representation of the bride can be inferred by a 'bridging inference' from the activated wedding referent. Definite articles, in languages that have them, like English, can thus be used to encode referents that are uniquely identifiable but not familiar.

[4] This term was first used in Gundel et al. 1988, e.g. in Table 2, p. 219.

Another example of condition (ii), from Japanese, is given in (30). Here a representation of the remaining half of the crumb can be inferred from mention of the first half.

(30) nanka pan ka nanka no kakera o eeto hanbun kurai tabete.
something bread QP something GM crumb OM well half about eat
'He eats about half of a crumb of bread.'

de **ato** **hanbun** o nanka jibun no kotori ni yatte ita mitai desu ne
and remaining half OM something self GM bird to was giving seem be EM
'And it seems that he was giving his bird *the remaining half*.'

It is often reported, e.g. recently by Diessel 2012, that definite articles frequently historically develop from demonstratives. In some languages, such as Mandarin, the distal demonstrative determiner seems to be taking on some properties of definite articles in that the distal form can be used in some instances for non-familiar referents, as in example (11) above.

3.5. Referential. The Coding Protocol states that a referent exists, is referential, if it meets one of two criteria: (i) "it is mentioned subsequently in the discourse", and (ii) "it is evident from the context that the speaker intends to refer to some specific entity." The examples given are shown in (31). (31b) illustrates a use of "indefinite *this,*" which is used to indicate referential entities in casual, spoken English.

(31) a. When my youngest child was three or so, we were at **a friend**'s house visiting and my friend was babysitting her infant nephew.
b. I want to tell you about **this strange guy I saw today**.

Mandarin quite often seems to use the numeral 'one' for referential uses, as in (32), where the main character of the boy is first introduced. By contrast, in the Japanese story, (33) shows a bare nominal used to first introduce the boy.

(32) zhei-ge shihou you **yi-ge** **haizi** qu dushu.
this-CLS time exist one-CLS child go study
'At this time, there was *a child* going to school.'

(33) de sono naka de hitori mu **chuugokujin ga nanka** **no otokonoko** ga
and that among PREP person um Chinese SM something GM boy SM

eeto ooutsushi ni natte
well close.up into brought

'And among them *one Chinese boy* is brought into close up.'

This usage in Mandarin, which is widespread in the story analyzed, perhaps indicates that the number 'one' in Mandarin is in the process of grammaticalizing into an indefinite article, as was proposed in Li and Thompson 1981.

The Salish languages of the Pacific Northwest of North America contain a system of frequently used articles that indicate referentiality, also called 'specificity' (Matthewson 1998, Gillon 2006). The referential article in the Salish language Sechelt (Sháshíshálh) is illustrated in (34), where the data is drawn from a story, 'The Beaver', published in Beaumont 1985[5]. A primary character in the story is the snake woman, introduced in (34a) by means of the feminine form of the referential article. Here the DP is referential but not uniquely identifiable. Later in the story, as in (34b), when the snake woman is referred to again several times, the referential article is again used, here for a familiar or activated referent. Since these articles explicitly encode the cognitive status 'referential' in Salish languages, it is not surprising that they can be used in encoding entities that have a status higher than referential as well as ones that are referential, but not uniquely identifiable.

(34) a. tí súxw-t-as **ɬe ʔulqay slánay**
AUX see-TR-3ERG ART snake woman
'He saw a snake woman.'
…

b. tí λum s-qwál-s **ɬe slánay** …
AUX then NOM-speak-3SG.POSS ART woman
'Then the woman said…'

In contrast to the Salish case, use of indefinite articles in English and other languages of often implicates that a higher cognitive status does *not* obtain. For example, such an implicature is cancelled in (35):

(35) I met a student before class**. A student** came to see me after class as well--in fact it was the same student I had seen before. [Hawkins 1991]

More research on different languages is needed before an explanation can be given of exactly when a given form will implicate that a higher status does not obtain.

3.6. Type Identifiable. Finally, the Coding Protocol states that, "an interpretation is type identifiable if the sense of the phrase (the descriptive/conceptual content it encodes) is understandable." The two examples shown in (36) are given:

(36) a. I don't have **a VCR** and neither does my neighbor.
b. Whenever Mary passes that store, she always picks up **a newspaper**.

It is clear that no unique referent exists when the indefinite phrase is in the scope of negation or a quantifier.

An example from Mandarin in (37), repeated from (9a), shows use of a bare NP for a non-referential, type-identifiable entity:

[5] Examples from this story given here were re-glossed by Kaoru Kiyosawa, PhD. student at the time at Simon Fraser University.

(37) shang jie qu mai **jiu**.
mount street go buy wine
'[He] went out to buy *some wine*.'

At this point in the story the speaker is not intending to refer to any particular quantity of wine. Hence, this is a type-identifiable, non-referential use of the bare NP.

Because all cognitive statuses entail 'type identifiable', indefinite article phrases can also have higher cognitive statuses. Thus, the indefinite article phrase in (38) is compatible with an interpretation where the speaker expects the addressee to be familiar with the car in question, e.g. because it is the speaker's own car.

