To appear in: Proceedings of SALT V, Cornell University, 1995

The Pragmatics of Word Meaning

Alex Lascarides
University of Edinburgh

Ann Copestake
University of Cambridge

Abstract

In this paper, we explore the interaction between lexical semantics and pragmatics. Linguistic processing is informationally encapsulated and utilises relatively simple 'taxonomic' lexical semantic knowledge. On this basis, defeasible lexical generalisations deliver defeasible parts of logical form. In contrast, pragmatics is open-ended and involves arbitrary knowledge. Two axioms specify when pragmatic defaults override lexical ones. We demonstrate that modelling this interaction allows us to achieve a more refined interpretation of words in a discourse context than either the lexicon or pragmatics could do on their own.

1 Introduction

Much recent work on lexical semantics has been concerned with accounting for the flexibility of word meaning. For example, there are many cases of regular polysemy, which includes not only the familiar verb alternations and nominalisations but also related senses of nouns such as the mass/portion senses *beer*, *coffee* etc and the container/contents alternation found with *box*, *case* and so on. There is also the phenomenon which Pustejovsky (e.g. 1991, in press) has called *logical metonymy* where additional meaning seems to arise for particular verb/noun or adjective/noun combinations. For example, (1a) usually has the same interpretation as (1b):

(1) a. Mary enjoyed the book.

b. Mary enjoyed reading the book.

Syntactic realisation and lexical semantics are closely interrelated: it is not an arbitrary fact that *beer* is a mass noun when it refers to an unbounded quantity of liquid, and a count noun when it refers to the portion, or that *enjoy* can take a nominal complement. The challenge is to account for these processes compositionally in a way which allows for their partly conventional nature, within a general framework of linguistic description that recognises the role of pragmatics.

Briscoe *et al* (1990) and Copestake (1992) argued for an interaction between lexical semantics and pragmatics in which purely linguistic processing is informationally encapsulated and utilises relatively simple 'taxonomic' lexical semantic knowledge. Lexical semantic information and real world knowledge were not seen as necessarily distinct. Instead, linguistic processes had limited access to world knowledge, which could therefore interact with knowledge of language and possibly become conventionalised in various ways. For processes such as logical metonymy, linguistic processing delivered a partly defeasible logical form, which could be overridden by open-ended pragmatic reasoning. This work concentrated on providing an account of lexical semantics which was integrated with syntax and compositional semantics by utilising a uniformly unification-based approach. However, the account was incomplete, because the interaction with pragmatics was left open. Defaults were simply used to aid in the encoding of static lexical generalisations. Thus the use of lexical defaults in syntax and morphology (e.g. Flickinger (1987), Evans and Gazdar (1989)) was extended to lexical semantics. But it was not related to the notion of defeasibility in the logical form, making it unclear how the unification based techniques served to distinguish defeasible from indefeasible parts of logical form.

Here we review the earlier account and argue for a new treatment of defaults, which allows default results of lexical generalisations to *persist* as default beyond the lexicon and thus be available to the interface with pragmatic reasoning. We will make specific proposals for the formalisation of the pragmatic component, and illustrate how this allows us to account for alternative interpretations of words in a discourse context. The decision as to whether the lexical default survives at the discourse level or not will be modeled in a formally precise way in the nonmonotonic logic for pragmatic reasoning. Just two rules will be needed to encode the communication link between default reasoning in the lexicon on the one hand, and default reasoning at the discourse level at the other. By providing this link between lexical operations and discourse ones, we will explain how words are interpreted in discourse, in a way that neither the lexicon nor pragmatics could achieve on their own.

2 Generalisations with Exceptions

Verbs such as *enjoy* can be described as selecting semantically for an event in examples such as (2):

(2) Mary enjoys playing the guitar.

However *enjoy* can also take an NP complement, and in (1a) the complement *the book* denotes a physical object:

(1) a. Mary enjoyed the book.

b. Mary enjoyed reading the book.

Traditionally the only way to handle this is to assume two lexical entries for *enjoy* and to relate the different senses by meaning postulates. However, quite apart from the undesirability of proliferating senses, this does not explain why the usual reading of (1a) is (1b), and it misses the generalisation to other cases where a noun phrase is interpreted as an event, such as those in (3).

(3) a. John began a new book.

b. John finished the beer.

c. Bill enjoyed the film.

d. After three glasses of champagne, John felt much happier.

It also does not allow for cases where an NP and a VP are conjoined, such as (4):

(4) Mary enjoys books, television and playing the guitar.

Pustejovsky (e.g., 1991) proposes that examples such as (1a) be treated as involving *logical metonymy*. He treats nouns as having *qualia structure* as part of their lexical entries which specifies possible events associated with the entity. For example, the telic (purpose) role of the qualia structure for *book* has a value equivalent to *reading*. When combined with *enjoy*, a metonymic interpretation is constructed where the particular sort of event which is likely to be involved can be determined from the qualia structure, which results in an interpretation for (1a) equivalent to (1b). In §3, we outline an account which is broadly similar to Pustejovsky's. In our treatment of (1a), the verb provides the basic metonymic interpretation, which can be glossed as (5a) with the logical form shown in (5b):[1]

(5) a. Mary enjoyed some event associated with the book.

b. $\exists y, e, e'[\mathit{enjoy}(e, \mathrm{Mary}, e') \wedge \mathit{object}(e, e') \wedge \mathit{act\text{-}on\text{-}pred}(e', \mathrm{Mary}, y) \wedge \mathit{book}(y)]$

c. $\exists y, e, e'[\mathit{enjoy}(e, \mathrm{Mary}, e') \wedge \mathit{object}(e, e') \wedge \mathit{read}(e', \mathrm{Mary}, y) \wedge \mathit{book}(y)]$

The constant *act-on-pred* is general over a broad class of predicates which we will not attempt to precisely delimit here, but which includes *watch*, *eat*, *smoke* and so on. However, the noun phrase provides the specific predicate involved, thus giving the interpretation shown in (5c) (which corresponds to (1b)). If the noun does not have a conventionalised telic role, the sentence is odd (out of context), as in (6):

(6) ? Mary enjoyed the pebble.

2.1 Exceptions to the generalisation

This generalisation about the interpretation has two classes of exceptions. The first case is exemplified by (7):

(7) ? Mary enjoyed the dictionary.

This is odd (again out of context) because dictionaries are usually used as reference books, and so its telic role *referring*, which is point-like, doesn't combine easily with *enjoy*, which has to be true of an event with significant duration. Thus (8) is only natural on an iterative interpretation:

(8) Mary enjoyed referring to the dictionary.

