A compositional treatment of polysemous arguments in Categorial Grammar

Anne-Marie Mineur* Paul Buitelaar†

(mineur@coli.uni-sb.de ; paulb@zag.cs.brandeis.edu)

Abstract

We discuss an extension of the standard logical rules (functional application and abstraction) in Categorial Grammar (CG), in order to deal with some specific cases of polysemy. We borrow from Generative Lexicon theory which proposes the mechanism of *coercion*, next to a rich nominal lexical semantic structure called *qualia structure*.

In a previous paper we introduced coercion into the framework of *sign-based* Categorial Grammar and investigated its impact on traditional Fregean compositionality. In this paper we will elaborate on this idea, mostly working towards the introduction of a new semantic dimension. Where in current versions of sign-based Categorial Grammar only two representations are derived: a prosodic one (form) and a logical one (modelling), here we introduce also a more detaled representation of the lexical semantics. This extra knowledge will serve to account for linguistic phenomena like *metonymy*.

1 Introduction

Categorial Grammar is a compositional theory. The representation of the meaning of a whole is a function of the representation of the meaning of its parts. Since Categorial Grammar has only been interested in modelling predicate-argument structure, scope and the like, a simple representation of the nominal lexical semantics was sufficient in most cases[1]. As observed by several researchers however, such a basic Fregean concern with composition is not enough for a cognitive or computational system (Pustejovsky 1995; Sowa 1992). Frequently there is additional information in between the composing parts, which is inferred by heuristic reasoning, e.g., by *metonymy* as in (1), which has to be interpreted as 'John beginning to do something with/to a novel', probably 'reading' or 'writing' it (Pustejovsky 1995).

*Computational Linguistics, University of the Saarland, Postfach 1150, 66041 Saarbrücken, Germany

†Computer Science, Brandeis University, Waltham, MA 02254, USA

[1] Although some of the work on *meaning postulates* developed a more elaborate lexical semantics (Dowty 1979).

(1) John began a novel.

To account for this kind of inferences in Categorial Grammar some additions have to be made to its present way of representing word meaning. We borrow from Generative Lexicon theory (Pustejovsky 1995) in defining the multiple meanings of a word by the use of a polymorphic representation. Instead of giving a traditional enumeration of word senses, we now relate them to one another into one coherent structure (*qualia structure*). This also allows for a generation of senses in context, that is in a dynamic fashion by means of *coercion*. We see this in an object-oriented way, where objects derive their meaning from the procedures that operate on them. Each procedure addresses a different aspect or quality (*'qual'*) of an object.

We want to stress that we are not talking about syntactic polymorphism, like a noun that is ambiguous between nominative and accusative (Pullum and Zwicky 1986; Ingria 1990; Bayer 1994). In this paper we are interested in how the meaning of one word influences that of another and thus how semantics is built up by association while still maintaining compositionality. In fact what we are saying, is that there are no such things as uninterpretable, syntactically well-formed sentences. To our minds the sentence 'Colourless green ideas sleep furiously', that Chomsky (1957) claimed proved the autonomy of syntax, only proves that people assign an interpretation to almost any syntactically well-formed sentence—and that they have a very flexible imagination.

In short, form (morphology) and structure (syntax) serve meaning, in the sense that they indicate to the hearer how to interpret the concepts that are communicated, and in what way those are related. The less well-formed a sentence, the less clear it is which relation exists between which objects, and the more complicated to fit subcategorization and argument. Once it is clear which argument belongs to which subcategorization, the limits to coercion seem very flexible.

From a computational angle however certain limits have to be dealt with. Currently it does not seem feasible to compute whatever interpretation for any given syntactically well-formed sentence. This is why computational linguistic theories like Categorial Grammar seem rather to ignore semantic anomalies as they occur in sentences like (1). In this paper however we try to deal with them in some computationally acceptable way, while at the same time acknowledging the fact that not every possible interpretation can or even should be computed. If one assumes a semantic network for a given language in which words are ultimately (through recursive links) related to every other word in the language, it must be obvious that not every interpretation to be obtained from this network should be computed. Therefore we take the simple but effective heuristic of computing only one level deep. Only the semantic elements to be found in the immediate qualia structure of a word are to be considered for interpretation, which takes the form of a selection of the right interpretation(s) from this lexical semantic structure. The compositional process selects the appropriate semantic element(s) (*sorts*) out of the qualia structure, which is represented as a *set*. Such an approach reinterprets Pustejovsky's *coercion* as a form of type *selection* rather than type *shifting*, which makes it also a monotonic operation[2].

The reason for choosing Categorial Grammar as the grammar formalism is rather arbitrary, although not entirely. Both Categorial Grammar and Generative Lexicon theory are descendants from Montague Grammar, and share a number of assumptions (lexicalism, compositionality). The multi-level approach that is pursued nowadays in sign-based Categorial Grammar facilitates adding an extra autonomous level, and constitutes an interesting extension to Categorial Grammar.

2 Compositionality

2.1 Functors and Arguments

In Categorial Grammar one defines a category in terms of its domain and its yield. An intransitive verb requires a noun phrase to yield a sentence; a transitive verb requires a noun phrase to yield an intransitive verb; an article requires a noun to yield a noun phrase. Even complex arguments are possible—take for instance a modifier to a VP (i.e., an NP\S): it requires an NP\S to yield an NP\S.

(2)	Intransitive verb:	NP\S
	Article:	NP/N
	Transitive verb:	(NP\S)/NP
	VPmodifier:	(NP\S)/(NP\S)

Directionality is indicated by the direction of the slash; an X/Y ('X *over* Y') seeks a Y to its right to yield an X, whereas a Y\X ('Y *under* X') seeks a Y to its left to yield an X.

