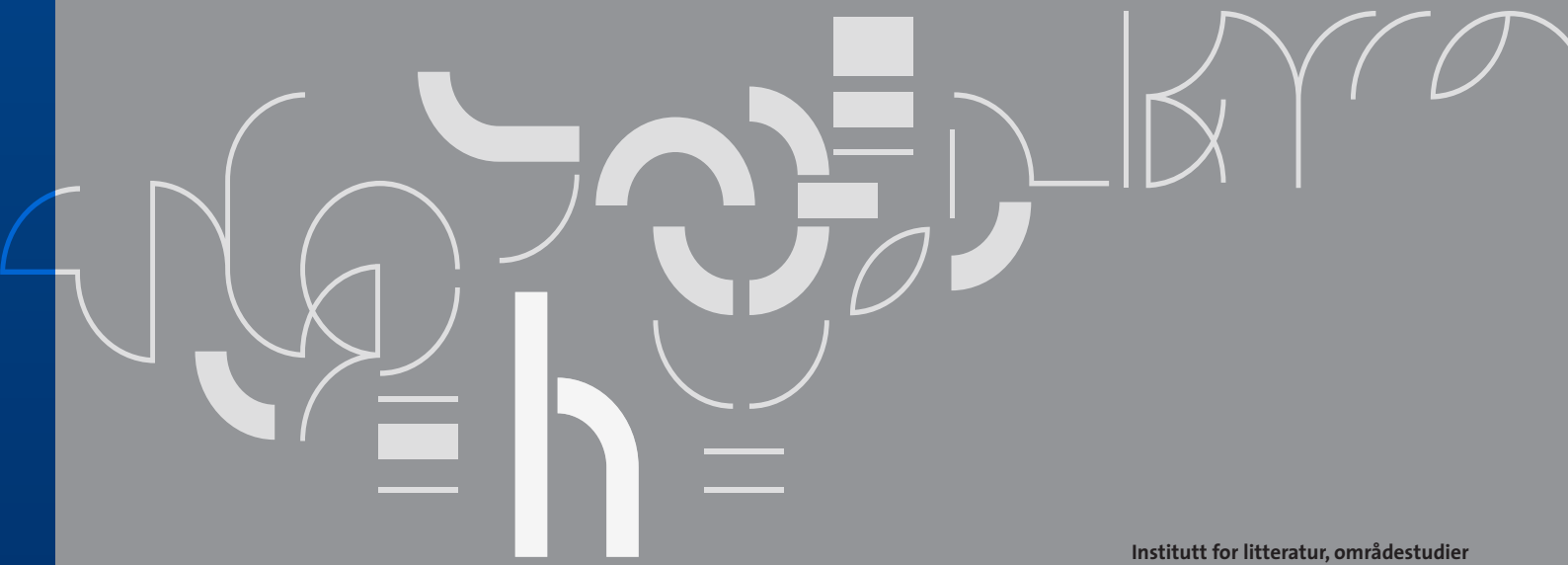
UNIVERSITY OF OSLO
FACULTY OF HUMANITIES

A Semantic Explication of Information Status and the Underspecification of the Recipients' Knowledge

Arndt Riester
Institute for Natural Language Processing (IMS)
Universität Stuttgart



Institutt for litteratur, områdestudier
og europeiske språk

Department of Literature, Area Studies and
European Languages

Grønn, Atle (ed.): Proceedings of SuB12, Oslo: ILOS 2008 (ISBN 978-82-92800-00-3), 508-522.

A Semantic Explication of *Information Status* and the Underspecification of the Recipients' Knowledge*

Arndt Riester
Institute for Natural Language Processing (IMS)
Universität Stuttgart

arndt.riester@ims.uni-stuttgart.de

Abstract

This article presents a survey of and an investigation into the notion of *information status*. Based on insights from DRT and presupposition theory a new variant of IS taxonomis is developed, considering issues such as accommodation and underspecification of text with regard to hearer knowledge.

1 Introduction

Compared to the often (and sometimes sloppily) used notion of *givenness*, *information status* is a more general concept. I will provide the reader with an overview on the most important aspects of this notion starting with Prince (1981, 1992). I will discuss what has become of Prince's key insights in the contemporary IS annotation literature and furthermore point out a number of unsolved problems relating to accommodation and textual underspecification. These problems can be tackled when considering various kinds of progress that have been achieved in presupposition theory and Discourse Representation Theory. Following my survey I propose an annotation scheme that integrates those findings. The annotation system is currently applied in a research project in the Stuttgart SFB 732.