In Briscoe *et al* (1990) and Copestake (1992), such cases are allowed for by using a default inheritance hierarchy in the lexicon. So, although *dictionary*, like *book*, could inherit its lexical semantic characteristics from a more general class such

[1] Here and in the following examples we ignore temporal information for the sake of simplicity.

as **literature**, the telic role of the qualia structure specified for *dictionary* corresponds to *refer to*, and this overrides the inherited value *read*. The use of defaults in the lexicon was taken to be strictly part of the description language, and led to a conventional lexical entry expressed as a typed feature structure. Using defaults is an important part of our theory of lexical structure, since it allows concise specification of lexical entries and avoids redundancy. However, purely lexical defaults do not extend to the second class of exceptions, which are triggered by context, or wider world knowledge. For example, (9a) means (9b) and not (9c):

(9) a. My goat eats anything. He really enjoyed your book.

b. The goat enjoyed eating your book.

c. The goat enjoyed reading your book.

Similarly, our interpretation of *Mary enjoyed the book* is different if we know that Mary is a goat and is revised if we subsequently learn this.

Briscoe *et al* (1990) allow for the second type of defaults by introducing a consistency operator M (cf. Reiter 1980) into the part of the logical form derived from the telic role. Thus the logical form of (1a) is (10); it can be glossed as "the event enjoyed is reading, in the absence of information to the contrary":

(10) $\exists y, e, e'[enjoy(e, Mary, e') \wedge object(e, e') \wedge M read(e', Mary, y) \wedge book(y)]$

This account was intended as a placeholder in the absence of a proper treatment of pragmatics. Even so, it has some major disadvantages. Firstly, the assumption that goats don't read is itself default, because of contexts such as fairy stories. Assuming that this default is encoded in the same logic, it is unclear how one could ensure that the axioms on M resolve the conflict between the default logical form and the default world knowledge in favour of the latter, especially since the two defaults are logically unrelated. Secondly, the consistency operator is introduced into the grammar as an ad-hoc stipulation. There is no connection between the defeasibility of the telic role with respect to its inheritance in the *dictionary* case and its defeasibility in the logical form. The pragmatic overriding in the goat example is due to the subject of *enjoy*. But the object can also have this effect, as shown in the examples in (11), given that *book made out of marzipan* and *book with blank pages* can't be lexicalised (unlike *dictionary*).

(11) a. John enjoyed the book made out of marzipan.

b. ? John enjoyed the book with blank pages.

Intuitively, these cases are just like the dictionary one, in that they arise because the object is an abnormal book. In fact, we hypothesise that all cases of overriding of the logical form arise because the context is such that the entity is being used in an abnormal way. Ideally, therefore, we would like the defeasibility in the logical form to arise from the default nature of the usual purpose specification made in the lexicon. But, because defaults in Briscoe *et* *al* (1990) are simply part of the lexical description language, they could not persist beyond the lexicon, and the defeasibility in the logical form had to be stipulated.

It is implausible that these problems could be resolved by adopting a purely lexical account, since arbitrarily complex reasoning could be involved in deciding that the subject can't read or that the object is unreadable. The alternative would be to claim that the interpretation of the event was purely pragmatic (i.e. that the logical form for (1a) was simply (5b), with the interpretation of the predicate *act-on-pred* being completely pragmatically determined). Such an approach is suggested by Hobbs *et al.* (1990) who use weighted abduction on pragmatic knowledge to determine the value of the underspecified predicate. But serious challenges to this line exist (see also Briscoe *et al* (1990)).

First, an adequate theory has to account for the usual interpretations. The corpus analysis described in Briscoe *et al* (1990) showed that for most metonymic examples the telic role of the noun gives an appropriate reading. What's more, the explicit mention of the verbal predicate is relatively rare in such cases—that is, examples such as (1a) are more common than (1b). On the other hand, the contexts in which the interpretation would not have been predicted by the qualia structure were informationally-rich (a concept which we will be able to formalise in §4). A purely pragmatic theory could only account for this data by assuming that some interpretations were privileged; for example, one would need a rule that encapsulates that *enjoy the book* by default means *enjoy reading the book*. But this would cause the same problems with prioritising defaults which we mentioned above. That is, one would have to impose prioritorisations on world knowledge that weren't independently motivated, because the conflicting knowledge that was pertinent to the case would be logically unrelated. In the case of weighted abduction, it is thus unclear how one can assign the weights that guide inference in a principled way.

Furthermore, there is some evidence which suggests that logical metonymy is partially conventionalised and triggered by the lexical item, rather than knowledge of the context. For example, (12) is strange, even if the hearer and the speaker both know that the doorstop is a book, which would not be predicted if the purpose were pragmatically determined by real world knowledge of the entity:

(12) ? John enjoyed the doorstop.

There are also examples where the coerced form is less acceptable than might be expected if interpretation was simply a matter of finding a possible event: for example, (13) seems relatively strange, even though hiking is a common recreational activity:

(13) ? John enjoyed the path.

Godard and Jayez (1993) give some data for *commencer* which leads them to suggest that the telic interpretation is only available for objects which are being in some sense consumed or affected by the action. However they then have to assume that books are affected by being read. Since it is unlikely that real world properties of books would necessarily lead to this conceptualisation, these exceptions support the hypothesis that logical metonymy is partially conventionalised. Conventionalised exceptions can be lexically encoded, but resist a purely pragmatic treatment.

We therefore did not want to reject the hypothesis that the lexicon proposes a partially defeasible logical form. Instead, we make use of a new formalisation of defaults, which allows them to persist beyond the lexicon. The default nature of the part of the logical form contributed by the telic role is not simply stipulated, but arises directly from the lexical default. The interface with pragmatics is set up so that reasoning with real world knowledge can override the defaults proposed lexically. Thus we can provide an integrated account of the interaction of lexical semantics and pragmatics. We describe this account in §3 and §4 but first we briefly review some other data which requires this sort of treatment.

2.2 Adjectives and compound nouns

Some examples of adjective interpretation can be treated along the same broad lines as *enjoy* and other event invoking predicates. Pustejovsky (e.g., 1991, in press) and others have argued against distinct lexical entries for *fast*, for each of its senses in *fast car*, *fast typist*, *fast motorway* and so on. Instead, it is possible to assume just a single lexical entry for *fast*, where its different 'senses' arise from the process of syntagmatic co-composition. The lexical generalisation is much like that for *enjoy*: adjectives like *fast* predicate over the telic role of the artefact (although *fast* can also apply to other parts of the qualia). So the lexical account predicts that *fast car* means *a car which goes fast*, and *fast typist* means *a typist who types fast*, via the same entry for *fast*.

