J. Linguistics **34** (1998), 387–414. Printed in the United Kingdom
© 1998 Cambridge University Press

Pragmatics and word meaning[1]

ALEX LASCARIDES

University of Edinburgh

ANN COPESTAKE

Stanford University

(Received 23 July 1996; revised 18 March 1997)

In this paper, we explore the interaction between lexical semantics and pragmatics. We argue that linguistic processing is informationally encapsulated and utilizes relatively simple 'taxonomic' lexical semantic knowledge. On this basis, defeasible lexical generalisations deliver defeasible parts of logical form. In contrast, pragmatic inference is open-ended and involves arbitrary real-world 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. Some cases of this involve regular polysemy, where words systematically have multiple senses. This covers a diverse range of phenomena including verb alternations (e.g. causative-inchoative), denominalised verbs of various types (e.g. *whistle*, *fax*), and the less well-studied noun alternations: count/mass senses of words denoting animals and their meat (*lamb*, *turkey*, *haddock*, etc.), the container/contents alternation (*box*, *case*, etc.) and so on. In this paper, however, we will be mainly concerned with more subtle cases where it is inadequate to postulate distinct word senses and where the interaction between words is crucial. One example is the phenomenon which Pustejovsky (e.g. 1991, 1995) has called logical metonymy where additional meaning seems to arise for particular verb/noun or adjective/noun combinations in a systematic way. For example, (1a) usually has the same interpretation as (1b).

[1] An earlier version of this paper appeared in the Proceedings of SALT V. We are grateful to Ted Briscoe, Nicholas Asher, Dan Flickinger, Danièle Godard, Ivan Sag 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. We have also benefited from anonymous *JL* comments. This work was partially supported by the ESPRIT Acquilex-II, project BR-7315, grant to Cambridge University, and by the ESRC grant, project number R000236052, to The University of Edinburgh.

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

(1) (a) Mary enjoyed the book.
(b) Mary enjoyed reading the book.

The apparent generalisation about *enjoy* and similar verbs is that semantically they always take eventualities. When the syntactic complement is an object-denoting NP, as in (1a), the sentence is nevertheless interpreted to mean that an event is enjoyed.

Pustejovsky argues that it is inadequate to postulate different word senses to account for logical metonymy and that the meaning of examples such as (1a) must be derived compositionally, as a result of the interaction of the semantics of the verb and the noun. The specific enjoyed event is supplied by the noun based on its lexical semantic structure. However, as we will discuss in detail below, context can affect the interpretation: (1a) will probably not be interpreted as (1b) if we know that Mary is a goat, for instance. Thus a purely lexical account of logical metonymy is inadequate. The challenge is to account for processes such as logical metonymy compositionally in a way which allows for their partly conventional nature, within a general framework of linguistic description that recognizes the role of pragmatics.

The starting point for our approach to the lexicon is the recognition that syntactic realisation and word meaning are often closely interrelated: for instance, it is not an arbitrary fact that *lamb* is a mass noun when it refers to (an unbounded quantity of) meat, and a count noun when it refers to the animal, or that *enjoy* can take a nominal complement. An account of the lexicon which does not incorporate lexical semantic information is inadequate because it misses generalisations in syntactic and morphological behaviour. There are, of course, exceptions to generalisations based on semantics and cases where grammatical behaviour has to be stipulated: because of this lexical representation must utilize a formalism which allows for defaults. But we do not want to postulate an unconstrained account of lexical semantics which involves arbitrarily complex inference and open-ended world knowledge. Because of this we take a methodological position where lexical semantic information is only postulated if it is required to account for generalisations about grammatical behaviour or if a purely pragmatic account seems untenable because an effect is (partially) conventional. We will try and make this approach more concrete in this paper, by a detailed discussion of borderline cases, such as the lexicon's contribution to logical metonymy.

We make the same basic assumptions about the lexicon as Briscoe et al. (1990) and Copestake (1992), which argue for an interaction between lexical semantics and pragmatics in which purely linguistic processing is informationally encapsulated and utilizes relatively simple 'taxonomic' lexical semantic knowledge. Lexical semantic information and real world knowledge are not seen as necessarily distinct. Instead, lexical semantic information is a strictly limited fragment of world knowledge, encapsulated in the lexicon,

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

which interacts with knowledge of language and may be partially conventional. For example, the entry for the animal sense of *goat* might locate it in a lexical semantic taxonomy under a node **animal**. It is, of course, part of real world knowledge that goats are animals, but **animal** is only included in the lexical semantic hierarchy because it is an appropriate locus for linguistically-relevant generalisations (for instance, animal-denoting nouns are generally count rather than mass). In contrast, as far as we know, there is no linguistic justification for a node **bovid**. In this account, lexical semantics is integrated with morphology, syntax and compositional semantics by utilising a uniformly unification-based approach comparable to that assumed in HPSG. The formalism is not extended for lexical semantics: for example, the lexical semantic hierarchy is encoded using types in the same manner as the syntactic hierarchy. Lexical semantics contributes to processes such as logical metonymy, allowing linguistic processing to deliver a partly defeasible logical form, which can be overridden by open-ended pragmatic reasoning.

However, the account in the earlier work 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 & 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 revised 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. This extends the formalism, but as we have argued elsewhere (Lascarides, Briscoe et al. 1996), this extension is desirable for purposes other than encoding lexical semantics. 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 modelled 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.

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

2. Generalisations with exceptions

We will begin this section by looking at logical metonymy in more detail. Traditionally the only way to handle the dual behaviour of verbs such as *enjoy* is to assume that it has two lexical entries, one of which takes a VP complement and the other an NP, 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), repeated below, is (1b), and it misses the generalisation to other cases where a noun phrase is interpreted as an event, such as those in (2).