(38) I'm ready to get some exercise. I've been sitting in **a car** all day.

4. Referring Expression Usage in Languages with and without Definite and Indefinite Articles.

Hedberg 1996 compares one goldfish story in each of English, Japanese and Mandarin in order to get an idea about how different referring expression types are distributed in a text. I argued that the stories indicate that the speakers of Japanese and Mandarin used different strategies in referring to entities that were marked with a definite article in English.

In (39)-(41), I show the description of the same episode in each of the three languages. While the English speaker used a definite article to refer to the activated main characters, all of which had been introduced already and the Japanese speaker used bare NPs, the Mandarin speaker frequently uses demonstrative phrases.

(39) *English*

All the while **the cat** now, which had gotten in,
didn't seem to notice **the fish**,
which was laying out of **the bowl**,
but was trying to get through, **the cat**, at **the bird**.

(40) *Japanese*

de sorekara **neko** ga haittekitte
And then cat SM come.in
'And then the cat comes in.'

sorede hajime ni **kingyo** ni kigatsuku ka to omottara
and.then stare at goldfish at notice QP QU thought
'At first (I) thought (it) would notice the goldfish.'

saki ni **kotori** no hoo o mite
first at bird GM direction OM look
'But (it) first looked at the bird.'

(41) *Mandarin*

suoyi **hei mao** jinlai de shihou
so black cat enter PRT time
'So when (the) cat, (the) black cat came in,'

zhei-zhi jinyu qiaqiao shi zai **zhoumian** shang
this-CLS goldfish by.chance be on table.top on
'this goldfish happened to be on the table.'

keshi **hei mao** meiyou faxian **zhei-zhi jinyu**
but black cat neg past notice this-CLS goldfish
'but the black cat didn't notice the goldfish,

yushi jiu xiang chi **nei-ge niaor.**
since then think eat that-CLS bird
'Since he was thinking of eating that bird..'

Table 1 shows the distribution of full definite DP forms across the three stories. It can be seen that demonstrative determiners are more frequent in Mandarin than in English and Japanese.

	Demonstrative determiner + NP	Definite determiner + NP	Definite bare NP	Total definite full DP	Total referring expressions
English	7 (5%)	58 (44%)	n.a.	65 (49%)	133 (100%)
Japanese	27 (14%)	n.a.	78 (40%)	105 (54%)	193 (100%)
Mandarin	37 (24%)	n.a.	45 (28%)	82 (52%)	158 (100%)

Table 1. The distribution of full definite DPs according to form of determiner.

When eight uses of *sono* in Japanese are eliminated because they occur in a use special to that language, the percentage of demonstrative determiners goes down to 10% in Japanese. This special use is illustrated in (42). Here, the medial demonstrative determiner *sono* is used to mark a discourse-oriented relational noun.

(42) sorede **sono naka** ni eeto hotondo ga kuroi kingyo nanda
then that inside at well almost.all SM black goldfish be
'Among them almost all are black goldfish.'

About half the demonstratives in the Mandarin story (most proximal) marked post-verbal important activated entities. Perhaps such referents occur in post-verbal position in Mandarin because they constitute part of the speaker's comment (i.e. they are relationally new, although referentially given). Because post-verbal bare NPs are interpreted as indefinite in Mandarin (Li & Thompson 1976), it seems plausible to propose that an explicit marker of activation is used in this position to block an implicature of non-activation. Japanese doesn't have such a word-order restriction on the use of bare NPs, so bare NPs are used even for activated referents that are part of the speaker's comment.

5. Conclusion

This paper has reviewed the theory of cognitive status of Gundel, Hedberg and Zacharski 1993, which is a theory of referential givenness. Examples were given to illustrate how to assign cognitive status to referents in accordance with the Coding Protocol of Gundel et al. 2006, with illustrations from Japanese and Mandarin as well as English. Finally, a brief example was given to show how the cognitive status constraints on the use of referring forms in a language interacts with the inventory of referring forms and with constraints on the realization of relational givenness to determine the distribution of the various referring forms in actual texts.

6. References

Ariel, Mira. 1990. Accessing Noun-Phrase Antecedents. London: Routledge.

Beaumont, Ronald C. 1985. She shashishalhem, the Sechelt language: Language, stories and sayings of the Sechelt Indian people of British Columbia. Penticton, British Columbia: Theytus Books.

Borthen, Kaja. 2003. Norwegian Bare Singulars. Doctoral Dissertation, Norwegian University of Science and Technology.

Chafe, Wallace. 1994. Discourse, Consciousness and Time. Chicago University of Chicago Press.

Chomsky, Noam. 1971. Deep structure, surface structure, and semantic interpretation. In Steinberg, Danny D. and Leon A. Jakobovits (eds.). Semantics: An Interdisciplinary Reader in Philosophy,Linguistics and Psychology. Cambridge: Cambridge University Press.180-216.

Diessel, Holger. 2012. Deixis and demonstratives." In Maienborn, Claudia, Klaus von Heusinger, and Paul Portner (eds.). Semantics: An International Handbook of Natural Language Meaning. Volume 3. Berlin/Boston: De Gruyter. 2407-2432.