But as before, some discourse contexts trigger exceptions to this generalisation. In (14), *fast typist* means *typist who runs fast*, and not *typist who types fast*.

(14) a. All the office personnel took part in the company sports day last week.

b. One of the typists was a good athlete, but the other was struggling to finish the courses.

c. The fast typist came first in the 100m.

As in the *enjoy* examples, the pragmatic component needs to know that the interpretation of *fast typist* as *a typist who types fast* is a default.

Another case where a default interpretation apparently arises from the lexicon/grammar is the interpretation of compounds. For example, there appears to be a generalisation that when a noun that refers to a solid substance combines with a noun that refers to a solid artefact, the compound refers to the artefact made of the substance (*wickerwork chair*, *plastic toy*, *wrought iron table*, *mahogany dresser*). On the other hand, some compounds can only be interpreted in context. Downing attests (15) in a context where there was a table already set with a glass of orange juice by three places and apple juice by the fourth:

(15) Please sit in the apple juice chair.

Here *apple juice chair* means "chair in front of a place setting with apple juice", but obviously this meaning cannot be listed in the lexicon.

Examples like (15) have led to the suggestion that noun-noun compounds should be assigned a representation where the relationship between the two halves of the compound is left completely unspecified and further interpretation should be left to the pragmatic component (e.g., Bauer, 1983). There are, however, serious objections to this application of the pragmatic dustbin. Without further elaboration it gives no explanation of the fact that the majority of compound nouns behave in a *semi*-regular manner. But if the above generalisation about solid substance/artefact compounds were encoded via a standard default inheritance mechanism, default and non-default information would not be distinguished in the result. But the "made-of" relationship between the nouns in compounds like *wickerwork chair* can be overridden in discourse:

(16) At school, everyone worked on crafts in groups round a big table, sitting on brightly coloured chairs. To make sure everyone could reach the materials, the groups used particular chairs: the wickerwork chairs were made of red plastic, for example.

These observations make noun-noun compounds a good candidate for the use of defaults which persist beyond the lexicon, along broadly similar lines to the discussion of logical metonymy above.

3 Persistent Default Unification and the Lexicon

We use a typed feature structure formalism comparable to that used in HPSG (Pollard and Sag, 1994) to implement the grammar and the lexicon. The standard method of implementing default inheritance within unification-based approaches to linguistic representation is to use some variety of default unification (see Copestake, 1993, for a overview). This is usually taken to be an operation in the description language, which allows one feature structure (FS) to incorporate only the consistent information from another FS. Inconsistent information is ignored, rather than causing failure of the operation as in normal unification. But since default unification returns a normal FS, there is no distinction between default and non-default information in the result. Thus, for example, there is no way of specifying that the telic role for the **literature** class is defeasible. The lexical entry for *dictionary* could override it (in fact *dictionary* could override any part of the information it was inheriting) but there is no way in which it can be stated to be defeasible more generally.

There is another problem with using this operation as a basis for lexical organisation. With the exception of Young and Rounds (1993), default unification is order-dependent. This somewhat compromises the declarativity of the formalism, but is not a serious problem for the lexicon itself, because all the entries to be unified are in a fixed hierarchy and an inheritance order can be imposed. But in a discourse situation, one cannot predict which pieces of information are to be unified, *in advance* of starting the discourse parsing process. So the interface between discourse processing and order dependent lexical

Where $\mathbf{t'}$ is more specific than ($\sqsubset$) $\mathbf{t}$:

$$\begin{bmatrix}\mathbf{t}\\ \mathrm{F} = \mathbf{a}\end{bmatrix} \overset{<>}{\sqcap} \begin{bmatrix}\mathbf{t'}\\ \mathrm{F} = /\mathbf{b}\end{bmatrix} = \begin{bmatrix}\mathbf{t'}\\ \mathrm{F} = \mathbf{a}\end{bmatrix} \qquad \text{Defeat of DMP}$$

$$\begin{bmatrix}\mathbf{t}\\ \mathrm{F} = /\mathbf{a}\end{bmatrix} \overset{<>}{\sqcap} \begin{bmatrix}\mathbf{t'}\\ \mathrm{F} = /\mathbf{b}\end{bmatrix} = \begin{bmatrix}\mathbf{t'}\\ \mathrm{F} = /\mathbf{b}\end{bmatrix} \qquad \text{Specificity/The Penguin Principle}$$

Figure 1: Some examples of PDU

processing would have to take into account the order in which the unification operations are done, which is impractical.

Lascarides *et al* (1994) have defined an order independent form of default unification over typed default feature structures (TDFSs). TDFSs are typed FSs where default information is marked as such, and the default unification operation is one where defaults in a TDFS, if they survive at all, survive with the marking that they are default. So this unification operation is one which permits defaults to *persist* as default beyond the lexicon's boundaries, in the sense that one can distinguish in the FS which parts are default. Because of this, the operation is known as Persistent Default Unification (PDU).

TDFSs are TFSs augmented with a slash notation which demarcates the indefeasible parts from the defeasible. Values to the left of the slash are indefeasible and those to the right defeasible (*indefeasible/defeasible*). We abbreviate this to */defeasible* where the indefeasible value is completely general, and omit the slash when the defeasible and indefeasible values are the same. So, for example, the TDFS (17) states that the value on the feature F is by default G:**a**, although the type of the FS and the existence of the feature F itself is non-default:

(17) $\begin{bmatrix}\mathbf{t}\\ \mathrm{F} = /\begin{bmatrix}\mathrm{G} = \mathbf{a}\end{bmatrix}\end{bmatrix}$

When a default value survives PDU (notated $\overset{<>}{\sqcap}$), it does so with the slash notation. The details of PDU are given in Lascarides *et al* (1994) but two examples are given in Figure 1. These indicate that PDU validates defeat of Defeasible Modus Ponens (DMP), and unlike Young and Rounds' definition, it also validates Specificity (i.e., defeasible information on more specific TDFSs overrides conflicting defaults on more general TDFSs).

Lascarides *et al* (1994) show one way of encoding the inheritance of telic roles in PDU (Figure 2). So, for example, the telic role of **literature** is **read** and this is inherited by **book**, but for the subclass **dictionary** it's **refer-to**. This is superficially similar to previous descriptions, apart from the slash, but here default inheritance can proceed in any order to compute the telic roles.