(1) (a) Mary enjoyed the book.
(b) Mary enjoyed reading the book.

(2) (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 (3).

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

Pustejovsky (e.g. 1991) proposes that examples such as (1a) involve logical metonymy. He treats nouns as having a 'qualia structure' as part of their lexical entries which, among other things, 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 involved is determined from the qualia structure, which results in an interpretation for (1a) equivalent to (1b).

In section 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 (4a) with the logical form shown in (4b).[2]

(4) (a) Mary enjoyed some event associated with the book.
(b) $\exists y, e, e'[enjoy(e, \text{Mary}, e') \wedge act\text{-}on\text{-}pred(e', \text{Mary}, y) \wedge book(y)]$

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 as well as *read*. We assume that the noun phrase provides the specific predicate involved, via the telic role of the qualia structure, but

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

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

unlike Pustejovsky we will treat this as a default. Thus (1a) has the logical form shown in (5), where * indicates defeasibility:

(5) $\exists y, e, e'[enjoy(e, \text{Mary}, e') \wedge *read(e', \text{Mary}, y) \wedge book(y)]$

Given the basic metonymic interpretation represented in (4b), the fact that the event that was enjoyed was done by Mary to the book is indefeasible, and so it can never be overridden. But the information that the event that's enjoyed is a reading one is defeasible, and in principle this default can be overridden. We will explicate in this paper the conditions under which this happens.

Note that 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.

On our account, such sentences will not be blocked by the grammar, but will result in a logical form which contains the very general *act-on-pred*, which will be pragmatically anomalous, unless context provides a more specific interpretation.

2.1 *Lexical and non-lexical exceptions*

The reason for the use of defaults is that the generalisation about the interpretation on the basis of the telic role has two classes of exceptions. The first case comprises lexical exceptions and is exemplified by (7):

(7) Mary enjoyed the dictionary.

Although dictionaries are books, (7) is unlikely (again out of context) to have the interpretation (8) because dictionaries are usually used as reference books rather than read.[3]

(8) Mary enjoyed reading 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 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

[3] Some people, including ourselves, find (7) and similar examples less than fully acceptable: in Copestake & Briscoe (1995) it was argued that this is because *enjoy* is not fully acceptable with point-like events such as *refer to*. However we will ignore this here, since it is not relevant to our main point, and acceptability judgements differ considerably between speakers.

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

lexical structure, since this 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 goats eat 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 Reiter's (1980) consistency operator M into the part of the logical form derived from the telic role. Thus the logical form given for (1a) was (10); it can be glossed as "the event enjoyed is reading, in the absence of information to the contrary".

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

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 a 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 lexicalized (unlike *dictionary*).

(11) (a) John enjoyed the book made out of marzipan.
(b) ?John enjoyed the book with blank pages.

Intutively, these cases are just like the dictionary one, in that they arise because the object is an abnormal book. In fact, we hypothesize 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.

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

It is implausible that these problems could be resolved by adopting a purely lexical account, at least on the assumptions about the nature of the lexicon that we discussed in the introduction, since arbitrarily complex reasoning on the context could be involved in deciding that the subject can't read or that the object is unreadable.[4] This should be apparent for the examples in (11), where no straightforward compositional mechanism would allow the modifiers to cause *book* to be marked as unreadable. In such cases, open-ended inference may be required about books and the nature of reading. Even for the simpler case of non-human subjects, it is implausible that a selectional restriction will work. Specifying that the subject of *read* must be human is both too general since it fails to explain the anomaly in (12), and too specific since it rules out the acceptability of (13) in the context of a fairy story.

(12) ?The illiterate man read the book.
(13) The goat put her spectacles on and started to read the book to her kids.

Furthermore, no purely lexical strategy can cope with uses of anaphor, such as (9a) (since marking the pronoun as human would cause it to fail to bind with the antecedent), or cases of revision of interpretation in the light of subsequent information (such as being told that Mary is a goat sometime after being told that she enjoyed the book).

One alternative to a purely lexical account would be to claim that the interpretation of the event in metonymic sentences was purely pragmatic (i.e. that the logical form for (1a) was simply (4b), with the interpretation of the predicate *act-on-pred* being completely pragmatically determined). Such an approach is suggested by Hobbs et al. (1990) for metonymy in general (although logical metonymy as such is not discussed), with weighted abduction on pragmatic knowledge being used to determine the value of the underspecified predicate. Copestake (1992) argues that this is inadequate for metonymy in general, because of syntactic effects that sometimes accompany metonymy. But serious challenges to this line also exist for the treatment of

[4] Of course, one could propose that the lexicon has access to arbitrary context and that open-ended lexical inference is possible. Pustejovsky & Boguraev (1993) assume a fully general knowledge representation language is needed for lexical representation and Strigin (1995) argues for abductive inference in the lexicon. In our view, there are two main problems with such proposals, or at least with the terminology, if such processing is described as purely lexical. The first is that open-ended reasoning is not solely connected with linguistic processing: it is needed for making inferences generally. Thus if the lexicon itself has this capability, it will be necessary to duplicate information and capabilities which are also available to the non-linguistic reasoning component. The second is that a non-trivial interface would be required between the sort of formalism necessary to implement open-ended inference and syntactic representation. It thus seems to us preferable to reserve the term lexicon for the component which integrates closely with the rest of the grammar and to assume that open-ended reasoning on context is part of the function of the pragmatic component and not of the lexicon.

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

logical metonymy (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. 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 section 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 problems with prioritorising defaults. 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. In section 3 we will offer an account where the priorities on interpretation necessary to account for these examples will follow from general principles about the integration of the lexicon and pragmatics.

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, (14) 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:

(14) ?John enjoyed that doorstop.