*The research on which this paper is based was enabled by the German Research Foundation (DFG) and is part of the Stuttgart Sonderforschungsbereich 732 "Incremental Specification in Context" (Project A1). It came about in cooperation with Katrin Schweitzer, Hans Kamp and Grzegorz Dogil. I would furthermore like to thank my student assistants, in particular Lorena Killmann and David Lorenz, for their clever remarks and observations. More thanks go to Michael Götze and Julia Ritz of the Potsdam SFB 632 (and Stefanie Dipper as a former member of it) and their colleagues for sharing thoughts and techniques on annotations of and in relation to information structure.

Grønn, Atle (ed.): *Proceedings of SuB12*, Oslo: ILOS 2008 (ISBN 978-82-92800-00-3), 508–522.

2 Approaches to Information Status

2.1 Prince (1981)

In her seminal paper, Ellen Prince (1981) proposes a classification system for referential expressions in text according to different degrees or ways of givenness and novelty. Prince draws on a number of earlier attempts from the literature to define what it means for an expression to be *given*. Prince's proposal basically distinguishes between three top-level categories: *new, inferrable* and *evoked*, as shown in figure 1.

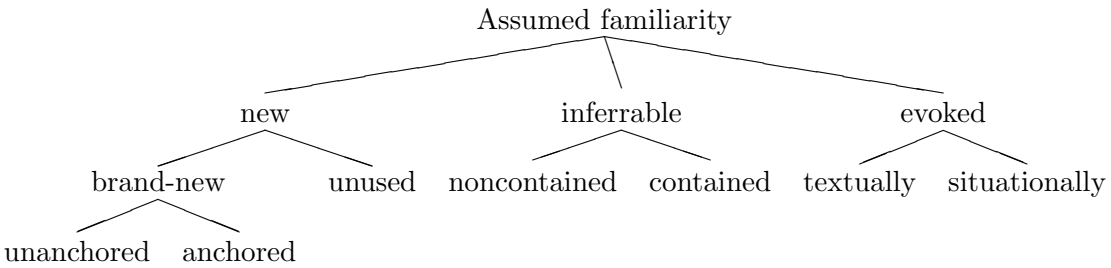


Figure 1: Familiarity taxonomy according to Prince (1981)

I need to emphasize that, although it looks as if we were dealing with a givenness continuum, it turns out that this is a question which is far more complex and should therefore not be answered prematurely. "Shared knowledge", as used in Clark and Haviland (1977), and defined as the speaker's assumptions about what the hearer knows, plays an important role in understanding most of the proposed categories.

Evoked

There are several reasons why the referent of an NP can be part of the shared knowledge. One is that it has been mentioned in the previous text, as (1), or it is available as part of the dialogue situation, which includes the discourse participants as well as individuals and objects to which they have visual, acoustic or some other form of direct access, like (2). Prince (1981) proposes the labels *textually/situationally evoked* to be assigned to the two types of NP uses.[1]

(1) Last week I had an argument with someone at the bus stop. The man was 1,90 m tall.

(2) *(pointing)* The girl with the bike is my niece Miriam.

[1]The expressions in question are marked by means of boxes, while the antecedent from which inferences are drawn is underlined.

Inferrable

An items whose referent is neither present in the discourse as yet nor in the situation but which is identifiable in the face of some other – already *evoked* – entity is called *inferrable*. An example is given in (3).

(3) George returned his laptop [to the dealer] because [the keyboard] was defective.

Although the "antecedent" of the *inferrable* item is typically present in the previous text it might also be situationally evoked, as in a case where the speaker points to a car while uttering (4).

(4) [The battery] is dead.

A group of constructions, of which possessives may be the most prominent, represents a special type of inferrables. They are characterized by their property of *containing* the entity (or set of entities) from which their referent can be inferred, as in (5). In other words, the antecedent of these *contained inferrables* is part of the expression itself rather than appearing elsewhere in the text or the environment.