The basic categories we use are N (for noun), NP (for noun phrase) and S (for sentence). All basic categories are categories. Complex categories are of the form X/Y or X\Y, where X and Y must also be categories.

This categorial approach has two immediate effects, first that all the combinatorial information is transferred to the lexicon, so that every entry has

[2] The observation that in such an approach coercion becomes a monotonic operation is due to Hans Uszkoreit (personal communication).

all the information it needs (*lexicalization*), secondly that the set of rules is reduced to two rules, that do nothing more than describe the behavior of the connectives / ('slash') and \ ('backslash'):

(3) $\mathrm{X/Y, Y} \rightarrow \mathrm{X}$
$\mathrm{Y, Y\backslash X} \rightarrow \mathrm{X}$

2.2 Logical Representation

It follows naturally to extend this compositional approach of constituent structure with a logical representation. Since in Categorial Grammar the choice to call something a functor is usually based both on content and on form, we find the same division as is found in the categories to be reflected in the logical representation. The result of applying a functor to an argument in terms of a logical form is the *logical representation* of the functor applied to the *logical representation* of the argument.

(4) $\langle\ (\mathrm{X/Y}) : \mathit{functor}\ \rangle,\ \langle\ \mathrm{Y} : \mathit{argument}\rangle \rightarrow \langle\ \mathrm{X} : \mathit{functor(argument)}\rangle$
$\langle\ \mathrm{Y} : \mathit{argument}\ \rangle,\ \langle\ (\mathrm{Y\backslash X}) : \mathit{functor}\ \rangle \rightarrow \langle\ \mathrm{X} : \mathit{functor(argument)}\rangle$

As we use these rules recursively, and as each of the lexical entries carries its own bit of information, the logical representation of the whole is composed *exclusively* of the logical representation of the subparts. If we follow the traditional Montagovian way of representing meaning, then the lexical entries for 'explained, 'a speaker' and 'an example' would be:

(5)

explained	$\rightsquigarrow$	$\mathrm{(NP\backslash S)/NP}$	$: \lambda qp.explain(p, q)$
an example	$\rightsquigarrow$	NP	$: \lambda R.\exists y.example(y) \wedge R(y)$
a speaker	$\rightsquigarrow$	$\mathrm{S/(NP\backslash S)}$	$: \lambda P.\exists x.speaker(x) \wedge P(x)$

Then the representation of (6a) would be derived straightforwardly as (6b). (Albeit with the intervention of some simple meaning postulates to account for the equivalence of a term and its η-normal form, see (Montague 1974):

(6) a. A speaker explained an example.
b. $\exists x.speaker(x) \wedge \exists y.example(y) \wedge explain(x, y)$

This lambda-term can on the one hand be seen as the output of the derivation procedure (which functors applied to which arguments) and on the other hand as the input for the interpretation function (the interpretation in the model).

However, there are some shortcomings to this approach. First, (6b) can hardly be considered an exhaustive representation of (6a). It does not show that a 'speaker' is a person; nor does it represent the information that 'to explain' means that you exemplify something to an audience; nor that an 'example' is something that is chosen because of its typical properties or behaviour. Secondly, nothing prevents sentences like (7a) through (7c):

(7) a. An example explained an example
b. A speaker explained a speaker
c. An example explained a speaker

These sentences have some semantic mismatch in them, but—and this is an important observation—we still can construct some meaning, albeit with increasing difficulty. That is, if we use all the knowledge that we have available on each of these words. As was mentioned before, meaning is more than an addition of the single atomic senses.

The idea of a semantic mismatch is mirrored quite well in the way some semanticists (e.g., Pollard and Sag 1987) formulate it: 'explain' has not just syntactic expectations considering its arguments, but also expects their meaning to meet certain *selection restrictions*, see (8). The first SUBCAT NP has to be fulfill the role of the explainer, the second that of the explained.

(8) $\text{explain} \rightarrow [\ \text{SUBCAT} : \langle\ \text{NP}_{\boxed{1}\ explainer},\ \text{NP}_{\boxed{2}\ explained} \rangle]$

If we look at the representation in (6b) none of this shows. All (6b) does is indicate what the predicate is and what the arguments, and what relations there are between them.

3 Lexical Semantics

3.1 Towards a Lexical Semantics

If we want to talk about natural language not only in terms of predicate-argument schemata but also in terms of associations and the bridging of semantic gaps, then we have to consider *all* the semantic relations that exist between the words in the lexicon. We have to take into account not only verbs and their arguments, but also the lexical semantics of other categories, more in particular of nouns.

We claim that it is virtually impossible for any word to have only one strict interpretation. By their very nature words adapt to their context. The hearer as well as the speaker use their imagination and select the one aspect of the concept (relating to a particular word) that fits best to the requirements of the situation (see also (Bartsch 1987)). In short: polysemy is everywhere.

In the following sections we will discuss the systematic nature of this semantic polymorphism, as well as the way in which it interacts with context.

3.2 Polysemy and Complementary Versus Contrastive Senses

As observed by Weinreich (1964), distinctive interpretations or *senses* of words are either of a *contrastive* or a *complementary* nature. Contrastive senses (or *homonyms*) are unrelated to each other, see for instance the different meanings of the word 'bank' in (9a) and (9b):

(9) a. We walked along the bank of the Charles river
    b. Did he have an account at the HBU bank?