Further support for the idea that there is some conventional aspect to logical metonymy is given by data discussed by Godard & Jayez (1993). Pustejovsky argues that for verbs like *begin*, two possible interpretations arise from the qualia structure of the noun: besides the telic interpretation there is an AGENTIVE interpretation, which corresponds to the event which is characteristically involved in the creation of an entity. We have only considered the telic interpretation so far, since sentences with *enjoy* do not usually appear to have a default agentive interpretation. However for many other verbs both telic and agentive interpretations may be possible, depending on context. There are however some restrictions on this, for example, the sentences in (15) only get the agentive reading (i.e. begin constructing) and not the telic one (begin travelling through/over/along).

(15) (a) Kim began the tunnel.
(b) Kim began the bridge.
(c) Kim began the freeway.

Consideration of comparable examples with *commencer* leads Godard & Jayez (1993) to suggest that the telic interpretation is only available for

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

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 lead to this conceptualisation, these exceptions support the hypothesis that logical metonymy is partially conventionalized.

As far as we know, no pragmatic theory which accounts for this data has been proposed, but we can sketch what it would have to include. As outlined above, there would have to be a means of providing an ordering on possible instantiations of the underspecified metonymic predicate, either by weights or defaults. Predicting the instantiation could not be done solely on the basis of world knowledge of the object denoted, but would require access to the description that was used in that utterance, as shown by examples such as (14). In order to predict the oddity of (14), the computation of the underspecified predicate would have to depend on knowledge of the usual purpose of the class of objects denoted by *doorstop*, not on real world properties of the doorstopping book itself. To account for the examples in (15), we would have to assume that the instantiation process had additional constraints which would have to be lexically or syntactically triggered, since *begin travelling through the tunnel*, etc. are perfectly acceptable. Thus a purely pragmatic account is only possible if pragmatics has access to language-specific information and once this move is made it is not clear how the account could be constrained or falsified.

Therefore we wanted to pursue the alternative hypothesis that the interface between the lexicon and pragmatics is via a partially defeasible logical form, where the nature of the metonymic event is proposed by the lexicon. Instead of the account proposed in Briscoe et al. (1990), 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 that are proposed lexically. Thus we can provide an integrated account of the interaction of lexical semantics and pragmatics. We describe this account in sections 3 and 4, but first we briefly review some other data which we believe require a similar treatment.

2.2 *Adjectives, compound nouns and null complements*

Some examples of adjective interpretation can be treated along the same broad lines as *enjoy*. Pustejovsky (e.g. 1991, 1995) and others have argued against distinct lexical entries for *fast*, for its usages in *fast car*, *fast programmer*, *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

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

artifact (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 programmer* means *a programmer who programs fast*, via the same entry for *fast*.

But as before, some discourse contexts trigger exceptions to this generalisation. In (16), *fast programmer* means *programmer who runs fast*, and not *programmer who programs fast* (cf. Pollard & Sag (1994: 330) and the interpretation of *good linguist*).

(16) (a) All the office personnel took part in the company sports day last week.
(b) One of the programmers was a good athlete, but the other was struggling to finish the courses.
(c) The fast programmer came first in the 100m.

As in the *enjoy* examples, the pragmatic component needs to know that the interpretation of *fast programmer* as *a programmer who programs 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 artifact, the compound refers to the artifact 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 (1977) gives an attested example where someone was asked to sit in the *apple juice seat* in a situation where there was a table already set with a glass of apple juice by one place. Here *apple juice seat* means "seat with a glass of apple juice in front", but obviously this meaning cannot be listed in the lexicon. Even if a compound has an established interpretation, in context there may be another possibility. In (17), taken from Bauer (1983: 86), *garbage man* can be taken to mean 'a man made out of garbage' by analogy with snowman:

(17) In the back street where I grew up, everybody was poor. We were so poor that we never went on holiday. Our only toys were the garbage cans. We never built sandcastles, only garbage men.

Examples like these 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. Downing 1977, Bauer 1983). Although it is undoubtably true that pragmatics and context play a major role in interpreting novel compounds and that there are pragmatic constraints on the possible interpretations, there are serious objections to suggesting that this is the only mechanism involved. Without further elaboration this gives no explanation of the fact that the majority of

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

compound nouns behave in a SEMI-regular manner. Some compounds which should be allowed on pragmatic grounds do not occur: for example, **blacksmith hammer* and other such compounds are not acceptable when taken as referring to an instrument used by a person with the given occupation. The possessive is used instead (*blacksmith's hammer*). Furthermore, languages vary in productivity with respect to compounds: Italian, for instance, has much more restricted compounding than English, but does not prohibit it completely. In contrast, in German compounds are formed even more freely than in English: for example, a noun-noun compound could be used rather than a possessive for the *blacksmith's hammer* example. There are other cases where literal translations from German to English, even of non-lexicalised compounds, are strange: for example, *Terminvorschlag* has to be translated as *suggestion for a date* rather than *?date suggestion*. It seems unlikely that this can be explained by any cultural or pragmatic effects. So, just as with logical metonymy, there are some linguistic constraints on compounds which must be represented.

Even if a purely pragmatic account were attempted for unestablished compounds, others, such as *garbage man*, must be explicitly listed. It is unlikely that it is the combination of denotations with the underspecified predicate that has an established interpretation, since in British English *rubbish* is the normal term rather than *garbage*, but *rubbish man* is not established in all dialects. So we must assume that the lexicon has to contain some compounds with their established meaning. But sentences containing such compounds would be ambiguous, because the corresponding productively generated underspecified compound would still have to be available. And if established compounds are listed in the lexicon, then any generalisation about the behaviour of classes of compounds should be accessible to the lexicon, since many established compounds have an interpretation that belongs to one of the standard patterns.