(5) a. [one **of these eggs**]
b. [the door **of the Bastille**]

New

Finally, if an entity is neither *evoked* nor can be inferred from another available entity it is *new*. However, a distinction is made – and here, shared knowledge plays a role again – between those entities that the hearer knows from some earlier experience and those which he learns about for the first time. The former are called *unused* (example (6)), whereas the latter are referred to as *brand-new*, as in (7). *Brand-new* entities, especially indefinites as in (8), can however be *anchored* to some other entity.

(6) the sun

(7) a guy

(8) a. [a guy **I** work with]
b. [a friend **of mine**]

It is not entirely clear how to demarcate the uses where the referent of an expression should be said to be "*inferrable* from some *contained* entity" from those cases where the referent should be described as "*brand-new* and *anchored* within that entity".

2.2 Prince (1992)

A reorganisation of the categories is undertaken in Prince (1992). New is the emphasis on two dimensions of what is now called *information status*, viz. the status of the hearer and that of the discourse. However, not all categories from the previous proposal can be fitted into the scheme in table 1.

	Discourse-new	Discourse-old
Hearer-new	*brand-new*	–
Hearer-old	*unused*	*textually evoked*

Table 1: Information Status dimensions from Prince (1992)

Some remarks about table 1 are in order. First, as Prince remarks, the two dimensions are not independent of each other. If an entity is *discourse-old*, it is also necessarily *hearer-old*, hence the empty right upper area. Second, a classification as in the above table, can also be unfolded into a givenness hierarchy as in table (9).

(9) HN/DN *(brand-new)* < HO/DN *(unused)* < HO/DO *(textually evoked)*

Third, it is apparent that neither *situationally evoked* items nor *inferrables* can be sufficiently described by means of these two dimensions. As we can also not subsume them under any of the above categories we end up having five main classes.

2.3 Nissim et al. (2004)

In Nissim et al. (2004), an attempt is made, on the one hand, to provide finer distinctions (e.g. whether an *evoked* entity is expressed by means of a pronoun or else) on the other hand to further integrate the zoo of information status categories. One of their main intentions is to arrive at more consistent annotations by bringing down the number of top-level categories to three again.

Their newly introduced category *mediated* subsumes Prince's categories *unused*, *inferrable* and also *situationally evoked*.[2] A translation guide is provided in table 2.

2.4 Götze et al. (2007)

In the "Potsdam guidelines" (Götze et al., 2007), an elaborate annotation system for information structure is presented, one part of which is also concerned with *information status*. Like in Nissim et al. (2004) a three-way classification is employed, however Götze et al. (2007) use a different terminology: *old/textually evoked* is now simply called

[2]It should be pointed out that Nissim et al. (2004, p. 1023) contains two severe misrepresentations: *mediated* does **not** as it is claimed in a footnote "[correspond] to Prince's (1981; 1992) *inferrables*" and it is also not true that "generally known entities [...] such as 'the sun', or 'the Pope" can normally be said to be *inferrable* "from the previous conversation".

Nissim et al. (2004)	Prince	explanation
new	brandnew-unanchored (HN/DN)	
mediated-poss	brandnew-anchored/ inferrable-containing	possessee in possessive NP
mediated-part/ -situation/-event/-set	inferrable-noncontaining	"bridging"
mediated-general	unused (HO/DN)	
old-general	(a subset of) situationally evoked	discourse participants
old-* (several types)	textually evoked (HO/DO)	

Table 2: Comparison of Nissim et al. (2004) and Prince (1981, 1992)

given and in close (but not identical) correspondence to *mediated*, they postulate a category *accessible*[3]. By using less subcategories the classification of Götze et al. (2007) suggests itself to be easily applicable. However, providing too few categories may on the other hand carry the risk of annotators being forced into unintuitive decisions and the annotations are in danger of becoming less meaningful than they could have been otherwise. These are certainly important issues to consider though I won't have anything really substantial to say about them here and the current paper is not meant to provide an empirical investigation into the subject of inter-annotator agreement. The relation between the "Potsdam system" and Prince's original proposal is shown in table 3.