Complementary senses on the other hand do not contrast each other, but seem to be related in a systematic way. For instance a *brand name* like 'BMW' has at least the following complementary senses[3]:

(10) a. *the company that produces it*:
     BMW stocks gained two points yesterday.
  b. *the company building*:
     BMW takes up half this block.
  c. *a spokesperson with the company*:
     BMW announced a new model last week.
  d. *the product*:
     This year around 10,000 BMWs will be sold.
  e. *the design or production process*:
     They started a new BMW last year.

The systematic relation between complementary senses has to some extent been investigated in the literature (Weinreich 1964, Apresjan 1973, Nunberg 1979, Bierwisch 1982); most recently by Pustejovsky (1995) who termed it *logical polysemy*. Some other examples are:

INSTITUTIONS In the case of *institutions* we quote the well known 'school' example (Bierwisch 1982):

(11) a. *as a group of people*:
     The school went for an outing
  b. *as a learning process*:
     School starts at 8.30
  c. *as an institution*:
     The school was founded in 1910
  d. *as a building*:
     The school has a new roof

[3]Note that (10a) through (10c) refer to the company, hence 'BMW' is used as a proper name, where (10d) and (10e) refer to the car BMW, and hence 'BMW' is used as a common noun.

Artifact-Event In the following sentences we see a *metonymic* relation between an *artifact* ('novel') and an *event* ('reading') that typically involves that *artifact*.

(12) a. John began the novel
b. John began reading the novel

This is a general pattern, which is productive for all *artifact* denoting words. Another example is a 'model' and the *event* of 'producing' it.

(13) a. The president of BMW announced a new model
b. The president of BMW announced that they will produce a new model

Figure-Ground Pustejovsky and Anick (1988) give several examples of so-called *figure-ground* nominals that are either to be interpreted as the physical object themselves (figure) or the open space they leave behind upon removal (ground).

(14) a. *figure*:
John painted the door blue
b. *ground*:
John walked through the door quickly

Interestingly, we can also construct the following perfectly acceptable example where *figure* and *ground* interpretations occur at the same time:

(14) c. Go through the red door on your right.

That this can not always be done is shown by the following two examples with *book*, which is both of sort *information* and *physical object*—apparently *to buy* doesn't trigger the *physical object*-aspect of *book*, or not as much as *to throw* does.

(15) a. I love the book you bought me.
b. ?? I love the book you threw at me.

Animal-Food *Grinding* (see e.g., (Copestake and Briscoe 1992)) is the systematic relation between an *animal* and the *food* those animals produce after being killed.

(16) a. *the animal*:
You won't find *badgers* living around here.
b. *the food it produces*:
Badger is a delicacy in China.

3.3 Lexical Semantics and Context

Natural language is at the same time the most powerful and the most limited knowledge representation language. Most powerful, because in no artificial language can we express what we mean so vividly and full of interpretations. Most limited, because the number of possible interpretations that are somehow conveyed in natural language is just too big to yield any precision. This would mean that one is never able to point out what exactly the meaning of an expression is. However, humans are able to communicate in a sensible way most of the time. This leads us to believe that in fact there are two levels of semantic reasoning, one operating on information that can be obtained from the lexicon, and one that involves real-world knowledge. We illustrate this distinction with the following example:

(17) The newspaper that fell off the table, fired its chief editor.

Technically speaking sentence (17) could have a meaning: one might imagine a talking, acting, living newspaper that can both fall off the table and fire people[4], but given our knowledge about the physical world, it is not a very likely one. In other words: in our theory *newspaper* will be coerced to *human* and then some interpretation can be assigned to the sentence, but unless we use it in an Alice-in-Wonderland setting, real-world knowledge will rule it out.

In this paper we will only deal with the compositionally construable aspects of systematic polysemy without looking at any specific context dependent interpretation.

4 Lexical Semantic Structure

4.1 Four Levels of Interpretation

We adopt Pustejovsky's (1995) model of lexical semantic structure, which introduces four interrelated levels of interpretation:

Argument Structure Specification of number and type of logical arguments, and how they are realized syntactically[5].

[4]The proposition that we do not allow any word to have an ambiguous meaning, or, in our case to use multiple aspects of its meaning at a time, is motivated extensively by Bayer (1994). There is reason to believe that this proposition does not hold for all situations, take for instance example (15), as well as the following sentence: "The school that starts at 9 am, burnt down this morning", which uses two aspects of the word's meaning that apparently do unify.

[5]Notice that there is a difference between Pustejovsky's (1995) notion of Argument Structure, where three types of arguments are distinguished (logical, default and shadow

Event Structure Definition of what sort of event a lexical item represents. Sorts include STATE, PROCESS and TRANSITION. Events may have sub-eventual structure.

Qualia Structure Modes of explanation, composed of FORMAL, CONSTITUTIVE, TELIC and AGENTIVE roles. In more common AI terms these roles correspond to the IS-A, PART-OF/HAS-A, PURPOSE and CAUSE relations respectively.

Lexical Inheritance Structure Identification of how a lexical object is hierarchically related to other objects in a lattice that constitutes the global organization of the lexicon.

4.2 Qualia Structure

In Pustejovsky 1995 qualia structure is seen as a polymorphic representation with the different qualia roles expressing different aspects of the lexical semantic object that it represents. An example is a qualia structure for 'BMW' given in figure (18). The values of the qualia roles are meant to be defaults (the "..." are to be filled in depending on context). They reflect typical knowledge related to the concept 'BMW'.

$$(18)\quad \begin{bmatrix} \text{FORMAL} & : company(bmw) \\ \text{CONSTITUTIVE} & : \exists y.spokesperson\_of(y, bmw), \ldots \\ \text{AGENTIVE} & : \exists e.\exists x.human(x) \wedge establish(e, x, bmw), \ldots \\ \text{TELIC} & : \exists e.\exists z.produce(e, bmw, z), \ldots \end{bmatrix}$$

This structure reads as follows: If we take BMW as a company (FORMAL), then depending on what qualia role is highlighted, it can also be seen as the group of people who constitute it, one of whom is the spokesperson (CONSTITUTIVE); it entails the process of being established (AGENTIVE); and it entails its default purpose, which is the production of cars (TELIC). In other words, in this case 'BMW' is polymorphic between the predicates 'company', 'spokesperson_of', 'establish' and 'produce'.