We therefore assume that generalisations about the interpretation of classes of compounds, such as substance/artifact compounds, are made in the grammar. However, such generalisations are only defaults. For instance, the "made-of" relationship between the nouns in compounds like *wickerwork chair* can be overridden in discourse:

(18) 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.

The last case we will consider here is that of verbs such as *drink*, which has an intransitive use that implies a narrow-scope, existential drinkee. However,

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

there is a strong preference that in the absence of information to the contrary that (19a), (20a) means (20b), and (21a) means (21b) (Fillmore 1986).

(19) (a) John drinks all the time.
(b) John drinks alcohol all the time.
(20) (a) We've already eaten.
(b) We've already eaten a meal.
(21) (a) I spent yesterday afternoon baking.
(b) I spent yesterday afternoon baking cakes or bread. (as opposed to ham or potatoes)

However, in context these preferences can be overridden by potentially aribitrary background information:

(22) The doctor thinks that John might have diabetes. He drinks all the time and excessive thirst is a symptom of diabetes.
(23) My tongue is no longer paralysed so I can eat again.
(24) As long as we're baking anyway, we may as well do the ham now too. (Silverstein, cited in Fillmore 1986)

It seems implausible to assume that pragmatics alone is responsible for the association of alcohol with the verb *drink*, in preference to other drinkable substances, or that flour-based products are pragmatically more likely to be cooked by baking than other foodstuffs. Under a pragmatic account it would also be difficult to explain why, for example, *Were you guzzling?* does not imply a meal, even though *guzzle′* implies *eat′*. We therefore assume that these default preferences for the null complements in (19a), (20a) and (21a) have been established as part of the conventional meanings of the relevant verbs. They therefore make a good candidate for the use of persistent defaults in the lexicon. The preferences for particular interpretations of the null complement can then be encoded as conventional, while ensuring that these interpretations are overridable by pragmatic information.

3. Persistent Default Unification and the lexicon

We use a typed feature structure formalism comparable to that used in HPSG (Pollard & 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 an 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

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

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. Any definition of default unification which does not distinguish non-default and default information will be order-dependent, as shown by Young & Rounds (1993) and Lascarides, Briscoe et al. (1996). This compromises the declarativity of the formalism, but is not an insuperable problem for the lexicon, because all the entries to be unified are in a fixed hierarchy and an inheritance order can be stipulated. But in a discourse situation, one cannot predict which pieces of information are to be unified, IN ADVANCE of starting the discourse interpretation process. So the interface between discourse processing and order dependent lexical processing would have to take into account the order in which the unification operations are done, which is impractical. It is therefore necessary for our account to distinguish default from non-default information, and also useful to do so, since it allows generalisations to be stated as exceptionless where this is appropriate.

Lascarides, Briscoe et al. (1996) 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 GS 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 (25) states that the value on the feature F is by default G:**a**, although the type of the FS (**t**) and the existence of the feature F are non-default.

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

When a default value survives PDU (notated $\overset{<>}{\sqcap}$), it does so with the slash annotation. The details of PDU are given in Lascarides, Briscoe et al. (1996) but two examples are given in (26).

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

(26) *Some examples of PDU*
Where $\mathbf{t}'$ is more specific than ($\sqsubset$) $\mathbf{t}$:

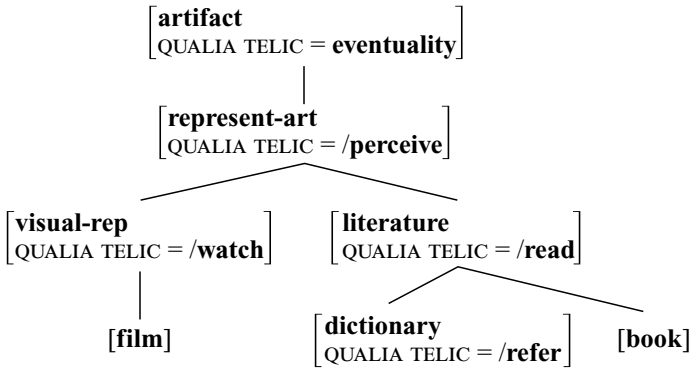
$$\begin{bmatrix}\mathbf{t}\\ \mathrm{F} = \mathbf{a}\end{bmatrix} \overset{<>}{\sqcap} \begin{bmatrix}\mathbf{t}'\\ \mathrm{F} = /\mathbf{b}\end{bmatrix} = \begin{bmatrix}\mathbf{t}'\\ \mathrm{F} = \mathbf{a}\end{bmatrix} \qquad \text{Defeat of DMP}$$

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

These indicate that PDU validates defeat of Defeasible Modus Ponens (DMP), and unlike Young & Rounds' definition, it also validates Specificity (i.e., defeasible information on more specific TDFSs overrides conflicting defaults on more general TDFSs).

Lascarides, Briscoe et al. (1996) show one way of encoding the inheritance of telic roles in PDU. This is given in (27).[5]

(27) *The telic role of artifacts*



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 the previous descriptions in Briscoe et al. (1990) and Copestake (1992), apart from the slash, but here default inheritance can proceed in any order to compute the telic roles. This use of defaults thus allows for exceptions due to lexically specified classes of 'abnormal' books, such as dictionaries: but unlike the previous account the use of persistent defaults extends this so that individual abnormal books and normal books put to abnormal uses are also allowed for, as we will see below.

We should emphasize that our concern here is not to represent the meaning of artifact-denoting nouns, but only to represent aspects of semantics which

[5] For expository purposes, the QUALIA feature is shown at the top-level of the sign, but it should actually be taken as being a part of the semantics. The use of distinct types for individual lexical entries is also a simplification which we adopt here for convenience, since the details of the feature structure geometry are irrelevant to this paper.