Götze et al. (2007)	Prince
new	brandnew-unanchored (HN/DN)
accessible-situation	situationally evoked
accessible-inferrable	inferrable-noncontaining
accessible-general	unused (HO/DN)
given	textually evoked (HO/DO)

Table 3: Comparison of Götze et al. (2007) and Prince (1981, 1992)

2.5 On the appropriateness of information status taxonomies

After having presented a number of different taxonomies it is certainly appropriate to pose the question according to what criteria one should decide between them. There are several arguments that come to mind. First, as I already remarked, practical considerations such as clarity, coverage and the possibility of obtaining a high agreement among annotators is certainly an issue. Second, the choice of a certain classification may obtain independent support if some or all proposed categories possess their characteristic reflexes in syntax (such as a strong tendency of occupying a certain position in word order with respect to other categories) or prosody (like a characteristic pitch accent selection or a specific fine-grained acoustic profile). A lot of research is currently

[3]from Chafe (1994)

being done on these topics. Third, cross-linguistic applicability is yet another factor: is our classification motivated by specific structures of a particular language or is it a general purpose tool that can be used for annotations in any language? And fourth, what can semantic theory tell us about the preference of one system over another? Are there semantic objects that have been proposed elsewhere in the literature that can be brought in accordance with information status categories? It is this fourth group of questions that we shall deal with in the remaining parts of this paper.

In the present section we shall have a look at different kinds of definite descriptions in order to demonstrate how the annotation systems I have mentioned differ with regard to their category assignments and to point out phenomena for which none of them suggests a satisfactory treatment. Let's first have a look at definites with a "familiar" referent as in (10) and assume that they occurred in a text for the first time.

(10) a. the Pope
b. the moon

The classification systems (abbreviations) introduced above would, as expected, assign these items the following labels.

(11) P: unused (HO/DN)
N04: mediated-general
G07: accessible-general

I take it that proper names should be treated as one kind of definite NP and certainly almost on a par with those in (10). This is a position which is not animously agreed upon. Kripke (1972), for instance, defends the view that names, other than definite descriptions, are rigid designators. I shall, however, follow Geurts (1997, p. 320), who claims that "[...]names must be expected to be used and interpreted like other definite NPs." Examples like those in (12) show that sometimes the distinction between what is "a name" or what is simply "used as a name" is impossible to draw.

(12) a. the Netherlands
b. the Tower of London
c. the Holy Spirit

Hence, what follows is that also a familiar name as in (13) should obtain the same tags that occur in (11).

(13) Johnny Depp

This, however, is not as straightforward as one might expect and although nothing contrary is claimed explicitly in the abovementioned literature, examples like (13) are likely to trick annotators into a confusion of form and function. We can see the problem

more clearly if we compare (13) to names like in (14), again assumed to occur in the given discourse or text for the first time.

(14) [John] loves [Mary].

"John" and "Mary" are *unfamiliar* in the sense that they do not refer to persons in our world knowledge (which is very common in typical examples from the semantics literature), hence they would receive the labels below.

(15) P: brandnew-unanchored (HN/DN)
N04: new
G07: new

In other words, "John" and "Mary" will receive the same information status as the *unanchored* indefinites in (16).

(16) [A man] loves [a woman].

In general, there is nothing wrong with an approach like that. After all, very much in the spirit of Ellen Prince's early conception, *information status* is to be kept separate from the formal feature *definite/indefinite*. Yet, it is also well-known that the story about the *novelty* of indefinites (since e.g. Heim (1982)) has been told in a completely different way than the story about the occasional *novelty* of definites as in (14), treated under the phenomenon of *accommodation* (Lewis, 1979, van der Sandt, 1992, Beaver and Zeevat, 2007). Note that an analogous confusion pertains to the notions *inferrable* and maybe also *textually evoked/old/given* (which I shall henceforth call *discourse-given* or – in short –*d-given*).

(17) a. Fred went to a pub late last night. When he arrived [the door] was closed.
b. John walked past the museum. [A painting] had just been stolen.

(18) a. Yesterday, I met my dentist. [The poor chap] just got divorced.
b. Agatha exhibited perfect manners, exactly as one would have expected it from [a lady].

Both *inferrables* as in (17) and *d-given* items as in (18) may occur with either *definite* or *indefinite* marking, an insight which, at first, might be much more puzzling than the by now well-established facts about definites conveying new information. The findings clash with the traditional picture from the dynamic and discourse-semantic literature (*indefinites* introduce "new" discourse referents, while *definites* pick up "given" ones, which is of course true for the prototypical cases.) The examples from (14) to (18), however, demonstrate why it is wise to keep the distinctions *given/new* (or similar) and *definite/indefinite* separate from each other.