Copestake and Briscoe (1995) show how to state the lexical generalisation concerning *enjoy*, that it predicates over the telic role of the artefact as shown in Figure 3.[2] When *enjoy* takes an artefact-denoting object (which instantiates the

[2] Unlike Pustejovsky (in press) and Briscoe *et al* (1990), this account assumes that the FS for *enjoy* when it takes an object which denotes an individual entity is distinct from the form which takes an event (although both inherit from a common underspecified form). The 'coercion' from object to event is represented as internal to the verb semantics. Some of the reasons for preferring this account are given in Copestake and Briscoe (1992, 1995) and Godard and Jayez (1993). However the differences between this and the alternative account

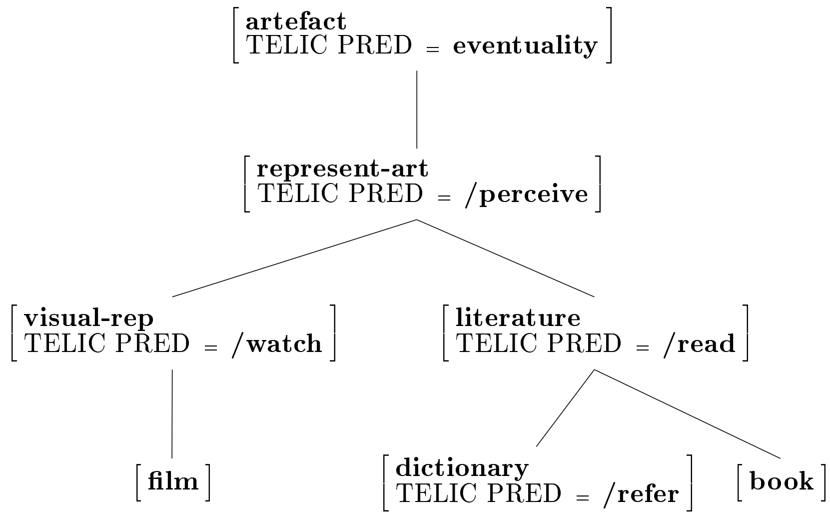

Figure 2: The Telic Role of Artefacts

$$\left[\begin{array}{l}\textbf{coercing}\\ \text{CAT SUBCAT} = \left\langle \left[\begin{array}{l}\textbf{np}\\ \text{SEM} = \boxed{\text{m}}\ [Q(y)]\\ \text{QUALIA TELIC PRED} = \boxed{\text{P}}\end{array}\right]\right\rangle\\ \text{SEM} = [e][\mathrm{R}(e,x,e') \wedge \boxed{\text{P}}\,\text{act-on-pred}(e',x,y) \wedge \boxed{\text{m}}\,]\end{array}\right]$$

Figure 3: The generalisation for verbs like *enjoy*.

CAT SUBCAT 'slot'), the event that is enjoyed is instantiated via the telic role, as indicated by the coindexation $\boxed{\text{P}}$ in Figure 3. In the figure, R is the predicate associated with the verb itself (e.g., **enjoy**) $\boxed{\text{P}}$ and $\boxed{\text{m}}$ indicate coindexation (we are using letters here for readability rather than the conventional integers). The instantiated form is shown in Figure 4.

In these figures the logical form is shown in a linearised notation for readability, rather than in its actual encoding in TDFSs. It is important, however, that we use the same formalism throughout, since it means we can use PDU to construct the semantics, just as normal unification is often used in FS based frameworks. We have shown the path QUALIA TELIC PRED explicitly, to illustrate that it is the predicate itself which is slashed. The semantic representation assumed is InL (Indexed Language, Zeevat *et al* 1987), which has a direct equivalence to DRT. We'll assume the use of DRT here, since this is the semantic representation scheme that underlies the pragmatic component DICE (Lascarides and Asher, 1991, 1993) that we'll link the grammar to. We assume that DRS-conditions that arise from elements on the RHS of the slash notation

---

where the NP itself undergoes coercion are largely irrelevant here.

$$\left[\begin{array}{l}\textbf{coercing}\\ \text{CAT SUBCAT} = \left\langle \left[\begin{array}{l}\textbf{np}\\ \text{SEM} = \boxed{\text{m}}\\ \text{QUALIA TELIC PRED} = \boxed{\text{P}}\textbf{act-on-pred/read}\end{array}\right]\right\rangle\\ \text{SEM} = [e][\text{enjoy}(e,x,e') \wedge \boxed{\text{P}}\,(e',x,y) \wedge \boxed{\text{m}}\,\text{book}(y)]\end{array}\right]$$

Figure 4: The sign for *enjoy* instantiated with information from the NP for *the book* (ignoring tense and the determiner)

are embedded in an operator * in the DRS conditions, and this will affect their truth conditional status. So the logical form of (1a) derived via PDU is (1a$'$):

(1) a. John enjoyed the book.

a$'$.
$$\begin{array}{|l|}\hline e, e', x, y \\ \hline \textit{john}(x) \\ \textit{enjoy}(e, x, e') \\ \textit{book}(y) \\ \textit{act-on-pred}(e', x, y) \\ *\textit{read}(e', x, y) \\ \hline \end{array}$$

For brevity, we have omitted WFFs of the form $*\phi$ when ϕ also holds.

We now have the task of assigning a semantics to DRS-conditions of the form $*\phi$. This must indicate that they're derived via defaults in the lexicon. PDU is formalised in a conditional logic. So the way defaults behave in PDU is determined by constraints on a function $*_{pdu}$ that's part of the model, and which takes worlds and propositions to propositions. $*_{pdu}$ represents assumptions about the behaviour of defaults in the lexicon: $*_{pdu}(w, p)$ encodes what according to w, normally follows from p. So, let K be DRS, and let K^- be the DRS K with all the DRS-conditions of the from $*\psi$ removed. Then we can define the semantics of $*\phi$ as follows:

- $M, w \models_f *\phi$ in DRS K just in case for all w' in $*_{pdu}(w, [\![K^-]\!])$, there is a $g \supseteq f$ such that $M, w' \models_g \phi$.

DRS conditions of the form $*\phi$ aren't asserted to be true in the actual world w, since according to the assumptions about $*_{pdu}$ in PDU, it's not necessarily the case that $w \in *_{pdu}(w, p)$. So in (1a$'$), the logical semantics *doesn't* entail that the event that was enjoyed was a reading; however, it does entail that an event was enjoyed by John. Thus we have utilised the fact that defaults persist, by assigning the default results of PDU a different truth conditional status in logical semantics from the indefeasible results. It is now up to the pragmatic component to assess whether *read* should be inferred as the appropriate event in the discourse context. The lexicon has proposed this, but clues from the more open ended pragmatic reasoning may dispose of this proposal, and replace it with another. We'll come to this in the next section.