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

contribute to processes such as logical metonymy. A philosophically adequate theory of meaning would probably treat knowledge of the purpose of artifacts as analytic: for example, it is an essential property of books as a class that they are readable, while it is not necessary that they have a realisation as solid individual physical objects. But we make no attempt at representing such distinctions here.

(28) *The generalisation for verbs like* enjoy

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

Copestake & Briscoe (1995) show how to state the lexical generalisation concerning *enjoy*, that it predicates over the telic role of the artifact as is shown in (28).[6] When *enjoy* take a non-event denoting object (which instantiates the CAT SUBCAT 'slot'), the event that is enjoyed is instantiated via the telic role, as indicated by the coindexation $\boxed{\text{P}}$ in (28). In the figure, R is the predicate associated with the verb itself (e.g. **enjoy**) $\boxed{\text{P}}$ and $\boxed{\text{n}}$ indicate coindexation (we are using letters rather than integers here for readability). The instantiated form is shown in (29). 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.

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

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

We have shown the path QUALIA TELIC explicitly, to illustate that it is the predicate **read** which is slashed. The information that there is some metonymic event and that the predicate involved is a subtype of **act-on-pred** is not default, and cannot be overridden. Nor can one override the values of

[6] Unlike Pustejovsky (1995) 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 & Briscoe (1992, 1995) and Godard & Jayez (1993). However the differences between this and the alternative account where the NP itself undergoes coercion are largely irrelevant here.

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

the arguments in the atomic formulae. However the effect of the slash will be that of the contribution of *read* to the semantics of the sentence can be contextually overridden.

The semantic representation assumed is InL (Indexed Language, Zeevat et al. 1987), which has a direct equivalence to DRT. We use DRT here, since this is the semantic representation scheme that underlies the pragmatic component DICE (Discourse in Commonsense Entailment, Lascarides & Asher 1991, 1993; see section 4.1 below) that we link the grammar to. We assume that DRS-conditions that arise from elements on the RHS of the slash notation are embedded in an operator $*$ in the DRS conditions, and this operator receives a semantics which ensures that $*\phi$ doesn't logically entail ϕ; rather, $*\phi$ is true just in case ϕ is suggested by the lexicon as the *default*. So the logical form of (1a) repeated here as (30a), derived via PDU is (30b).

(30) (a) Mary enjoyed the book.

(b)

e, e', x, y
$mary(x)$
$enjoy(e, x, e')$
$book(y)$
$act\text{-}on\text{-}pred(e', x, y)$
$*read(e', x, y)$

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 are 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 is 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 form $*\psi$ removed. Then we can define the semantics of $*\phi$ as follows:

(31) $M, w \models_f *\phi$ 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 assserted to be true in the actual world w, since according to the assumptions about $*_{pdu}$ in PDU, it is not necessarily the case that $w \in *_{pdu}(w, p)$. So in (30b), the logical semantics DOES NOT entail that the event that was enjoyed was a reading. Rather, the formula $*read(e', x, y)$ records that the lexicon suggests this. However, (30b) does entail that an event was enjoyed by Mary.

Thus we have utilized the fact that defaults persist, by assigning the default results of PDU a different truth conditional status in logical semantics from the indefeasible results. The partial defeasibility of the logical form indicates that *read* is the best guess on the basis of lexical information for the specific

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

enjoyed event (the information arises from the lexical semantics of book, but overriding it in pragmatics does not entail any abnormality of the specific book involved, just that the book was being enjoyed in an unusual way). It is up to the pragmatic component to assess whether *read* should be inferred as the appropriate event in the discourse context. The lexicon has suggested 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.

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 & 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 is 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 rhetorical relations like *Elaboration*, *Contrast* and so on. These relations impose coherence constraints on the discourse, by imposing restrictions on the semantic relationships between the propositions being connected. The details of these are in Asher 1993, Lascarides & Asher 1993 and Asher & Lascarides 1995. In these papers, we have exploited the semantics of these relations, as specified by their coherence constraints, to model the way the truth conditional semantic content of a sentence is affected by the way it connects to the discourse context. Modelling this captures the intuition that speakers expect hearers to accommodate semantic content during discourse processing, that is additional to the compositional semantic content provided by the grammar. Indeed, modelling this accommodation of semantic content is the primary motivation for using rhetorical relations in the representation of the semantics of discourse. However, here we use the coherence constraints on rhetorical relations for the more specific purpose of reasoning 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 & 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`: $goat(x) > \neg read(e, x, y)$

Although this rule stipulates knowldge that is intuitively compelling, it is unliikely that one would want to record a rule like this directly in a practical

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

system for pragmatic reasoning, since it's too specific. However, it can be derived automatically via the axioms in CE from a more general default such as "Only humans read, normally", so long as this more general default was represented in CE in the appropriate way.[7] At any rate, we use 'Goats Don't Read' for illustrative purposes here, for investigating the circumstances when pragmatics overrides lexical defaults.

DICE also uses default rules to compute the rhetorical relations that connect DRSs together to form an SDRS. All these rules are of the form given in (26). Here $\langle \tau, \alpha, \beta \rangle$ is the update function which can be glossed "β is to be attached to α with a rhetorical 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 rhetorical relation:

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

Details of these discourse attachment rules appear in Lascarides & Asher (1991, 1993) and Asher & 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 is 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 $\beth(A, B)$. So $\beth(A, B)$ means $A |\!\approx B$. This amounts to a weak deduction theorem. The object language formula $\beth(A, B)$ means that A nonmonotonically yields B in the meta-language.

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 (33), for example, where the PDU prediction,

[7] Compare "Only humans read, normally" with the proposed selectional restriction on *read* mentioned in section 2.1. We assume this rule is DEFAULT, and it isn't conventionalized in the lexicon for the reasons given earlier.

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

that the event enjoyed was a reading, is overridden by the conflicting pragmatic information stipulated in the $>$-rule 'Goats Don't Read'.