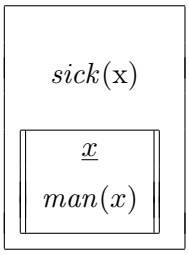
On the other hand, practice also shows that when annotating information status it is helpful to keep two separate sets of labels for either definite or indefinite expressions. Not only does this reduce the error rate of the annotators, it also enables faster access to potentially interesting data. For instance, empirical investigations into the specific syntactic or phonetic properties of information status categories should not only be carried out with regard to *new, inferrable* or *d-given* but one should also take into account more fine-grained classes by distinguishing between *new definites, new indefinites, inferrable definites, inferrable indefinites* and so forth. Such a classification system will be introduced below, but before we do that we shall consider some insights from DRT and presupposition theory.

3 Semantic background: definites, presuppositions & DRT

A well-known approach to the treatment of definite NPs in Discourse Representation Theory (Kamp, 1981, Kamp and Reyle, 1993) is the one in van der Sandt (1992). Definite descriptions, as the prototypical presupposition triggers, are represented as so-called embedded presuppositions (or "A-structures"), which I indicate by means of specially marked boxes. The sentence in (19) generates the *preliminiary DRS* in (20).[4]

(19) The man is sick.

(20)



It is a wide-spread and fairly uncontroversial assumption that presuppositions want to be *bound* to an antecedent, although some of them fail to do so. As discussed in van der Sandt (1992), in some cases where binding fails the referent of the presupposition may get *accommodated* (for instance) in the main DRS.

A more controversial claim is van der Sandt's frequently quoted dictum that *presuppositions* are *anaphors*.[5] Geurts (1999, p. 83) has criticised this claim for its "[inflating] the traditional concept of anaphora beyond recognition". The view which Geurts (1999) defends is rather that **some** presuppositions end up being anaphoric, while others do not. On the other hand, everything which is anaphoric may be represented by means

[4] I shall adopt the convention from Geurts (1999) as well as Kamp's current work to underline the "anaphoric" referent of an embedded presupposition, in order to distinguish it from other, existential discourse referents. More on the problematic notion "anaphoric" below.

[5] Actually, in van der Sandt (1992) it is claimed that presuppositions form a **subset** of the set of anaphors ($P \subset A$) and that pronouns, though anaphoric, aren't presuppositions. This has no influence on what I have to say here. For a longer discussion on these matters see Riester (2008).

of embedded presuppositions ($A \subset P$), including pronouns. One class of presupposition triggers that cannot be called anaphoric are certainly those which get accommodated and, hence, by definition do not have an antecedent.

But even for some of those definite descriptions that end up getting *bound* it is not clear whether we should necessarily call them *anaphoric*. The most prominent examples are those already discussed in (10), repeated in (21) (primary occurrences).

(21) a. the Pope
b. the moon

As is indicated in Geurts (1999), such definite descriptions are *bound* in the hearer's world knowledge. This claim is of course similar as the one from the literature on *information status* that these entities be identifiable on the basis of general knowledge. The choice is now whether to extend the meaning of *anaphoricity* from *being bound in the discourse context* to *being identifiable* or to remain more conservative and translate *anaphoric* as exclusively *discourse given*, while using *bound* in the general case. At the moment, I am undecided on this point, as I see advantages in both options.

3.1 Context Theory

The question of binding in contexts other than the actual discourse leads to a second one: which and how many different "contexts" do we have to assume? A recent paper by Kamp (ms.) addresses exactly this question. Kamp draws our attention to the fact that beyond the discourse context, which in the DRT literature has played an almost exclusive role, there are a number of other contexts that need to be taken into account as information sources in which expressions find their referents. The sum of these contexts is referred to as the "articulated context" and consists of the 4-tuple given in (22).

(22) $\langle K_{dis}, K_{env}, K_{gen}, K_{enc} \rangle$

As expected, K_{dis} is the familiar dynamic discourse context, representing a "protocol" of the previous spoken or written conversation.[6]

There is, furthermore, the *environment context* comprising all elements in the immediate dialogue situation. These are the elements that are typically, though not exclusively, picked up by means of demonstratives.

