Lexical Approaches to Argument Structure

Stefan Müller  
Freie Universität Berlin

Stephen Wechsler  
University of Texas at Austin

October 19, 2012
Abstract

We conclude that argument structure properties should be represented together with lexical items.

Contents

1 Introduction 2

2 Lexicalist approaches 4
   2.1 Predicate argument structure 4
   2.2 Views on Lexical Rules 6

3 Non-lexical approaches 8
   3.1 Preliminary theory comparison 8
   3.2 Non-lexical approaches that refer to phrase structure 9

4 The pendulum of lexical and phrasal approaches 11
   4.1 Some historical notes 11
   4.2 Morphological Derivation 12
   4.3 Partial Frontings 13

5 Valence versus plugging 17
   5.1 Usage-based theories 17
   5.2 Coercion 20
   5.3 Valence and Coordination 22
   5.4 Valence and Derivational Morphology 28
   5.5 Simplicity and polysemy 29

6 Radical Underspecification: The End of Argument Structure? 31
   6.1 Little v and idiom asymmetries 33
   6.2 Deverbal nominals 35
   6.3 Idiosyncratic Syntactic Selections 38
   6.4 Expletives 42

†We thank Bill Croft, Charles Fillmore, Adele Goldberg, Paul Kay, Jean-Pierre Koenig, and Ivan Sag for extended discussions via email and in person and [your name here?] for comments on an earlier version of this paper.

We thank Patience Epps for forwarding our question regarding causatives and passivization to the LingTyp mailing list and Peter Arkadiev, Anna Bugaeva, Mark Donohue, Diana Forker, Alexander Letuchiy, Yury Lander, Geda Paulsen, Barbara Stiebels, Kashi Wali, and Bjoern Wiemer for examples and pointers to literature. We are grateful to all replies but decided to use only some of the data.

1 Introduction

Central to the mastery of a language is knowledge of the predicate-argument relations: an English speaker interpreting the sentence *The rabbit nibbled a carrot* knows that a nominal object following the verb *nibble* represents the food or other solid substance that is consumed, while a subject preceding it fills the role of the consumer of that substance. But the exact nature of that knowledge and how that information is represented within the grammar, remain matters of controversy within linguistics. Simplifying the current debate, one can distinguish *lexical* versus *phrasal* approaches.¹ In this paper we argue for a certain class of lexical approaches.

In *lexical* (or *lexicalist*) approaches, words are phonological forms paired with valence structures (also called *predicate argument structures*). For a certain head word (in our example above, the verb *nibble*) and certain argument phrases, the

¹ The phrasal approaches are usually called *constructional*, but we use that label cautiously since it is also used for approaches that are explicitly lexical. See for instance Kay (2005); Sag (To appear).
word’s predicate argument structure specifies the meaning of the combination as a function of the meanings of the parts. Lexical rules grammatically encode the systematic relations between cognate forms and diathesis alternations. The syntactic combinatorial rules are usually assumed to be very general and few in number.

In contrast, phrasal (or constructional; but see footnote 1) approaches eschew the use of lexical rules. Instead, different morphological cognates and diathesis alternants are captured by plugging a single word (or root) into different constructions. The construction carries a meaning that combines with the word’s meaning. In some versions the constructions are phrasal structures, while in others, they are non-phrasal grammatical constructs called argument structure constructions that resemble the lexicalist’s predicate argument structure, minus the specific verb or other predicator.

The lexical and phrasal approaches differ. The lexicalist’s predicate argument structure is an autonomous, reified grammatical entity, while the constructional approach seeks to avoid such entities. A phrasal construction or argument structure construction is more like a mechanism for the filling of argument roles during the generation or interpretation of a sentence. It is ‘grounded’ in actual sentences. Also, as noted above, the construction carries a meaning, and so some of the phrasal approaches would replace standard phrase structure rules or syntactic valence frames with meaningful constructions. For both of these reasons, constructional approaches are often affiliated with usage-based theories of human language that deny the existence, or downplay their importance, of ‘meaningless’ algebraic syntactic rules such as phrase structure rules defined purely on syntactic categories like V and NP. On the usage-based view, the progressive generalization over input patterns that explains language acquisition and use is incapable of abstraction to the point of removing communicative content entirely (Tomasello, 2003). Thus the resolution of the lexical-constructional debate has potentially broad theoretical consequences.