Copestake and Briscoe (1995) treat *fast* in a very similar way to *enjoy*. The coindexation between the telic role of the object NP in the subcat list and the event that *fast* predicates over in the semantics is inherited via PDU from a lexical generalisation over the class of adjectives of which *fast* is a member (other members are *slow*, *careful*, *long*). In this case the telic role of *typist* is $[x][/\textit{type}(e, x)]$, where x is coindexed with the 'normal' variable. But this is defeasible: it's on the RHS of the slash. The truth conditional effects of this is are represented in the DRS (18) for *fast typist*, where the formula $\textit{type}(e, x)$ is within the scope of $*$:

(18) $\begin{array}{|l|}\hline x, e \\ \hline typist(x) \\ fast(e) \\ act\text{-}pred(e, x) \\ *type(e, x) \\ \hline \end{array}$

So the lexicon proposes that the event *fast* that predicates over is *type*, but this may be overridden by pragmatic information.

4 Linking The Lexicon to Pragmatics

4.1 DICE

We'll link the lexicon and grammar to a theory of pragmatics: specifically DICE (Discourse in Commonsense Entailment, Lascarides and Asher 1991, 1993). This is a model of discourse interpretation which encodes real world knowledge like *goats don't read*, and more generally, it encodes background information that's used to compute the rhetorical links between segments of discourse. The representation of discourse structure produced by DICE are segmented DRSs (SDRSs) (Asher 1993). An SDRS is a recursively defined structure which connects DRSs together using discourse relations like *Elaboration*, *Contrast* and so on. These relations impose coherence constraints on the discourse. The details of these are in Asher 1993, Asher and Lascarides 1995, Lascarides and Asher 1993. We'll exploit these constraints to reason about when lexical defaults should be overridden. Simply put: lexical defaults will normally be overridden when they lead to a bad discourse.

DICE uses the default logic Commonsense Entailment (CE) (Asher and Morreau, 1991) to reason about pragmatic interpretation. This logic exploits conditions of the form: $A > B$, which means *If A then normally B*. So one could represent *goats don't read* as the schema:

- `Goats Don't Read:` $goat(x) > \neg read(e, x, y)$

Furthermore, all default rules for computing the rhetorical relation that connects DRSs together are of the form given in (19). Here $\langle \tau, \alpha, \beta \rangle$ is the update function which can be glossed "β is to be attached to α with a discourse relation, where α is part of the discourse structure τ built so far". "Some stuff" stands for syntactic and semantic information about τ, α and β and R is a particular discourse relation:

(19) $(\langle \tau, \alpha, \beta \rangle \wedge \text{ some stuff}) > R(\alpha, \beta)$

Details of these discourse attachment rules appear in Lascarides and Asher (1991, 1993) and Asher and Lascarides (1995).

The nonmonotonic validity of CE ($\approx\!\!|$) has several nice properties. There are three that are relevant here. First, it validates DMP: if one default applies and its consequent is consistent with the KB, then it's nonmonotonically inferred. Second, it validates the Specificity Principle: if conflicting defaults have their antecedents verified, then the consequent of the default with the most specific antecedent is preferred. Finally, for each deduction $A \approx\!\!\!| B$ there is a corresponding embedded default in the object language (that is, a formula in which one $>$ occurs within the scope of another) which links boolean combinations of the formulae A and B, and which is verified to be true. We gloss this embedded default formula as $\sqsupset(A, B)$. So $\sqsupset(A, B)$ means $A \approx\!\!\!| B$. This amounts to a weak deduction theorem. The object language formula $\sqsupset(A, B)$ means that A nonmonotonically yields B in the metalanguage.

4.2 Linking PDU and DICE

To link the PDU treatment of lexical productivity to pragmatic knowledge, we add two axioms to DICE. First, Defaults Survive captures the intuition that defaults in the lexicon normally survive at the discourse level:

- `Defaults Survive:` $*\phi > \phi$

Second, we need an axiom that ensures that when the consequents of discourse processing and lexical processing conflict, the discourse processing wins. This is what happens in (20), for example, where the PDU prediction, that the event enjoyed was a reading, is overridden by the conflicting pragmatic information stipulated in the $>$-rule Goats Don't Read.

(20) The goat enjoyed the book.

Let KB_h be obtained from the knowledge base KB, by removing all the DRS conditions of the form $*\phi$ (h stands for "hard information"). Then Discourse Wins states: when this KB yields a nonmonotonic conclusion ψ, then normally this survives the KB with conditions like $*\phi$ added to it:

- `Discourse Wins:` $(*\phi \wedge \sqsupset(KB_h, \psi)) > \psi$

This rule is called Discourse Wins, because by the Specificity Principle with Defaults Survive, if ψ conflicts with ϕ, then ψ is nonmonotonically inferred and ϕ is not, even if $*\phi$ was in the KB. In other words, the clues from discourse context, if there are any, override conflicting results of PDU. On the other hand, if ϕ and ψ are compatible, they will both be inferred by DMP. So Discourse Wins also serves to model how discourse information can further refine the information about meaning obtained from the lexicon.

Let's now investigate how this affects the interpretation of the above examples. First, consider (1a), whose logical form expressed in DRT is (1a$'$):

(1) a. John enjoyed the book.

a$'$.

e, e', x, y
john(x) *enjoy*(e, x, e') *book*(y) *act-on-pred*(e', x, y) $*$*read*(e', x, y)

There are no $>$-rules which give information about the kinds of things that John enjoys. Moreover, Defaults Survive applies with the following instantiation of the schema: $*read(e',x,y) > read(e',x,y)$. So by DMP on this rule, one infers that John enjoyed reading the book.

Now compare this with (20), whose logical form is similar to (1a$'$):

(20) The goat enjoyed the book.

(20$'$)
$$\begin{array}{|l|}\hline e, e', x, y \\ \hline goat(x) \\ enjoy(e,x,e') \\ book(y) \\ act\text{-}on\text{-}pred(e',x,y) \\ *read(e',x,y) \\ \hline \end{array}$$

First consider the nonmonotonic consequences on KB_h. Goat's Don't Read applies, but Defaults Survive doesn't with respect to KB_h, because KB_h contains no conditions of the form $*\phi$. So by DMP on Goats Don't Read, $\neg read(e',x,y)$ follows nonmonotonically from KB_h. That is, $\sqsupset(KB_h, \neg read(e',x,y))$ holds. In the KB as a whole, the instantiation of Defaults Survive given in (21) applies just as before. But in contrast to (1a), so does the instantiation of the schema Discourse Wins given in (22):

(21) $*read(e',x,y) > read(e',x,y)$

(22) $(*read(e',x,y) \wedge \sqsupset(KB_h, \neg read(e',x,y))) > \neg read(e',x,y)$.