(33) 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 (whatever the logical relation between ϕ and ψ):

- `Discourse Wins`: $(*\phi \wedge \boldsymbol{\square}(KB_h, \psi)) > \psi$

This rule is called Discourse Wins, because by the Specificity Principle with Defaults Survive, if ψ conflicts with ϕ – e.g., ψ is $\neg\phi$ – 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 or logically unrelated, 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) again, whose logical form expressed in DRT is (30b). Both are repeated as (34a, b).

(34) (a) Mary enjoyed the book.

(b)

e, e', x, y
$mary(x)$
$enjoy(e, x, e')$
$book(y)$
$act\text{-}on\text{-}pred(e' x, y)$
$*read(e', x, y)$

There are no $>$-rules which give information about the kinds of things that Mary 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 Mary enjoyed reading the book.

Now compare this with (33), whose logical form is similar to (34b). Both are presented in (35).

(35) (a) The goat enjoyed the book.

(b)

e, e', x, y
$goat(x)$
$enjoy(e, x, e')$
$book(y)$
$act\text{-}on\text{-}pre(e', x, y)$
$*read(e', x, y)$

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

First consider the nonmonotonic consequences on KB_h. 'Goats 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 (36) applies just as before. But in contrast to (34a), so does the instantiation of the schema Discourse Wins given in (37):

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

So by the Specificity Principle on (36) and (37), $\neg read(e', x, y)$ is inferred.

4.3 *Discourse context*

We would need more $>$-rules to infer that the event enjoyed as an eating in (35a). But in (38), 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 & Lascarides 1995).

(38) My goat ate the whole library.
He really enjoyed your book.

The relevant rules for discourse attachment, which are taken from Asher & 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`:
  $\Box(Narration(\alpha, \beta) \rightarrow \exists\gamma(\gamma \Downarrow \alpha \wedge \gamma \Downarrow \beta \wedge \neg(\alpha \Downarrow \beta) \wedge \neg(\beta \Downarrow \alpha)))$
- `Elaboration`: $\langle \tau, \alpha, \beta \rangle \wedge Subtype(\alpha, \beta) > Elaboration(\alpha, \beta)$
- `Axiom on Elaboration`: $\Box(Elaboration(\alpha, \beta) \rightarrow \alpha \Downarrow \beta)$

We write the formal rules here so that the logical relations between them (for example, which rules are default and which aren't, and which defaults are more specific than others) are clear. However, it is easiest to understand what they convey in words. Narration stipulates that if one is attaching β to α with a rhetorical relation, then normally, that relation is *Narration*. The Axiom on Narration and Distinct Common Topic stipulate coherence constraints in using this relation. The axiom states that when $Narration(\alpha, \beta)$ holds, then indefeasibly, α's event precedes β's (written $e_\alpha \prec e_\beta$). Distinct Common Topic stipulates that α and β must have a distinct common topic γ ($\gamma \Downarrow \alpha$ means γ is a topic for α). So, Narration together with its Axiom and Distinct Common Topic capture the intuition that normally the textual order of events match their temporal order and the propositions have a distinct common topic. Elaboration states that if β is to be attached to α and β is a

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

subtype of α, then normally *Elaboration*(α, β) holds; its Axiom says that α must be a topic of β. *Subtype*(α, β) can be inferred via rules given in Asher & Lascarides (1995), but we omit them here for the sake of simplicity. Intuitively, *Subtype*(α, β) holds just in case the event described in β is a subtype of that described in α in that the former entails the latter. For example, *the goat devoured the library* is a subtype of *the goat ate the book* and *the goat enjoyed eating the book* is also a subtype of *the goat ate the book*.

Consider how these rules apply in (9a) above. The DRS β representing the second sentence in (9a) must be attached on 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. So we can assume that β represents *the goat enjoyed your book*, whatever rhetorical relation is used to attach β to α. Therefore, the default rule 'Goats Don't Read' applies just as in the analysis of (27). But the discourse context α can also be used to provide clues about how to expand the metonymy. Suppose one were to resolve the metonymy to something of the form *enjoy* V-*ing your book*, where V is not related to eating. Then the only default rule that would apply for computing the rhetorical relation would be Narration. But *Narration*(α, β) can hold only if a distinct common topic can be found for α and β. In SDRT, this is obtained by generalising the propositions in the narrative to produce a single predicate argument structure. If the V is unrelated to eating, however, this topic will be very general: it's something like *the goat did things*. In Lascarides, Copestake & Briscoe (1996), we argued that if the reader can't compute an explanation for why a particular proposition is the topic of the discourse, then the discourse is at best coherent but weak.

In DICE, pragmatic interpretations of sentences that lead to weak discourse coherence are avoided if possible, via the Interpretation Constraint below (Lascarides, Copestake & Briscoe 1996):

(39) *Interpretation Constraint*
  (a) $(\langle \tau, \alpha, \beta \rangle \wedge \mathit{Info}(\alpha, \beta)$
  (b) $\mathbb{B}_{KB}(\beta', \mathit{weak}(\tau \cup \beta)))$
  (c) $> \neg\beta'$

In this schema, *Info*(α, β) is a gloss for all monotonic information about α and β, and $\mathbb{B}_{KB}(A, B)$ means $\mathbb{B}(KB \wedge A, B)$ and $\neg\mathbb{B}(KB, B)$ (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 rhetorical 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 (38) whenever β' is an assumption that the metonymy in β is resolved to an event that is unrelated to eating, because as we have stated,

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

such an assumption produces a weak narrative on the grounds that no explanation can be computed for why the topic of the discourse is so general. However, the Intepretation Constraint does not 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. Because of this $Subtype(\alpha, \beta)$ is true, and so Elaboration as well as Narration applies. By Specificity, $Elaboration(\alpha, \beta)$ is inferred, and 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. This change in topic improves the coherence 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):[8]

(40) $*read(e', x, y) > read(e', x, y)$
$(*read(e', x, y) \wedge \mathbf{\square}(KB_h, eat(e', x, y))) > eat(e', x, y)$

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 the nonmonotonic logic which is used to compute pragmatic inference would have to compute the relevant predicate of the event that is enjoyed, rather than checking that conventional clues about this predicate are coherent. In other words, we would need to replace the information in (27) and (28) above with $>$-rules in DICE, because this information wouldn't be conventionalized anymore, and the fact that the usual purpose of a book is to read it needs to be represented somewhere in order to interpret the metonymic construction *enjoy the book*.