Next, there are two contexts containig the "shared assumptions" of speaker and addressee. The reason for having two contexts is that they differ with regard to the type of knowledge they contain. K_{gen}, the *generic context component*, contains conditional

[6]In Kamp's setup, K_{dis} subsumes part of the *utterance context*, K_{utt}, in the sense of Kaplan (1989) (that part which consists of speaker, addressee and utterance time). On Kaplan's own account these are treated as forming one context together with Kamp's *environment*. However, as Kamp argues, there is a clear distinction between the former entities and the environment context in that the former are always available, while the latter is not, for instance not in telephone conversations or in writing.

information describing, for instance, causal relations between events and other entities. As we are primarily interested in individual type entities, we shall not be concerned with this context component.

More relevant for the present paper, however, is K_{enc}, the *encyclopaedic context component*, which consists of the entities that the speaker may assume his addressee to have knowledge about.

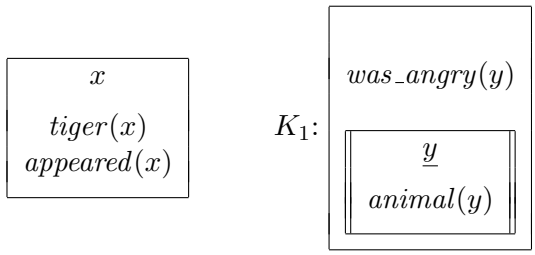
3.2 Binding in different context components

The advantage of a complex "articulated context" is that we need not worry about spelling out the semantics of different types of definite descriptions in the semantics but in fact may stick to one simple and general representation as e.g. the one in (20) above. The only thing that is needed in addition is an external resolution mechanism that guides an embedded presupposition through the different contexts in search of an antecedent. Due to space limitations, I shall only compare the resolution of definites in two different context components, viz. in the discourse context (example (23)-(25)) and in the encyclopaedic context (example (26)-(28)).

(23) a. A tiger appeared.
b. [The animal] was angry.

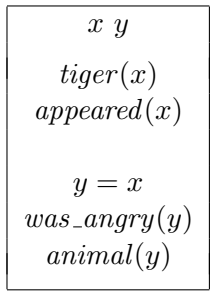
Assume that sentence (23b), represented as the preliminary DRS K_1 in (24), is uttered in the discourse context K_{dis_1}, made up of the content of sentence (23a).

(24)



As is standard procedure in contemporary DRT, first the embedded presupposition of K_1 is resolved in K_{dis_1} and then the resulting representations are merged, yielding the DRS K'_{dis_1} depicted in (25).

(25)



Now compare this to the processing of sentence (26) (K_2 in (27)) in an empty discourse context K_{dis_2}.

(26) The Pope was angry.

Since the embedded presupposition of K_2 cannot be resolved in K_{dis_2}, the encyclopaedic context component K_{enc_2} is consulted instead. Actually, the form in which I have depicted K_{enc_2} in (27) is necessarily a tremendous simplification. As the word "encyclopaedic" suggests, this context component is supposed to contain the entire set of individual entities the hearer has knowledge about, in combination with all sorts of information related to these entities. It is evident that in a paper presentation such a resource can only be displayed partially.

(27) K_{dis_2}: [ ] K_{enc_2}: [x … ; $Pope(x)$; …] K_2: [$was\_angry(y)$; [$\underline{y}$; $Pope(y)$]]

As soon as a referent for the embedded presupposition in K_2 is found in K_{enc_2} this referent (x) is copied (along with its associated property of being Pope) into the discourse context. I shall call this copying process *activation*. Finally, the discourse referent y from K_2 is linked to x and the DRSs are merged, the result of which is shown in (28).

(28) K'_{dis_2}: [[x *(copy from* K_{enc}*)*] y ; $Pope(x)$; $y = x$; $was\_angry(y)$]

3.3 Accommodation

Note that although from the perspective of the discourse context the process described in (27)-(28) looks like a case of *accommodation* it is actually a different matter. As I already mentioned in section 2.5, Geurts (1999) describes it as a type of *binding*. An instance of an NP that is likely to get accommodated is the one in (29).[7]

[7]The indigenous population in the German Southwest is rather fanatic in their habit of scrubbing the communal areas in and in front of their houses on a weekly basis. Nevertheless, an association as in (29) has not been heard of. I take it, though, that a cooperative dialogue partner, at least one having lived there, would not hesitate to accommodate its existence upon hearing (29).

(29) [The chairman of the Stuttgart street sweeping association] fell ill.