4.3 Coercion

If qualia structure is to be seen as a polymorphic representation, then *coercion* is the generation process that produces each individual sense that is represented in the qualia structure. This notion of *coercion* is similar to the original notion of the same name in the context of object-oriented and functional programming languages (Cardelli and Wegner 1985).

arguments) and the CG notion of predicate-argument structure, which only accounts for those arguments that are actually realized—logical arguments, in Pustejovsky's terms.

Essentially, the original notion of coercion is restricted to changing the type of an object if a particular function that takes the object as input requires it. From the standpoint of natural language semantics this can be seen as some form of interpretation. The context of the object forces, *coerces* it to be interpreted differently, to take on another denotation.

In natural language interpretation this happens all the time, but unlike Pustejovsky 1995, we prefer to see this as a selection out of a set of possible interpretations rather than as some form of meaning shift. In the sentences on BMW above we saw an example of such a set of interpretations to choose from: a company, a car, a building, a collection of people and a process.

In the following section we will try to formalize this in a Categorial Grammar framework.

5 A Categorial Treatment of Polysemous Arguments

5.1 Motivation

The rules for functional application in *traditional* Categorial Grammar (Bar-Hillel 1953), when labeled with a logical representation, as given in (4) are repeated here:

(19) $\langle\ \mathrm{X/Y} : f\ \rangle, \langle\ \mathrm{Y} : a\ \rangle \rightarrow \langle\ \mathrm{X} : f(a)\ \rangle$
$\langle\ \mathrm{Y} : a\ \rangle, \langle\ \mathrm{Y\backslash X} : f\ \rangle \rightarrow \langle\ \mathrm{X} : f(a)\ \rangle$

That is, a type X can be derived from the concatenation of a functor X/Y with an argument Y on its right, or a functor Y\X with an argument Y on its left. In terms of its logical representation the functor is applied to the argument.

But now we are proposing to insert lexical semantic information (qualia structure) into these combinatorial rules. We want to do it in such a way that it does not influence the derivability of a sentence, nor the compositionality of the logical representation. Whether we want the lexical semantic information to be reflected in the logical representation is another matter. We insist on maintaining compositionality. However, since our representation of the composing parts (i.e., the lexical objects) will be more elaborate, this may effect the representation of the whole (i.e., the sentence), and it may in fact influence its denotation.

5.2 The Categorial Sign

The nucleus of the categorial sign is the category, which describes the combinatorial behaviour. To it may be attached a number of labels, which represent different dimensions (e.g., denotational semantics, prosody) (Gabbay 1991). The rules that describe the behaviour of the categories will also specify the combinatorial behaviour of the representation of the other dimensions. Properties that apply to the category need not apply to the labels: the way a sequence is proven to be derivable does not matter for the prosodics of the utterance, and for the representation of the denotation of an utterance it does not matter in which order the arguments were found. In other words: bracketing can be omitted in the prosodics dimension, order is irrelevant for predicate-argument structure.

In sign-based Categorial Grammar (Moortgat 1992, Morrill 1994), every lexical entry is represented by a sign that contains two dimensions: prosody and denotational semantics. The sign, the basic element of the framework, looks as follows:

(20) $\langle$ C : (λ,P) $\rangle$

where C = combinatorics;
λ = denotational semantics;
P = prosody

The combinatorics is the category as it is inductively built up from the basic categories NP, N and S. The basic categories will carry with them those features that can influence the combinatorial properties of the word, whether they are semantically relevant or not. Number, person, gender and case are typically combinatorially relevant properties. The denotational semantics is in fact the predicate-argument structure of the sentence; the constants that are used are only mnemonic identifiers, as we know them from the Montague-literature (Montague 1974). The prosodic dimension represents word order, and could be extended to represent constituent structure and intonation.

Denotational semantics is a standard label in the existing work on sign-based Categorial Grammar. For the sake of simplicity we will omit prosody. The categorial sign as we use it then becomes:

(21) $\langle$ C : (λ,Q) $\rangle$

where C = combinatorics;
λ = denotational semantics;
Q = qualia structure

The introduction of a qualia structure is new. The qualia structure of a phrase will consist of the information of its composing parts; it grows as more lexical objects are added. The interpretation of the lexical objects however gets more precise: context disambiguates.

5.3 The Representation of Qualia Structure

In (Buitelaar and Mineur 1994) we explored the consequences of including the qualia structure in the *combinatorics*, together with all the other features that described the category's lexical properties. In this way, we gave the qualia structure combinatorial power, and made it possible for an application to fail because of conflicting qualia structures. However, this is not consistent with what we stated above, namely that people will assign an interpretation to any syntactically well-formed sentence[6]. Chomsky's example shows that its lexical semantics will not make a sentence ungrammatical, the arguments will always adapt to their selectional restrictions—this clearly illustrates the appropriateness of the term *coercion*.

In our current proposal qualia structure is an extra semantic dimension, next to the logical form. It serves to license coercions, where otherwise a semantic mismatch would have occurred between the argument and the selection restriction of the functor for that argument.