So by the Specificity Principle on (21) and (22), $\neg read(e',x,y)$ is inferred.

4.3 Discourse Context

We would need more $>$-rules to infer that the event enjoyed is an eating in (20). But in (23), we could infer that the goat enjoyed eating the book via the rhetorical structure of the discourse and the existing DICE rules which compute that rhetorical structure (Asher and Lascarides, 1995).

(23) My goat ate the whole library. α
He really enjoyed your book. β

The relevant rules for discourse attachment, which are taken from Asher and Lascarides (1995), are given below:

- `Narration:` $\langle \tau, \alpha, \beta \rangle > Narration(\alpha, \beta)$
- `Axiom on Narration:` $\Box(Narration(\alpha,\beta) \rightarrow e_\alpha \prec e_\beta)$
- `Distinct Common Topic:`
$\Box(Narration(\alpha,\beta) \rightarrow \exists\gamma(\gamma \Downarrow \alpha \wedge \gamma \Downarrow \beta \wedge \neg(\alpha \Downarrow \beta) \wedge \neg(\beta \Downarrow \alpha)))$
- `Subtype:`
$\Box((\theta_i(e\text{-}condn_\alpha, \alpha, \gamma_1) \wedge \theta_i(e\text{-}condn_\beta, \beta, \gamma_2) \wedge$
$e\text{-}condn_\beta \sqsubseteq e\text{-}condn_\alpha \wedge \gamma_2 \sqsubset \gamma_1) \rightarrow Subtype(\alpha,\beta))$

- `Elaboration:` $\langle\tau, \alpha, \beta\rangle \wedge \mathit{Subtype}(\alpha, \beta) > \mathit{Elaboration}(\alpha, \beta)$

- `Axiom on Elaboration:` $\Box(\mathit{Elaboration}(\alpha, \beta) \rightarrow \alpha \Downarrow \beta)$

Narration together with its Axiom capture the intuition that normally the textual order of events match their temporal order. Distinct Common Topic acts as a coherence constraint: it stipulates that a narrative must have a distinct common topic ($\gamma \Downarrow \alpha$ means γ is a topic for α). Elaboration states that if β is to be attached to α and β is a subtype of α, then normally $\mathit{Elaboration}(\alpha, \beta)$ holds; its Axiom says that α must be a topic of β. $\mathit{Subtype}(\alpha, \beta)$ can be inferred via the monotonic rule Subtype, which states: if (a) the DRSs γ_1 and γ_2 respectively identify the thematic role θ_i in α and β, with respect to the event conditions $e\text{-}condn_\alpha$ and $e\text{-}condn_\beta$, and (b) $e\text{-}condn_\beta$ is a subtype of $e\text{-}condn_\alpha$ (for example, *The goat ate* x is a subtype of *The goat enjoyed eating* x), and (c) γ_2 is part of γ_1 (e.g., *book* is part of what makes up *library*), then (d) β is a subtype of α.

Consider how these rules apply in (9a). The DRS β representing the second sentence in (9a) must be attached to the DRS α representing the first. The anaphor *he* must be identified with an accessible antecedent, and the SDRT constraints on accessibility restrict this to being the goat. Now, if the metonymy in β is resolved to *enjoy reading a book*, then the only candidate discourse relation according to the above rules for discourse attachment is *Narration*. By Distinct Common Topic, this relation requires a distinct common topic, which in SDRT is obtained by generalising the propositions in the narrative to produce a single predicate argument structure. But the resulting topic is too general to really establish anything better than weak coherence: it's something like *the goat did things*. Indeed, resolving metonymy to anything of the form *enjoy VP-ing your book*, where VP is not related to eating, results in a similar interpretation of the discourse i.e., the coherence is weak.

In DICE, pragmatic interpretations of sentences that lead to weak discourse coherence are avoided if possible, via the Interpretation Constraint below (Lascarides *et al* 1995):

- `Interpretation Constraint`
  (a) $(\langle\tau, \alpha, \beta\rangle \wedge \mathit{Info}(\alpha, \beta)$
  (b) $\sqsupset_{KB}(\beta', \mathit{weak}(\tau \cup \beta)))$
  (c) $> \neg\beta'$

In this schema, $\mathit{Info}(\alpha, \beta)$ is a gloss for all monotonic information about α and β, and $\sqsupset_{KB}(A, B)$ means $\sqsupset(KB \wedge A, B)$ and $\neg(KB, B)$ (that is, B nonmonotonically follows from the KB augmented with A but not from the KB alone). So in words, the Interpretation Constraint states that if (a) β is to be connected to α with a discourse relation, and β and α are both true, and (b) if the KB that includes not only the update task of β to α, but also the information β' nonmonotonically leads to a discourse of only weak coherence or no coherence at all, then normally (c) β' doesn't hold. This rule applies to (23) whenever β' is an assumption that the metonymy in β is resolved to an event that's unrelated to eating, because as we've stated, such an assumption produces a weak narrative. However, the Interpretation Constraint doesn't apply if the metonymy is resolved to an event which is related to eating. This is because in this case, the event condition of eating in α is a subtype of the event condition of enjoy eating in β, and the book in β is taken to be a part of the library in α. So Subtype applies in the monotonic component, thereby yielding $Elaboration(\alpha, \beta)$ in the nonmonotonic component. So there's no need for a distinct common topic between α and β anymore: *Elaboration* dictates that α is the topic of the discourse.

Consequently, DMP on the Interpretation Constraint rules out all resolutions of metonymy apart from *eat*, and so KB_h yields a nonmonotonic conclusion that $eat(e', x, y)$ holds. Therefore at the discourse level, the following rules apply, and conflict (assuming e' can't be both a reading and eating):

(24) $*read(e', x, y) > read(e', x, y)$
$(*read(e', x, y) \wedge \sqsupset(KB_h, eat(e', x, y))) > eat(e', x, y)$

So by the Specificity Principle, $eat(e', x, y)$ is inferred. This leads to the nonmonotonic conclusion that $Elaboration(\alpha, \beta)$ holds via Subtype and Elaboration.