Sections 2 and 3 lay out these approaches in more detail and present preliminary arguments in favor of a particular lexical approach. Following that initial outline of the approaches, Section 4 provides a brief historical overview of the developments in theoretical linguistics of the last century, which shows that the development progressed in waves oscillating between phrasal and lexical approaches. We discuss the reasons for changes and thereby point to problems that still exist in current approaches, or have been reintroduced into them. Then we revisit some classic arguments for the lexical approach, present some new arguments, and answer challenges involving acquisition (Tomasello, 2003; Goldberg, Casenhiser and Sethuraman, 2004), psycholinguistics (Goldberg, 1995, 2006), neurolinguistics (Pulvermüller et al., To appear), statistical distribution (Stefanowitsch and Gries, 2009; Bod, 2009a,b), and coercion. We conclude in favor of the lexical approach.

2 Lexicalist approaches

2.1 Predicate argument structure

On lexical approaches, a word’s predicate argument structure, or valence structure, indicates the number and type of arguments, and specifies the meaning of the combination of the word and its argument phrases as a function of the meanings of the parts. The following entry for the word *nibble* indicates that when it appears together with certain arguments, the combination has a certain semantic CONTENT:

(1) A predicate argument structure:

\[
\begin{array}{c}
\text{PHON}\langle\text{nibble}\rangle \\
\text{ARG-ST}\langle\text{NP}\,\text{NP}\rangle \\
\text{CONTENT}\text{nibble} (\text{agent: } x, \text{ patient: } y)
\end{array}
\]

The rules of syntax specify the positions for ARG-ST list items, thus interacting with this structure to license a grammatical clause or other phrasal construction with the right meaning. Note that the predicate argument structure is slightly abstract: it does not directly encode the phrase structure or precedence relations between this verb and its arguments. This abstraction captures the commonality across different syntactic expressions of the arguments of a given root.

(2) a. The rabbits were nibbling the carrots.
   b. The rabbits were nibbling at/on the carrots.
   c. The rabbits were nibbling.
   d. The carrots were being nibbled (by the rabbits).
   e. a large, partly nibbled, orange carrot
   f. the quiet, nibbling, old rabbits
   g. the rabbit’s nibbling of the carrots
   h. The rabbit gave the carrot a nibble.
   i. The rabbit wants a nibble (on the carrot).
   j. The rabbit nibbled the carrot smooth.

Verbs exhibit variable polyadicity, i.e. direct-oblique and other diathesis alternations (2a,b), argument optionality (2c), and morpholexical operations like passive (2d), as well as antipassive, causative, and applicative in other languages. They have cognates in other parts of speech such as adjectives (2e,f) and nouns (2g,h,i). Verbs have been argued to form complex predicates with resultative secondary predicates (2j), and with serial verbs in other languages.

On the view pursued here, the same root lexical entry *nibble*, with the same meaning, appears in all of these contexts. The effects of lexical rules together with the rules of syntactic mapping dictate the proper argument expression in each context. For example, if we call the first two arguments in an ARG-ST list (such as the one in (1) above) Arg1 and Arg2, respectively, then in an active transitive sentence Arg1 is the subject and Arg2 the object; in the passive, Arg2 is the subject. When adjectives are derived from verbal participles, whether active (*a nibbling rabbit*) or passive (*a nibbled carrot*), the rule is that whichever role would have been expressed as the subject of the verb is assigned by the participial adjective to the noun that it modifies.

Summarizing so far, a predicate argument structure specifies the relation between a head word, its arguments, and the meaning that results when they are combined. Rules of syntactic mapping specify the way the arguments are realized (or suppressed) in the syntactic environment of the word. Lexical rules specify systematic relations *between* predicate argument structures, within the grammar of a language. The main issue we address in this paper is whether such lexical rules, as described in more detail in the next section, exist within the minds of speakers. We argue that they do.

In addition to predicate argument structures, we assume that grammars include meaningful phrasal constructions. That is, we agree with (Goldberg, 1995; Tomasello, 2003; Goldberg and Jackendoff, 2004; Jackendoff, 2011) that grammars should contain a phrasal component for certain constructions, such as the N-P-N construction of the kind in (3) discussed by Jackendoff (2008) and the verbless directives in (4) mentioned by Jackendoff and Pinker (2005, p. 220) and discussed in detail by Jacobs (2008).

(3) student after student
    [NP/advp N-P-N]

(4) a. Off with his head!
    b. Into the trunk with you!