Féry, Caroline and Manfred Krifka. 2008. Information structure: Notional distinctions, ways of expression. In van Sterkenburg, P. (ed.). Unity and Diversity of Languages, Amsterdam: John Benjamins, 123-136.

Fuller, Judith and Jeanette K. Gundel. 1987. Topic prominence in inter-language. Language Learning 37. 1-17.

Gillon, Carrie. 2006. The Semantics of Determiners: Domain Restriction in Skwxwu7mesh. Doctoral Dissertation, University of British Columbia.

Grice, H. P. 1975. Logic and conversation. In Cole, Peter, and Jerry L. Morgan (eds.), Syntax and Semantics, Volume 3: Speech Acts. New York: Academic Press. 41-58.

Gundel, Jeanette K. 1974. The Role of Topic and Comment in Linguistic Theory. Doctoral Dissertation, University of Texas, Austin.

Gundel, Jeanette K. 1988. Universals of topic-comment structure. In Hammond, M., E.

Moravscik and J. Wirth (eds.), Studies in Syntactic Typology. Amsterdam: John Benjamins. 209-239.

Gundel, Jeanette K., Mamadou Bassene, Bryan Gordon, Linda Humnik, and Amel Khalfaoui. 2010. Testing predictions of the Givenness Hierarchy framework: A crosslinguistic investigation. Journal of Pragmatics 42. 1770-1785.

Gundel, Jeanette K. and Thorstein Fretheim. 2004. Topic and focus. In Horn, Laurence R. and Gregory Ward (eds.). The Handbook of Pragmatics. Oxford: Blackwell. 175-196.

Gundel, Jeanette K., Nancy Hedberg, and Ron Zacharski. 1988. "On the generation and interpretation of demonstrative expressions," in D. Vargha, ed., Proceedings of the Twelfth International Conference on Computational Linguistics, John von Neumann Society for Computing Sciences: Budapest, Hungary, 216-221.

Gundel, Jeanette K., Nancy Hedberg, and Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. Language 69(2). 274-307.

Gundel, Jeanette K., Nancy Hedberg, Ron Zacharski, Ann Mulkern, Tonya Custis, Bonnie Swierzbin, Amel Khalfoui, Linda Humnick, Bryan Gordon, Mamadou Bassene, and Shana Watters. 2006. Cording Protocol for Statuses on the Givenness Hierarchy. Available at: http://www.sfu.ca/~hedberg/Coding_for_Cognitive_Status.pdf.

Hawkins, John A. 1991. On (in)definite articles: Implicatures and (un)grammaticality predictors. *Journal of Linguistics* 27. 405-442.

Hedberg, Nancy. 1996. Word order and cognitive status in Mandarin discourse. In Fretheim, Thorstein and Jeanette K. Gundel (eds.). Reference and Referent Accessibility. John Benjamins, 173-192.

Hedberg, Nancy, Emrah Görgülü, and Morgan Mameni. 2009. On Definiteness and specificity in Turkish and Persian. Proceedings of the 2009 Annual Conference of the Canadian Linguistic Association.

Himmelmann, Nikolaus P. 1996. Demonstratives in narrative discourse: A taxonomy of universal uses. In Fox, Barbara (ed.), Studies in Anaphora. Amsterdam: John Benjamins. 205-254.

Horn, Laurence R. 1972. On the Semantic Properties of Logical Operators in English. Doctoral Dissertation, University of California at Los Angeles.

Jackendoff, Ray. 1972. Semantics and Generative Grammar. Cambridge, MA: MIT Press.

Jarbou, Samir Omar. 2010. Accessibility vs. physical proximity: An analysis of exophoric demonstrative practice in Spoken Jordanian Arabic. Journal of Pragmatics 42. 3078-3097.

Lambrecht, Knud. 1994. Information Structure and Sentence Form. Cambridge: Cambridge University Press.

Li, Charles N. and Sandra Thompson. 1981. Mandarin Chinese: A Functional Reference Grammar. Berkeley: University of California Press.

Li, Charles N. and Sandra Thompson. 1975. The semantic function of word order: a case study in Mandarin. In Li, Charles N. (ed.), Word Order and Word Order Change. Austin: University of Texas Press, 163-195.

Matthewson, Lisa. 1998. Determiner Systems and Quantificational Strategies: Evidence from Salish. The Hague: Holland Academic Graphics.

Mulkern, Ann. 2007. Knowing who's important: Relative discourse salience and Irish pronominal forms. In Hedberg, Nancy and Ron Zacharski (eds.). The Grammar-Pragmatics Interface: Essays in Honor of Jeanette K. Gundel. Amsterdam/Philadelphia: John Benjamins. 113–142.

Prince, Ellen F. 1981. Toward a taxonomy of given-new information. In Cole, Peter, Radical Pragmatics. New York: Academic Press. 223-255.

Reinhart, Tanya. 1982.Pragmatics and linguistics: An analysis of sentence topics. Philosophica 27. 53-94.

Vallduvi, Enric. 1992. The Information Component. New York: Garland.