Qualia structures are represented as sets of *sorts*[7] and are indexed by the *entity* they correspond to in the model. For instance:

(22) QS_{bmw}: $\{\mathbf{company}, \mathbf{spokesperson\_of}, \mathbf{establish}, \mathbf{produce}\}$

Note that with this representation of qualia structure we are abstracting away from the use of qualia role names like FORMAL, etc. (see figure (18). We acknowledge that subdividing the qualia set by role names adds additional structure to the lexical semantic representation, but we refrain from exploiting this for now, in order to keep the representation legible. Not adding role names, we think, does not damage the fundamental ideas behind a polymorphic lexical semantic representation.

Sorts are hierarchically organized in a lattice, where sub-sorts are subsumed by their super-sorts. In the case where a sub-sort is unified with its super-sort, the result will be the sub-sort. For instance: The sub-sorts **read** and **write** share the same super-sort **event**, and unification with **event** would result in **read** and **write**, respectively. In general:

(23) $\mathbf{x} \leq \mathbf{y} \Longrightarrow \mathbf{x} \sqcup \mathbf{y} = \mathbf{x}$.

[6] In this respect it is interesting to compare (Verschuur 1994), who explores metonymy in a many-sorted version of Categorial Grammar—using sorts instead of types as basic categories. Like in our current approach this system ties syntactically well-formedness to a flexible form of semantic interpretation.

[7] Specifications on a basic type that correspond to subsets of the set of entities that correspond to the basic type in the model.

Basic categories carry one qualia structure, while complex categories have lists of qualia structures (one for the functor and one for the argument[8]):

$$
\text{(25)} \quad
\begin{array}{ll}
\text{basic} & \langle\ \mathrm{X} : (\lambda,\ QS_a : \{q_1 \ldots q_n\})\ \rangle \\
\text{complex} & \langle\ \mathrm{X/Y} : (\lambda,\ [QS_f : \{q_1 \ldots q_m\}, QS_a : \{q_1 \ldots q_p\}])\ \rangle \text{ or} \\
 & \langle\ \mathrm{Y\backslash X} : (\lambda,\ [QS_f : \{q_1 \ldots q_m\}, QS_a : \{q_1 \ldots q_p\}])\ \rangle
\end{array}
$$

We chose to represent qualia structure as a set and not as a conjunction or disjunction. The use of conjunction would imply incorrectly that all values need to be present all the time. The use of disjunction would imply that values can be maintained, as long as one of the other values is applicable, even if they are not wanted. Clearly this is not correct also. We want to be able to make a selection, and only use those values that apply in a given context.

5.4 The Calculus

We will not go into detail about all the varieties of functional application that occur in the literature. For that we refer the reader to (Moortgat 1992) and (Morrill 1994) on Lambek-style Categorial Grammar. What we claim with respect to functional application of a functor to an adjacent argument, may be generalized to any form of adjacent or non-adjacent argument resolution.

In order to have a complete calculus, we not only need to have rules for functional application, but abstraction (or conditionalization) as well. The use of abstraction can best be demonstrated by an example: the derivation of a VP. A sequence of types that reduces to form a VP will in categorial terms be said to reduce to a sentence that lacks a noun phrase on its left (NP\S). This sequence can be reformulated in the following way. We can assume that the missing noun phrase is present as the leftmost type in the sequence and try to prove that this sequence combines to form a (complete) sentence. Like in any other equation, we can simply add a category on both sides. This category is hypothetical, and so are all the labels that are attached to it. In the denotational semantics dimension this materializes by the introduction of a lambda operator.

To represent the relation between the original sequence and the reformulated one in a rule, we need a notation that can show the derivability

[8]With the embedding of the categories, the qualia structures of complex categories will also be embedded—a transitive verb ((NP\S)/NP) will typically have a qualia structure of the form

$$
\text{(24)} \quad
\begin{bmatrix}
\mathrm{QS}_f\text{:} \begin{bmatrix} \mathrm{QS}_{f'}\text{:}\{\mathrm{x}_1 \ldots \mathrm{x}_n\}, \\ \mathrm{QS}_{a'}\text{:}\{\mathrm{x}_1 \ldots \mathrm{x}_m\} \end{bmatrix}, \\
\mathrm{QS}_a\text{:}\{\mathrm{x}_1 \ldots \mathrm{x}_p\}
\end{bmatrix}
$$

between two sequences. We also need such a notation to generalize the simple rule for application—we want to be able to capture any sequence of categories that resolves to the argument. Only in the final stage, when the categories are reduced to basic categories, will there be a check for identity between the given category and the required one—the axiom. The set of rules for the basic fragment are based on Moortgat (1988) and look as follows.

5.5 The Original Rules

Application is given in (26), where a functor X/Y (or Y\X, respectively), preceded by left context U and followed by right context V, takes some sequence of categories T on its right (or on its left, respectively) as its argument if T reduces to Y, and the resulting category X with U and V reduces to the same Z.

(26) [L/] $U, \langle X/Y{:}f \rangle, T, V \rightarrow Z$ if $T \rightarrow \langle Y{:}a \rangle$ and $U, \langle X{:}f(a) \rangle, V \rightarrow Z$ [L\\] $U, T, \langle Y\backslash X{:}f \rangle, V \rightarrow Z$ if $T \rightarrow \langle Y{:}a \rangle$ and $U, \langle X{:}f(a) \rangle, V \rightarrow Z$

Abstraction is given in (27), where a sequence of categories T reduces to some functor X/Y (or Y\X, respectively) if T followed by Y (or preceded by it, respectively) reduces to X.