Following this pragmatic strategy of encoding the information in (27) and (28) as $>$-rules is technically possible, but representation of pragmatic information will on the whole be much trickier. For example, to interpret (35a) correctly, the real world knowledge that goats don't read must win over the $>$-rules concerning generalisations about *enjoy* on telic roles, since these $>$-rules would apply when interpreting (35a), but we would not want to infer their consequents. This means that the antecedent of the rule that goats don't read would have to be more specific; otherwise the logic won't resolve the

[8] In fact, 'Goats Don't Read' applies as well, but we don't mention it here since the consequent of Discourse Wins in this case is strictly more specific than that of 'Goats Don't Read'.

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

conflict among the default rules that apply when interpreting (35a) 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 (42), so that it could compete with the $>$-rule (41) which would replace the information in (27) and (28) relevant to *enjoy the book*.

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

This rule is self-evidently extremely specific, but a rule of this form is required for Specificity to hold, and this inference pattern is required to obtain the right interpretation of (35a).

Instead of following this pragmatic strategy, we have spread the load between pragmatics and the lexicon, and we have encoded communication links between them. By doing this, we can 'loosen up' how we represent information. We can ensure that REGARDLESS of how the pragmatic information is represented relative to the lexical information – in other words, regardless of whether the pragmatic rules that apply are more specific than the relevant lexical rules, and regardless of whether they're not related at all – the pragmatic rules will ALWAYS win over conflicting lexical clues. This means the relevant rule for representing 'Goats Don't Read' can have a very general antecedent, and yet we guarantee that it will always win over conflicting lexical information, such as that given in (27) and (28).

4.4 *Adjectives, compounds and null complements revisited*

Having discussed the case of *enjoy* in some detail, we'll now revisit the other examples given in section 2.2 more briefly.

Copestake & 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 *programmer* is $[x][/program(e, x)]$, 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 are represented in the DRS (34) for *fast programmer*, where the formula $program(e, x)$ is within the scope of $*$:

(43)

x, e
$programmer(x)$
$fast(e)$
$act\text{-}pred(e, x)$
$*program(e, x)$

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

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

Consider (16), repeated here as (44), where *fast programmer* means *programmer who runs fast*.

(44) (a) All the office personnel took part in the company sports day last week.
(b) One of the programmers was a good athlete, but the other was struggling to finish the courses.
(c) The fast programmer came first in the 100 m.

The axioms Defaults Survive and Discourse Wins capture this. In outline, the Interpretation Constraint in DICE blocks the assumption that *the fast programmer* in (44c) is different from the programmers mentioned in (44a, b) because this would lead to a weak discourse. Consequently, Subtype and Elaboration yield the intuitive attachment that (44c) is an *Elaboration* of (44a, b).

As we have mentioned, *the fast programmer* must identify a unique programmer from (44a, b). There are two programmers, 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 $\supseteq(KB_h, fast(e') \wedge run(e', x))$ holds (where $programmer(x) \in KB_h$). So Defaults Survive and Discourse Wins both apply, and they have the consequents $program(e', x)$ and $run(e', x)$ respectively. Assuming that e' can't be both a programming and a running, these rules conflict. And so by the Specificity Principle, $run(e', x)$ is nonmonotonically inferred. In contrast, in 'neutral' (i.e., uninformative) discourse contexts, DMP on Defaults Survive will yield that *fast programmer* means *programmer who programs fast*.

Turning now to compounds, a general schema for endocentric compound interpretation is shown in (45), with an underspecified predicate, **pred**, relating the indices of the constituents.

(45) *General schema for endocentric noun-noun compounds*

compound-noun-schema < binary-rule

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

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

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

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 (46).

(46) *A compound noun subschema*

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

$$\left[\begin{array}{l}\textbf{lex-count-noun}\\ \text{SEM} = \boxed{3} \wedge \boxed{5} \wedge \textbf{pred/made-of-substance}\ (\boxed{x}, \boxed{y})\\ \text{QUALIA} = \textbf{artifact}\end{array}\right] \rightarrow$$

$$\left[\begin{array}{l}\textbf{lex-uncount-noun}\\ \text{SEM} = \boxed{3}\,\textbf{P}\ (\boxed{y})\\ \text{QUALIA} = \textbf{substance}\end{array}\right], \quad \left[\begin{array}{l}\textbf{lex-count-noun}\\ \text{SEM} = \boxed{5}\,\textbf{Q}\ (\boxed{x})\\ \text{QUALIA} = \textbf{artifact}\end{array}\right]$$

This schema defeasibly specifies that the compounding predicate is **made-of-substance**.

The structure in (47) shows the result of instantiating the schema in (46) with *wickerwork chair* (ignoring the substructure in *wickerwork*).

(47)

$$\left[\begin{array}{l}\textbf{lex-count-noun}\\ \text{SEM} = \textbf{wickerwork}(\boxed{4}) \wedge \textbf{chair}(\boxed{6}) \wedge \textbf{pred/made-of-substance}\ (\boxed{6}, \boxed{4})\\ \text{QUALIA} = \textbf{artifact}\end{array}\right]$$

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 (18), an alternative interpretation is found, since the default interpretation is contradicted by the context. (18) is repeated in (48).