In addition to cases like (3) and (4), the analysis of some idioms seems to call for *phrasal lexical items*, that is, phrases in which more than one word is fixed (Abeillé and Schabes, 1989; Richter and Sailer, 2009). Other classes of idioms can be handled by analyses in which words select particular lexemes in their valence features (Sag, 2007). Some combination of these two is often posited, in order to capture the full range of idiom types, from fixed phrases to syntactically analyzable idioms. See Sailer, 2000; Soehn and Sailer, 2008 for lexical approaches to idioms.

---

2See G. Müller (2011) for a lexical account of Jacobs’ data and Müller (2010a, Section 11.11.9.1) for discussion.
While we think grammars include meaningful phrasal constructions, we do not think lexical rules can or should be eliminated by representing argument structure phrasally.

2.2 Views on Lexical Rules

There are various understandings of what lexical rules are. The first dimension along which lexical rules could be classified was discussed by Jackendoff (1975). Jackendoff distinguishes between two conceptions: (i) lexical rules that relate two stored lexical entries and thereby capture redundancies in the lexicon; and (ii) lexical rules that license new lexical items. On the latter view, lexical rules can apply to stored lexical items (which are called lexical entries here) or to lexical items that are licensed by a lexical rule or a chain of lexical rules.

The second dimension was discussed in the more formal literature on lexical rules in the 90s (Copestake, 1992; Riehemann, 1993, 1998; Calcagno, 1995; Briscoe and Copestake, 1999; Meurers, 2001), but as it turns out there seem to be not just formal but also empirical differences between the approaches (Goldberg, To appear). Calcagno (1995) and Calcagno and Pollard (1995) argued for a view on lexical rules that was called the meta-level approach by Meurers (2001). This approach can be sketched as in (5):

\[
L_1 \rightarrow L_2
\]

Here \( L_1 \) and \( L_2 \) are descriptions of lexical objects. The rule states that if the language contains a lexical object satisfying \( L_1 \) then it contains another lexical object satisfying \( L_2 \). Thus the rule is not itself a description but a ‘meta-description.’ An alternative is the description level approach suggested by Copestake, Briscoe, Riehemann, and Meurers. There are two variants of this approach. Both are shown in (6):

\[
\begin{align*}
(6) \quad \text{a.} & \quad \begin{bmatrix} \text{OUT} & L_2 \end{bmatrix} \\
& \begin{bmatrix} \text{IN} & L_1 \end{bmatrix} \\
& \begin{bmatrix} \text{lr-type} \end{bmatrix}
\end{align*}
\]

\[
\begin{align*}
(6) \quad \text{b.} & \quad \begin{bmatrix} L_2 \end{bmatrix} \\
& \begin{bmatrix} \text{DTR} & L_1 \end{bmatrix} \\
& \begin{bmatrix} \text{lr-type} \end{bmatrix}
\end{align*}
\]

What is shown in (6) is two typed feature descriptions. In both versions \( L_1 \) is the value of a feature. In (6a) the output of the lexical rule is the value of a feature, while it is not in (6b). The lexical rule in (6b) licences an object that can be used in further analyses since it basically has the properties of \( L_2 \).

\[3\]For approaches that assume a representation like (6a) additional machinery is needed. Meur-
2.2 Views on Lexical Rules

In what follows, we mean (6b) when we speak of lexical rules. As suggested by the name of the feature DTR, which stands for DAUGHTER, this type of lexical rule is equivalent to a unary branching tree. In Figure 1 a lexical rule licenses the subtree where the daughter node is the verb BREAK and the mother has the 3rd person singular inflected form (breaks) and agreement features. In Figure 2 on the next page the passive lexical rule licenses a structure with the same verb BREAK is the sole daughter and the mother has the passive participle form and a valence feature in which the role of the object of the active is assigned to the subject of passive. Figure 3 on page 9 illustrates the participle-to-adjective conversion rule as applied to a passive verbal participle. The effect of this rule is that the semantic role that the participle would assign to its subject is assigned by the adjective to the noun it modifies.

\[
S \\
\text{NP} \quad \text{VP} \\
\quad \text{V} \quad \text{NP} \\
\text{Kim} \quad \text{breaks} \quad \text{toys} \\
\text{BREAK}
\]

Figure 1: Example of lexical rule for present tense verb.

ers suggested the following word principle to be able to state that lexical items that are licensed by lexical rules can be used as words.