(27) [R/] $T \rightarrow \langle X/Y{:}f \rangle$ if $T, \langle Y{:}a \rangle \rightarrow \langle X{:}\lambda a.f \rangle$ [R\\] $T \rightarrow \langle Y\backslash X{:}f \rangle$ if $\langle Y{:}a \rangle, T \rightarrow \langle X{:}\lambda a.f \rangle$

The axiom rule (28) shows that any basic category reduces to itself (identity).

(28) [Ax] $\langle X{:}t \rangle \rightarrow \langle X{:}t \rangle$

5.6 Qualia Structure in Categorial Grammar

Application In (29) and (30) we present a revised version of the rules for functional application. Note that the difference between the rule for left application (X/Y) and right application (Y\X) is only in where the argument is found. Neither the predicate-argument structure nor the qualia structure are order-sensitive (see subsection (5.2) at page 10).

(29) [L/] U, $\langle$ X/Y : (f, $[QS_f, QS_{a'}]$) $\rangle$, T, V $\rightarrow$ Z if
T $\rightarrow$ $\langle$ Y : (a, QS_a) $\rangle$ and
U, $\langle$ X : ($f(a)$, $[QS_f, QS]$) $\rangle$, V $\rightarrow$ Z

if Y is basic then $QS = \{\ q \sqcup q' \mid q \epsilon QS_a \ \&\ q' \epsilon QS_{a'} \ \&\ q \sqcup q' \neq \emptyset\ \}$;
QS_a otherwise

(30) [L\\] U, T, $\langle$ Y\\X : (f, $[QS_f, QS_{a'}]$) $\rangle$, V $\rightarrow$ Z if
T $\rightarrow$ $\langle$ Y : (a, QS_a) $\rangle$ and
U, $\langle$ X : ($f(a)$, $[QS_f, QS]$) $\rangle$, V $\rightarrow$ Z

if Y is basic then $QS = \{\ q \sqcup q' \mid q \epsilon QS_a \ \&\ q' \epsilon QS_{a'} \ \&\ q \sqcup q' \neq \emptyset\ \}$;
QS_a otherwise

The functor usually has limited selection restrictions ($QS_{a'}$) on its argument, whereas the argument itself tends to have a more elaborate qualia structure (QS_a). A functor may for instance require its argument to be of sort **human**, whereas the argument may have **boy** as one value of its qualia structure, next to a few other ones.

We take the qualia structure of the resulting type (the result of applying the functor to the argument) to be the set of qualia roles of the argument that unify with requirements of the functor. In other words: that subset of the qualia structure of the argument that meets some requirements of the functor. From the set of qualia roles of the argument it takes only those that unify with any of the requirements of the functor.

The case where the unification of $QS_{a'}$ and QS_a is empty—there is no value that unifies some qualia role from the argument and some qualia role from the functor—indicates that no interpretation can be found on the basis of the present information. Such a restriction of the interpretation process to coercion from the initial qualia structure, working only with stipulated values, implies accepting that the process may fail at some point.

A sentence like (31), where an address should be coerced to the function of the person who works there, the prime minister, and from there to his or her spokesperson, will fail.

(31) Downing Street denied all knowledge today.

That our theory does not capture all possible readings a sentence may have, is in contrast with what we stated earlier, but we would need a recursive application of coercion to avoid failing and derive values also from the embedded qualia structures. Such an approach may be computationally uncontrollable and we therefore pursue it no further here[9].

Abstraction The rules for abstraction are given in (32) and (33). Since the resulting category X/Y (Y\X, respectively) is hypothetical, the requirements ($QS_{a'}$) for the embedded argument can not be lexically given, but will result from the derivation. Similarly, the introduced category Y has only a hypothetical qualia structure QS_a, that must be instantiated *on the fly*, i.e., through argument cancelation elsewhere in the derivation.

(32) [R/] $\mathrm{T} \rightarrow \langle\ \mathrm{X/Y} : (f, [QS_f, QS_{a'}])\ \rangle$ if
$\mathrm{T}, \langle\ \mathrm{Y} : (a,\ QS_a)\ \rangle \rightarrow \langle\ \mathrm{X} : (\lambda a.f, [QS_f, QS_{a'}])\ \rangle$

(33) [R\] $\mathrm{T} \rightarrow \langle\ \mathrm{Y\backslash X} : (f, [QS_f, QS_{a'}])\ \rangle$ if
$\langle\ \mathrm{Y} : (a,\ QS_a)\ \rangle, \mathrm{T} \rightarrow \langle\ \mathrm{X} : (\lambda a.f, [QS_f, QS_{a'}])\ \rangle$

Axiom The identity case, the axiom, is given in (34). This now is the place where the final type-check occurs. Note that this is *not* where coercion takes place, coercion is typically a process between functor and argument.

(34) [Ax] $\langle\ \mathrm{X} : (t,\ QS_t)\ \rangle \rightarrow \langle\ \mathrm{X} : (t,\ QS_t)\ \rangle$

5.7 Examples

5.7.1 'begin a novel'

In the following example we demonstrate how the transitive verb *begin* combines with the noun phrase *a novel* to constitute the verb phrase *begin a novel*.

[9] In (31) one might think of looking into the qualia structure of *Downing Street*, finding the address of the Prime Minister; then looking into the qualia structure of *Prime Minister* and finding an entry for *spokesperson* there.