These examples provide further motivation for conventionalising some aspects of metonymy. For suppose we were to compute metonymy *solely* within pragmatics. Then we would need to replace the information in Figures 2 and 3 with $>$-rules in DICE. Such a strategy is technically possible, but representation of pragmatic information would be trickier. For example, to interpret (20) correctly, the real world knowledge that goats don't read must win over the $>$-rules concerning generalisations about *enjoy* on telic roles. This means that the antecedent of this rule would have to be more specific, otherwise the logic won't resolve the conflict in the right way. Indeed, there is currently no logic for nonmonotonic reasoning which resolves conflict between unrelated default rules without assuming prioritorisation mechanisms that are extraneous to the logic itself. So Goats Don't Read would have to be replaced with something like (26), so that it could compete with the $>$-rule (25) which replaces the information in Figures 2 and 3 relevant to *enjoy the book*:

(25) $(enjoy(e, x, e') \wedge literature(y)) > read(e', x, y)$

(26) $(enjoy(e, x, e') \wedge goat(x) \wedge literature(y)) > \neg read(e', x, y)$

This rule is self-evidently extremely specific, but a rule of this form is required for Specificity to hold. But by spreading the load between pragmatics and the lexicon, and having communication links between them, we can 'loosen up' how we represent information.

Now consider (14), where *fast typist* means *typist who runs fast.*

(14) a. All the office personnel took part in the company sports day last week.

b. One of the typists was a good athlete, but the other was struggling to finish the courses.

c. The fast typist came first in the 100m.

compound-noun $<$ **binary-rule**

$$\begin{bmatrix}\textbf{lex-noun}\\ \text{ORTH} = \boxed{1}, \boxed{2}\\ \text{SYN} = \textbf{noun-cat}\\ \text{SEM} = \boxed{3} \wedge \boxed{5} \wedge \textbf{pred}(\boxed{x}, \boxed{y})\\ \text{QUALIA} = \boxed{7}\ \textbf{nomqualia}\end{bmatrix} \rightarrow \begin{bmatrix}\textbf{lex-noun}\\ \text{ORTH} = \boxed{1}\\ \text{SYN} = \textbf{noun-cat}\\ \text{SEM} = \boxed{3}\ \mathbf{P}(\boxed{y})\\ \text{QUALIA} = \textbf{nomqualia}\end{bmatrix}, \begin{bmatrix}\textbf{lex-noun}\\ \text{ORTH} = \boxed{2}\\ \text{SYN} = \textbf{noun-cat}\\ \text{SEM} = \boxed{5}\ \mathbf{Q}(\boxed{x})\\ \text{QUALIA} = \boxed{7}\end{bmatrix}$$

Figure 5: General schema for endocentric noun-noun compounds

made-of-substance-schema $<$ **compound-noun**

$$\begin{bmatrix}\textbf{lex-count-noun}\\ \text{SEM} = \boxed{3} \wedge \boxed{5} \wedge \textbf{pred/made-of-substance}(\boxed{x}, \boxed{y})\\ \text{QUALIA} = \textbf{artefact}\end{bmatrix} \rightarrow$$

$$\begin{bmatrix}\textbf{lex-uncount-noun}\\ \text{SEM} = \boxed{3}\ \mathbf{P}(\boxed{y})\\ \text{QUALIA} = \textbf{substance}\end{bmatrix}, \begin{bmatrix}\textbf{lex-count-noun}\\ \text{SEM} = \boxed{5}\ \mathbf{Q}(\boxed{x})\\ \text{QUALIA} = \textbf{artefact}\end{bmatrix}$$

Figure 6: A compound noun subschema

The axioms Defaults Survive and Discourse Wins capture this. In outline, the Interpretation Constraint in DICE blocks the assumption that *the fast typist* in (14c) is different from the typists mentioned in (14a,b) because this would lead to a weak discourse. Consequently, Subtype and Elaboration yield the intuitive attachment that (14c) is an *Elaboration* of (14a,b).

As we've mentioned, *the fast typist* must identify a unique typist from (14a,b). There are two typists, who have been differentiated only on the grounds of their athletic ability. So verifying the uniqueness condition is possible only if *fast* is equated with athletic ability. Thus $\sqsupset(KB_h, \mathit{fast}(e') \wedge \mathit{run}(e', x))$ holds (where $\mathit{typist}(x) \in KB_h$). So Defaults Survive and Discourse Wins both apply, and they have the consequents $\mathit{type}(e', x)$ and $\mathit{run}(e', x)$ respectively. Assuming that e' can't be both a typing and a running, these rules conflict. And so by the Specificity Principle, $\mathit{run}(e', x)$ is nonmonotonically inferred. In contrast, in 'neutral' (i.e., uninformative) discourse contexts, DMP on Defaults Survive will yield that *fast typist* means *typist who types fast*.

4.4 Compound nouns

A general schema for endocentric compound interpretation is shown in Figure 5, with an underspecified predicate, **pred**, relating the indices of the constituents. Most compounds will instantiate one or more of the subschemata which inherit from this schema with the predicate relating the parts of the compound marked as persistently default. An example of a more specific schema is shown in Figure 6. This schema defeasibly specifies that the compounding predicate is **made-of-substance**.

The structure below shows the result of instantiating the schema in Figure 6 with *wickerwork chair* (ignoring the substructure in *wickerwork*).

$$\begin{bmatrix}\textbf{lex-count-noun}\\ \text{SEM} = \textbf{wickerwork}(\boxed{4}) \wedge \textbf{chair}(\boxed{6}) \wedge \textbf{pred/made-of-substance}(\boxed{6}, \boxed{4})\\ \text{QUALIA} = \textbf{artefact}\end{bmatrix}$$

In normal contexts, this interpretation will stand. However, since the compounding predicate is defeasible, it can be pragmatically overridden along the same lines as the examples discussed above. In a context such as (16), an alternative interpretation is found, since the default interpretation is contradicted by the context:

(16) At school, everyone worked on crafts in groups round a big table, sitting on brightly coloured chairs. To make sure everyone could reach the materials, the groups used particular chairs: the wickerwork chairs were made of red plastic, for example.

The pragmatic interpretation of *were made of red plastic* blocks the inference that the *chairs* were made of wickerwork. Moreover, the discourse structure of (16)—and in particular, the line of reasoning in DICE that leads to *Elaboration*—yields a nonmonotonic inference from KB_h that *wickerwork chair* is to be interpreted as *chair which is sat on by someone who works on wickerwork*. So by the Specificity Principle on Defaults survive and Discourse Wins, the established meaning of *wickerwork chair* is overridden in (16); instead it means *chairs made of red plastic, which are sat on by people working with wickerwork*.