In the case of (29) there is no context component that might possibly contain a referent for the subject NP, hence nothing can be copied into the discourse context and a discourse referent must be newly created.

4 A new classification

It should by now have become easy to anticipate what the enterprise of the current article is meant to be. The central claim that I would like to make is the one below.

> **Claim: from a semantic point of view, Information Status categories should reflect the context components in which the presuppositions triggered by referential expressions are bound.**

Two things have to be added at this point: (i) for cases as in (29), a separate category label should be introduced to capture the cases of unbound or "to-be-accommodated" expressions. Since the only information that a hearer has about such entities consists in the description itself, I propose the name *accessible-via-description* for this category. (ii) What is intuitively missing in Kamp's concept of the *articulated context* but what is needed to cover the whole variety of definite NPs are smaller scenario contexts which serve as the referent supply in *bridging* cases.[8] In the case of (3), repeated here as (30) the scenario would consist of the set of things commonly associated with laptops.

(30) George returned his laptop [to the dealer] because [the keyboard] was defective.

The proposed classification for definite NPs is the one in table 4.

binding context	definite
K_{dis}	d-given
K_{env}	situative
K_{enc}	accessible-general/ encyclopaedic
$K_{scenario}$	bridging
none	accessible-via-description/ accommodation

Table 4: Proposed Information status classification for definite NPs

As for the indefinite domain, I propose to use the labels *new* for hearer-new/discourse-new indefinite expressions, and *partitive* for indefinite inferrables, like the one in (17b). I must add, though, that there are some further cases of "anaphoric", "specific" or otherwise "non-novel" indefinites (cf. Krifka (2001), Portner and Yabushita (2001), Geurts

[8] Kamp (ms.) doesn't need such scenario contexts but uses so-called context predicates to capture the meaning of bridging NPs. His approach deviates from the one proposed here, cf. Riester (2008).

(2002)) whose exact classification in accordance with the scheme proposed here still has to wait.

5 Underspecified knowledge

As the last point of discussion consider the follwing expression taken from a corpus of German radio news.

(31) der EU-Außenbeauftrage Solana
the Representative for the Common Foreign Policy of the EU, Solana

The question whether this item should receive the label *accessible-general* or *accessible-via-description* will crucially depend on the knowledge of the addressee. Radio and other news are typically addressed at a rather inhomogeneous audience. While person A may be fully aware of Mr. Solana and the position he is currently in, person B may have just been familiar with the name, while person C may not have heard of Solana at all. What this means is that different people will need different amounts of accommodation when processing the phrase in (31).[9] It would be impossible and quite unfortunate to demand of annotators in such cases that they decide between one or the other label. The option that is pursued in the Stuttgart SFB corpus annotations at present is to assign ambiguous labels together with a little flag to indicate which option is the most likely one. An sample annotation using the SALTO tool (Burchardt et al., 2006) is shown in figure 2.

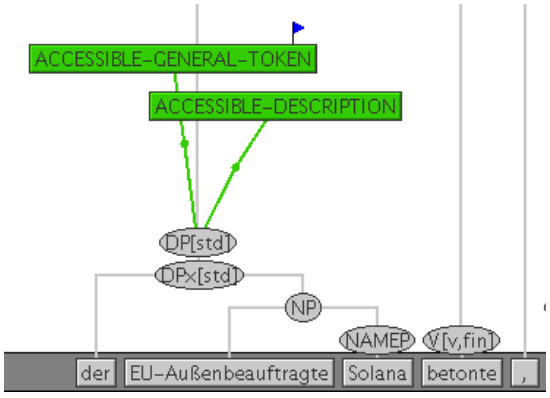


Figure 2: Ambiguous annotation of information status

[9]Kamp and Rossdeutscher (1994) use the notion *presupposition justification* for such hybrid cases between binding and accommodation.

6 Summary and outlook

In the present article I have discussed the notion of *information status* and different proposals as to how it should be annotated, including my own, which is fundamentally motivated by considerations based on DRT and presupposition theory. Developing a feasible annotation system is a prerequiste for creating resources for the investigation of prosodic and word order phenomena, which clearly depend on the concept of information status. My last remarks on the variation in knowledge among the intended recipients of certain text types (such as newspaper articles or radio news bulletins) and the underspecification of the encyclopaedic context for texts of this type that is entailed by this variation, however, should have made the reader alert that information status cannot always be unambiguously determined on the basis of the text alone.