(48) 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 (48) – 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 (48);

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

instead it means *chairs made of red plastic*, *which are sat on by people working with wickerwork*. In contrast, for the novel use (49), it is more plausible to assume that the interpretation suggested by the grammar is completely underspecified.

(49) Please sit in the apple juice seat.

The discourse information serves to instantiate this, in order to make an interpretation possible in this context. In other contexts, where it was not apparent that the seat could be distinguished in this way, interpretation would fail.

Finally, let us briefly discuss the treatment of null complements in this framework. We suggested in section 2.2 that one could use persistent defaults to encode the preferences for the interpretation of 'null complements' when *eat*, *drink* and *bake* are used intransitively. Such a treatment would produce the representation (50b) or (50a) (again, ignoring temporal information and making the simplifying assumption that the adverbial *all the time* can be interpreted as *always*):

(50) (a) John drinks all the time.

(b)

x, y, e
$john(x)$
$drink(e, x, y)$
$always(e)$
$*alcohol(y)$

In the absence of any discourse context, DMP on Defaults Survive will yield an interpretation where John drinks alcohol all the time. However, the interpretation of (50a) in the context given in (51) will be different (cf. (22) above).

(51) The doctor thinks that John might have diabetes. He drinks all the time.

Assuming that one knows that a symptom of diabetes is a continual thirst, background knowledge in this case supports as *Explanation* relation between the constituents, so long as the second sentence is interpreted as *John drinks fluids all the time*. So discourse information will override the lexical default in a similar manner as for previous examples.

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 $\mathbf{\Box}(KB_h, \neg\phi)$. 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.

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

5. Conclusion

Many lexical generalisations have exceptions, 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. 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.

REFERENCES

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

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

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

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

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

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

Copestake, A. (1993). Default in lexical representation. In Briscoe, E. J., Copestake, A. & de Paiva, V. (ed.), *Inheritance, defaults and the lexicon*. Cambridge: Cambridge University Press. 223–245.

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

Copestake, A. & Briscoe, E. J. (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. & Gazdar, G. (1989). Inference in DATR. *Proceedings of the 4th Conference of the European Chapter of the Association for Computational Linguistics* (EACL-*1989*). Manchester. 66–71.

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press

Fillmore, C. J. (1986). Pragmatically controlled zero anaphora. *Berkeley Linguistics Society* **12**. 95–107.
Flickinger, D. (1987). *Lexical rules in the hierarchical lexicon*. Ph.D. dissertation, Stanford University.
Godard, D. & Jayez, J. (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. 168–177.
Hobbs, J. R., Stickel, M., Appelt, D. & Martin, P. (1990). Interpretation as abduction. Technical Note No. 499. Menlo Park, CA: Artificial Intelligence Centre, SRI International.
Kamp, H. & Reyle, U. (1993). *From discourse to logic: an introduction to modeltheoretic semantics, formal logic and Discourse Representation Theory*. Dordrecht: Kluwer Academic Publishers.
Lascarides, A. & Asher, N. (1991). Discourse relations and defeasible knowledge. *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics* (ACL-*91*). Berkeley, CA. 55–63.
Lascarides, A. & Asher, N. (1993). Temporal interpretation, discourse relations and common sense entailment. *Linguistics and Philosophy* **16.5**. 437–493.
Lascarides, A., Briscoe, E. J., Asher, N. & Copestake, A. (1996). Persistent order independent typed default unification. *Linguistics and Philosophy* **19.1**. 1–89.
Lascarides, A., Copestake, A. & Briscoe, E. J. (1996). Ambiguity and coherence. *Journal of Semantics* **13.1**. 41–65.
Moens, M. & Steedman, M. (1988). Temporal ontology and temporal reference. *Computational Linguistics* **14**. 15–28.
Pollard, C. & Sag, I. A. (1994). *Head-driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
Pustejovsky, J. (1991). The generative lexicon. *Computational Linguistics* **17. 4**. 409–441.
Pustejovsky, J. (1995). *The generative lexicon*. Cambridge, MA: MIT Press.
Pustejovsky, J. & Boguraev, N. (1993). Lexical knowledge representation and natural language processing. *Artificial Intelligence* **63**. 193–223.
Reiter, R. (1980). A logic for default reasoning. *Artificial Intelligence* **13**. 81–132.
Strigin, A. (1996). Abductive inference during update: the German preposition *mit*. In Simons, M. & Galloway, T. (ed.), *Semantics and linguistic theory* (*SALT*) *V*. Ithaca, NY: Cornell University. 310–327.
Young, M. & Rounds, W. (1993). A logical semantics for nonmonotonic sorts. *Proceedings of the 31st Annual Meeting of the Association for Computational Linguistics* (ACL-*93*). Columbus, OH. 209–215.
Zeevat, H., Klein, E. & Calder, J. (1987). An introduction to unification categorial grammar. In Haddock, N., Klein, E. & Morrill, G. (ed.), *Categorical grammar, unification grammar, and parsing*. *Working Papers in Cognitive Science* **1**. Centre for Cognitive Science, University of Edinburgh. 195–222.

Authors' addresses: *(Lascarides)*
Centre for Cognitive Science and HCRC,
University of Edinburgh,
2, Buccleuch Place,
Edinburgh EH8 9LW,
Scotland, UK.
E-mail: alex@cogsci.ed.ac.uk

(Copestake)
CSLI,
Stanford University,
Cordura Hall,
Stanford, CA 94305,
U.S.A.
E-mail: aac@csli.stanford.edu

https://doi.org/10.1017/S0022226798007087 Published online by Cambridge University Press