(35) *begin* : $\langle$ (NP\S)/ S/(NP\S) : $(\lambda R\lambda x.R(\lambda y\exists e.begin(e,x,y))$,

$$\begin{bmatrix} \mathrm{QS}_f: & \begin{bmatrix} \mathrm{QS}_{f'}:\{\mathrm{x}_1 \ldots \mathrm{x}_n\}, \\ \mathrm{QS}_{a'}:\{\textbf{human}\} \end{bmatrix}, \\ \mathrm{QS}_a: & \begin{bmatrix} \mathrm{QS}_{f''}:\{\mathrm{y}_1 \ldots \mathrm{y}_n\}, & \\ \mathrm{QS}_{a''}: & \begin{bmatrix} \mathrm{QS}_{f'''}:\{\mathrm{z}_1 \ldots \mathrm{z}_n\}, \\ \mathrm{QS}_{a'''}:\{\textbf{event}\} \end{bmatrix} \end{bmatrix} \end{bmatrix})\rangle,$$

a novel : $\langle$ S/(NP\S) : $(\lambda P.\exists z.novel(z) \wedge P(z)$,

$$\begin{bmatrix} \mathrm{QS}_f:\{\mathrm{x}_1 \ldots \mathrm{x}_n\}, & \\ \mathrm{QS}_a: & \begin{bmatrix} \mathrm{QS}_{f'}:\{\mathrm{y}_1 \ldots \mathrm{y}_n\}, \\ \mathrm{QS}_{a'}:\{\textbf{artifact, read, write}\} \end{bmatrix} \end{bmatrix})\rangle \rightarrow$$

begin a novel : $\langle$ NP\S : $(\lambda x.(\exists z.novel(z) \wedge (\exists e.begin(e,x,z)))$,

$$\begin{bmatrix} \mathrm{QS}_f: & \begin{bmatrix} \mathrm{QS}_{f'}:\{\mathrm{x}_1 \ldots \mathrm{x}_n\}, \\ \mathrm{QS}_{a'}:\{\textbf{human}\} \end{bmatrix}, \\ \mathrm{QS}_a: & \begin{bmatrix} \mathrm{QS}_{f''}:\{\mathrm{y}_1 \ldots \mathrm{y}_n\}, & \\ \mathrm{QS}_{a''}: & \begin{bmatrix} \mathrm{QS}_{f'''}:\{\mathrm{z}_1 \ldots \mathrm{z}_n\}, \\ \mathrm{QS}_{a'''}:\{\textbf{read, write}\} \end{bmatrix} \end{bmatrix} \end{bmatrix})\rangle$$

where: **read, write** $\leq$ **event**.

The functor, with category (NP\S)/NP, applies to the argument, with category NP, to yield an NP\S. The semantics of the functor (f) applies to the semantics of the argument (a) and results in *f(a)*. The qualia structure of the functor QS_e remains unaffected, as well as the selection restrictions of the subject argument ({**human**}). The selection restrictions of the object argument ({**event**}) are intersected with the qualia structure of the NP ({**artifact, read, write**}) which results in {**read, write**}, following **read, write** $\leq$ **event**.

5.7.2 'BMW announced ...'

This example demonstrates how a predicate combines with its subject argument. The transitive verb 'announced' combines with the noun phrase 'BMW' to constitute the incomplete sentence 'BMW announced ...'[10]

(36) *BMW* : $\langle$ NP : (*bmw*, QS_a:{**company, spokesperson,...** }) $\rangle$,

announced : $\langle$ (NP\S)/NP : $(\exists e.\lambda yx.announce(e,x,y)$,

[10] Another way to look at it would be to assume a VP, and not a transitive verb, if one wishes to introduce bracketing. In that case the qualia structure of the object argument would have been specified. For our present purposes this is of no relevance.

$$\begin{bmatrix} \mathrm{QS}_f\!: \begin{bmatrix} \mathrm{QS}_{f'}\!:\!\{\mathrm{y}_1 \ldots \mathrm{y}_n\}, \\ \mathrm{QS}_{a'}\!:\!\{\mathbf{human}\} \end{bmatrix}, \\ \mathrm{QS}_a\!:\!\{\mathbf{event}\} \end{bmatrix}$$

BMW announced : $\langle$ S/NP : $(\exists e.\lambda y.announce(e, bmw, y)$,

$$\begin{bmatrix} \mathrm{QS}_f\!: \begin{bmatrix} \mathrm{QS}_{f'}\!:\!\{\mathrm{y}_1 \ldots \mathrm{y}_n\}, \\ \mathrm{QS}_{a'}\!:\!\{\mathbf{spokesperson}\} \end{bmatrix}, \\ \mathrm{QS}_a\!:\!\{\mathbf{event}\} \end{bmatrix} \rangle$$

where: **spokesperson** $\leq$ **human**.

The functor, with category (NP\S)/NP, applies to the argument, with category NP, to yield an S/NP. The semantics of the functor (f) applies to the semantics of the argument (a) and results in *f(a)*. Again, the qualia structure of the functor (QS_e) and the selection restrictions of the object argument ({**event**}) remain unaffected. The selection restrictions of the subject argument ({**human**}) are intersected with the qualia structure of the NP ({**company, spokesperson,**...}), which results in {**spokesperson**}, following **spokesperson** $\leq$ **human**.

6 Conclusion

The main criticism one may have against our work, is that it introduces a new combinatorial explosion in Categorial Grammar, that already suffers from the problem of spurious ambiguity (Hendriks 1993). A noun may have any finite number of values in its qualia roles, and this may cause computational problems. This is a valid objection. However, what we claim to derive is all the possible readings a sentence could have. In other words, we show how compositionality might work with functions that operate on polysemous arguments. However, any research into the modelling of constraints on interpretation that are set by our knowledge of the (im)possibilities of the physical world is beyond the scope of this paper.

Related to this, also the semantics / pragmatics interface remains a problem to be solved. It is not yet clear to what extent pragmatic inference is compositional or not. In this paper we tried to include at least that part which can be dealt with compositionally[11].