Briscoe *et al* (1990) claim that lexical generalisations are only cancelled in contexts that are informationally rich. We have illuminated in a formal setting exactly what this means. According to Defaults Survive and Discourse Wins, a lexical generalisation $*\phi$ can be cancelled only if $\sqsupset(KB_h, \neg\phi)$. So a discourse context is 'informationally rich' if, independently of all default lexical generalisations, there are discourse clues which enable one to nonmonotonically conclude the exception.

5 Conclusion

Many lexical generalisations are of the sort where there are exceptions to the rules, which are triggered by information outside the lexicon. This poses a challenge to monotonic accounts of the lexicon and to those which treat defaults as an abbreviatory convention and restrict their use to the description language.

Using an account of lexical organisation involving *persistent* default unification, we showed that links to a pragmatic component were possible with just two axioms: the first ensures that lexical generalisations normally apply in a discourse context, while the second ensures that normally, discourse information about how a word should be interpreted—if there is any—wins over defaults from the lexicon. This accounted for exceptions to lexical generalisations in a discourse context in two areas: logical metonymy and compound nouns. Moreover, the axioms clarified in a formal setting the claim in Briscoe *et al* (1990), that exceptions to lexical generalisations can only be triggered by discourse contexts which are informationally rich.

This is just a first step towards linking lexical and pragmatic knowledge. Much more needs to be done, to achieve a robust theory of lexical interpretation in a discourse context. Nevertheless, these first results indicate the kinds of operations that one needs in both components for them to communicate properly. In the grammar and lexicon, persistent defaults are needed, while in pragmatics, the Specificity Principle and embedded defaults are a crucial part of the account.

Acknowledgements

We are grateful to Ted Briscoe, Nicholas Asher, Dan Flickinger and Danièle Godard and to participants at SALT and the 1995 AAAI spring symposium on representation and acquisition of lexical knowledge for their helpful comments on material presented here. This work was partially supported by the ESPRIT Acquilex-II, project BR-7315, grant to Cambridge University.

References

Asher, N. (1993) *Reference to abstract objects in discourse,* Kluwer Academic Publishers.

Asher, N. and M. Morreau (1991) 'Commonsense entailment: a modal theory of non-monotonic reasoning', *Proceedings of the 12th International Joint Conference on Artificial Intelligence (*IJCAI*-91),* Sydney, pp. 387–392.

Asher, N. and A. Lascarides (1995) 'Lexical disambiguation in a discourse context', *Journal of Semantics,* **12(1)**, 69–108.

Bauer, L. (1983) *English word-formation,* Cambridge University Press, Cambridge, England.

Briscoe, E.J., A. Copestake and B. Boguraev (1990) 'Enjoy the paper: lexical semantics via lexicology', *Proceedings of the 13th International Conference on Computational Linguistics (*COLING*-90),* Helsinki, pp. 42–47.

Copestake, A. (1992) 'The representation of lexical semantic information', Doctoral dissertation, University of Sussex, Cognitive Science Research Paper CSRP 280.

Copestake, A. (1993) 'Defaults in lexical representation' in Briscoe, E.J., A. Copestake and V. de Paiva (ed.), *Inheritance, Defaults and the Lexicon,* Cambridge University Press, pp. 223–245.

Copestake, A. and E. J. Briscoe (1992) 'Lexical operations in a unification based framework' in J. Pustejovsky and S. Bergler (ed.), *Lexical semantics and knowledge representation. Proceedings of the first* SIGLEX *Workshop, Berkeley, CA,* Springer-Verlag, Berlin, pp. 101–119.

Copestake, A. and E.J. Briscoe (1995) 'Semi-productive polysemy and sense extension', *Journal of Semantics,* **12(1)**, 15–67.

Downing, P. (1977) 'On the Creation and Use of English Compound Nouns', *Language,* **53(4)**, 810–842.

Evans, R. and G. Gazdar (1989) 'Inference in DATR', *Proceedings of the 4th Conference of the European Chapter of the Association for Computational Linguistics (*EACL*-1989),* Manchester, England, pp. 66–71.

Flickinger, D. (1987) 'Lexical rules in the hierarchical lexicon', PhD thesis, Stanford University.

Godard, D. and J. Jayez (1993) 'Towards a proper treatment of coercion phenomena', *Proceedings of the Sixth Conference of the European Chapter of the Association for Computational Linguistics (*EACL*-93),* Utrecht, The Netherlands, pp. 168–177.

Hobbs, J.R., M. Stickel, D. Appelt and P. Martin (1990) 'Interpretation as abduction', Technical Note No. 499, Artificial Intelligence Centre, SRI International, Menlo Park, CA.

Kamp, H. and U. Reyle (1993) *From discourse to logic: an introduction to modeltheoretic semantics, formal logic and Discourse Representation Theory,* Kluwer Academic Publishers, Dordrecht, Germany.

Lascarides, A. and N. Asher (1991) 'Discourse relations and defeasible knowledge', *Proceedings of the 29th annual meeting of the Association for Computational Linguistics (ACL-91),* Berkeley, California, pp. 55–63.

Lascarides, A. and N. Asher (1993) 'Temporal interpretation, discourse relations and common sense entailment', *Linguistics and Philosophy,* **16.5**, 437–493.

Lascarides, A., E.J. Briscoe, N. Asher, and A. Copestake (in press) 'Persistent order independent typed default unification', *Linguistics and Philosophy*

Lascarides, A., A. Copestake and E.J. Briscoe (1995) 'Ambiguity and coherence', ACQUILEX working paper, available from the Computer Laboratory, University of Cambridge. Also submitted to *Journal of Semantics.*.

Pollard, C. and I. Sag (1994) *Head-driven phrase structure grammar,* CSLI lecture notes series, Chicago University Press, Chicago.

Pustejovsky, J. (1991) 'The generative lexicon', *Computational Linguistics,* **17(4)**, 409–441.

Pustejovsky, J. (in press) *The generative lexicon,* MIT Press, Cambridge, Mass.

Reiter, R. (1980) 'A logic for default reasoning', *Artificial Intelligence,* **13**, 81–132.

Young, M. and W. Rounds (1993) 'A logical semantics for nonmonotonic sorts', *Proceedings of the 31st annual meeting of the Association for Computational Linguistics (ACL-93),* Columbus, Ohio, pp. 209–215.

Zeevat, H., E. Klein and J. Calder (1987) 'An introduction to unification categorial grammar' in N. Haddock, E. Klein and G. Morrill (ed.), *Categorial grammar, unification grammar, and parsing: working papers in cognitive science, Vol. 1,* Centre for Cognitive Science, University of Edinburgh, pp. 195–222.