References

David Beaver and Henk Zeevat (2007) "Accommodation", in Gillian Catriona Ramchand and Charles Reiss (eds.) *The Oxford Handbook of Linguistic Interfaces*, Oxford University Press.

Aljoscha Burchardt, Katrin Erk, Anette Frank, Andrea Kowalski, and Sebastian Padó (2006) "SALTO: A Versatile Multi-Level Annotation Tool", *Proceedings of LREC-2006*, Genoa.

Wallace L. Chafe (1994) *Discourse, Consciousness, and Time*, University of Chicago Press.

Herbert H. Clark and Susan E. Haviland (1977) "Comprehension and the Given-New Contract", in R. Freedle (ed.) *Discourse Production and Comprehension*, Erlbaum, Hillsdale, NJ, 1–40

Bart Geurts (1997) "Good News about the Description Theory of Names", *Journal of Semantics* **14**, 319–348.

Bart Geurts (1997) *Presuppositions and Pronouns*, Elsevier, Oxford.

Bart Geurts (2002) "Specific Indefinites, Presupposition, and Scope", in U. Reyle R. Bäuerle and T. Zimmermann (eds.) *Presuppositions and Discourse*, Elsevier, Oxford.

Michael Götze et al. (2007) "Information Structure" in Stefanie Dipper, Michael Götze, and Stavros Skopeteas (eds.) *Information Structure in Cross-Linguistic Corpora: Annotation Guidelines for Phonology, Morphology, Syntax, Semantics and Information Structure*, volume 7 of Interdisciplinary Studies on Information Structure, Universitätsverlag Potsdam.

Irene Heim (1982) *The Semantics of Definite and Indefinite Noun Phrases*, PhD thesis, University of Massachusetts, Amherst.

Hans Kamp (1981) "A Theory of Truth and Semantic Representation", in J. Groenendijk, T. Janssen, and M. Stokhof (eds.) *Formal Methods in the Study of Language*, Mathematical Centre Tracts, Amsterdam.

Hans Kamp (2008) *Discourse Structure and the Structure of Context*, ms., IMS, Universität Stuttgart.

Hans Kamp and Uwe Reyle (1993) *From Discourse to Logic. Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*, Kluwer, Dordrecht.

Hans Kamp and Antje Rossdeutscher (1994) "DRS-Construction and Lexically Driven Inference", *Theoretical Linguistics*, **20**(2/3), 165–235.

David Kaplan (1989) "Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and other Indexicals", in J. Almog, J. Perry, and H. Wettstein (eds.) *Themes from Kaplan*, Oxford University Press, 481–563.

Manfred Krifka (2001) "Non-Novel Indefinites in Adverbial Quantification", in C. Condoravdi and G. Renardel de Lavalette (eds.) *Logical Perspectives on Language and Information*, CSLI Publications, Stanford, 1–40.

Saul Kripke (1972) *Naming and Necessity*, Harvard University Press, Cambridge, MA.

David Lewis (1979) "Scorekeeping in a Language Game", *Journal of Philosophical Logic* **8**, 339–359.

Malvina Nissim, Shipra Dingare, Jean Carletta, and Mark Steedman (2004) "An Annotation Scheme for Information Status in Dialogue", in *Proceedings of the 4th Conference on Language Resources and Evaluation (LREC2004)*, Lisbon.

Paul Portner and Katsuhiko Yabushita (2001) "Specific Indefinites and the Information Structure Theory of Topics", *Journal of Semantics* **18**, 271–297.

Ellen F. Prince (1981) "Toward a Taxonomy of Given-New Information", in P. Cole (ed.) *Radical Pragmatics*, Academic Press, New York, 233–255.

Ellen F. Prince (1992) "The ZPG Letter: Subjects, Definiteness and Information Status", in W. C. Mann and S. A. Thompson (eds.) *Discourse Description: Diverse Linguistic Analyses of a Fund-Raising Text*, Benjamins, Amsterdam, 295–325.

Arndt Riester (2008) *The Components of Focus and their Use in Annotating Information Structure*, PhD thesis, IMS, Universität Stuttgart.

Rob van der Sandt (1992) "Presupposition Projection as Anaphora Resolution", *Journal of Semantics* **9**, 333–377.