We did succeed however in increasing the semantic potential of Categorial Grammar, while maintaining compositionality. We have dealt with poly-

[11] Some constraints on the interface of semantics and pragmatics may come from presuppositions. Bos, Buitelaar and Mineur (1995) study the parallel between qualia roles and presuppositions and takes coercion to operate on the presuppositions that lexical entries trigger. This does not effect the truth values for the entities that are denoted by these lexical items.

morphism in lexical semantics. Instead of stipulating some n monomorphic types for one word, we adopted Pustejovsky's approach to stipulate one polymorphic type which through coercion will generate all n monomorphic types. Such an approach explains in part the creative use of language, which we find for instance in the use of metonymy. On the other hand we acknowledged that Pustejovsky's approach will not be helpful in cases of strict homonymy (bank/bank).

7 Acknowledgements

This paper has benefitted from discussions with Hiyan Alshawi, Johan Bos, Bob Carpenter, Ann Copestake, Kees van Deemter, David Milward, Michael Moortgat, Gerald Penn, James Pustejovsky, Hans Uszkoreit, Marc Verhagen, Kees Vermeulen, Leon Verschuur, our colleagues at Brandeis, Saarbrücken and Utrecht and from suggestions by an anonymous reviewer. As always we are solely responsible for any remaining errors. We would like to thank Kees van Deemter also for inviting us to write this paper.

References

Apresjan, J. 1973. Regular Polysemy. *Linguistics* (142).

Bar-Hillel, Y. 1953. A quasi-arithmetical notation for syntactic description. *Language* 29:47–58.

Bartsch, R. 1987. The construction of Properties under Perspectives. *The Journal of Semantics*, 5:293–320.

Bayer, S. 1994. The Coordination of unlike Categories. Paper of a conference presentation at the 1993 LSA meeting in Los Angeles, May.

Bierwisch, M. 1982. Semantische und konzeptuelle Repräsentation lexikalischer Einheiten. In *Untersuchungen zur Semantik*, ed. R. Ružička and W. Motsch, 61–99. Berlin: Akademie-Verlag.

Bos, J., P. P. Buitelaar and A.-M. Mineur. 1995. Bridging as Coercive Accommodation. In *Computational Logic for Natural Language Processing—Workshop Proceedings*, ed. Suresh Manandhar et.al. South Queensferry, Scotland, 3rd - 5th April, 1995

Buitelaar, P. P., and A. M. Mineur. 1994. Coercion and Compositionality in Categorial Grammar. In *Proceedings of the 9th Amsterdam Colloquium*, ed. P. Dekker and M. Stokhof, 175–188. Amsterdam. ILLC.

Cardelli, L., and P. Wegner. 1985. On Understanding Types, Data Abstraction, and Polymorphism. *Computing Surveys* 17(4):22–29.

Chomsky, Noam. 1957. *Syntactic Structures.* Janua linguarum, Ser. minor, Vol. 4. 's Gravenhage: Mouton.

Copestake, A., and E. Briscoe. 1992. Lexical Operations in a Unification-Based Framework. In *Lexical Semantics and Knowledge Representation. Lecture Notes in Artificial Intelligence 627*, ed. J. Pustejovsky and S. Bergler, 22–29. Berlin. Springer Verlag.

Dowty, David R. 1979. *Word Meaning and Montague Grammar*. Dordrecht, Holland: Reidel.

Gabbay, D. 1991. *Labelled Deductive Systems*. Oxford University Press. (to appear).

Hendriks, Herman. 1993. *Studied Flexibility*. Doctoral dissertation, Institute for Logic, Language and Information, Amsterdam, December.

Ingria, R. J. 1990. The Limits of Unification. In *Proceedings of the 28th Annual Meeting of the ACL*, 194–204.

Moortgat, M. 1992. Labelled Deductive Systems for categorial theorem proving. Technical Report OTS-WP-CL-92-003, Onderzoeksinstituut voor Taal en Spraak, Utrecht.

Moortgat, M. J. 1988. *Categorial Investigations. Logical and Linguistic Aspects of the Lambek Calculus*. Vol. 9 of Groningen-Amsterdam Studies in Semantics. Dordrecht, Holland: Foris.

Montague, Richard. 1974. In Richard Thomason, editor, *Formal Philosophy. Selected papers of Richard Montague*. New Haven: Yale University Press.

Morrill, Glyn V. 1994. *Type-logical Grammar*. Dordrecht: Kluwer Academic Publishers.

Nunberg, G. D. 1979. The nonuniqueness of semantic solutions: Polysemy. *Linguistics and Philosophy* 3.

Pollard, C., and I. A. Sag. 1987. *Information-based Syntax and Semantics*. Vol. 1, Fundamentals of CSLI Lecture Notes, Number 13. CSLI.

Pullum, G. K., and A. M. Zwicky. 1986. Phonological Resolution of Syntactic Feature Conflict. *Language* 62:751–773.

Pustejovsky, J. 1995. *The Generative Lexicon*. Cambridge: MIT Press.

Pustejovsky, J., and P. Anick. 1988. On the Semantic Interpretation of Nominals. In *Proceedings Coling88*. Boedapest.

Sowa, J. 1992. Logical structures in the lexicon. In *Lexical Semantics and Knowledge Representation. Lecture Notes in Artificial Intelligence*, ed. J. Pustejovsky and S. Bergler, Vol. 627, 577–613. Berlin: Springer Verlag.

Verschuur, L. 1994. Flexibly Disambiguating Metonymy. Technical Report 52, ITK, Tilburg, June. MMC Research Report.

Weinreich, U. 1964. Webster's third: A critique of its semantics. *International Journal of American Linguistics* 30:405–409.