\[
\text{word} \Rightarrow (L_1 \land \text{STORE}(\langle \rangle)) \lor \ldots \lor (L_2 \land \text{STORE}(\langle \rangle) \ldots) \lor \text{STORE} \left[ \text{OUT} \begin{split} & 1 \\ & \text{lex-rule} \end{split} \right]
\]

This word principle disjunctively lists all lexical entries and furthermore says that linguistic objects that correspond to the output of a lexical rule are words. The $1$ indicates structure sharing, that is, identity of values. So the linguistic object following the first occurrence of $1$ in (6) is identical to the OUT value of lexical rules like the one in (5a). In this sense the view of lexical rules in (5a) paired with a word principle like (6) is rather similar to the view in (5b). The approaches in (5b) and (5a)+(6) are called template-based approaches by Goldberg (To appear).

Sag (To appear) suggests an approach along the lines in (5a) together with a meta principle that states that the OUT value is a well-formed linguistic object licensed by the theory that can be used in other constructions. This approach is similar to the meta level approach in that the licensed objects do not contain the object from which they are derived. So (4) and Sag’s proposal correspond to what Goldberg calls lexical rules.
3 Non-lexical approaches

Instead of using lexical rules, non-lexical approaches capture morphological cognates and diathesis alternants for a single word (or root) by plugging the word into different phrasal constructions. The construction carries a meaning that combines with the word’s meaning. Phrasal constructions such as the Intransitive, Transitive, and Ditransitive constructions replace the phrase structure rules or valence frames of other syntactic theories; others include the Caused Motion and Resultative constructions. The ditransitive construction means ‘X caused Y to receive Z’ and can combine with either a 3-argument verb like \textit{fax} (Pat faxed Bill the letter) or a 2-argument verb like \textit{bake} (Pat baked Bill a cake). In the latter case the construction licenses the recipient argument.

3.1 Preliminary theory comparison

As noted in the introduction, there is a subtle but real theoretical difference between the lexical and non-lexical constructional approaches. The lexicalist’s predicate argument structure is an autonomous, reified grammatical entity that therefore need not immediately combine with its specified arguments, but can alternatively meet other fates: it can serve as the input to a lexical rule; it can combine first with a modifier in an adjunction structure or with a similar head in a coordinate structure; instead of being realized locally, one or more of its arguments can be effectively transferred to another head’s valence feature (raising or argument transfer); or arguments can be saved for expression in some other syntactic position (partial fronting). (These phenomena are discussed below.) But on the non-lexical constructional approach, the argument structure construction is more like a mechanism for the filling of argument roles during the generation or interpretation
3.2 Non-lexical approaches that refer to phrase structure

There are two major variants of phrasal approaches. In some versions the constructions are non-phrasal grammatical constructs called argument structure constructions (ASCs). An ASC contains roughly the same information as a lexicalist’s predicate argument structure but without a specific verb or other predicator. Goldberg’s (1995) ASCs contain grammatical relation names like SUBJ, OBJ, and OBL. Hence her ASC closely resembles an LFG functional structure, only without the verb specified. The verb is stored with some of its roles specified as profiled, which means they are destined for realization as direct grammatical relations (SUBJ or OBJ). Goldberg assumes that her phrasal constructions just specify grammatical functions that have to be realized together with a certain head. That...
is, such constructions can be underspecified with regard to linear order. The only requirement is that the parts of the construction have to be realized somewhere in a structure (Goldberg, 1995; Goldberg, 2006). How this comes about is not worked out in detail.

Other authors assume phrase structure-like objects, that is, a certain configuration with part of speech and structural information is paired with a certain meaning (Alsina, 1996; Goldberg and Jackendoff, 2004; Bergen and Chang, 2005; Culicover and Jackendoff, 2005; Asudeh, Dalrymple and Toivonen, 2008; Jackendoff, 2011). We illustrate the empirical consequences of the two variants with the example of verb-particle constructions in Dutch and German. Booij (2002, Section 2; To appear) and Blom (2005), working in the frameworks of Construction Grammar and LFG, respectively, assume that particle verbs are licensed by phrasal constructions (pieces of phrase structure) in which the first slot is occupied by the particle.

(7) \[ X \ [ ]_v \] \[ . \] where \( X = \text{P}, \text{Adv}, \text{A}, \text{or N} \]

Examples for specific Dutch constructions are:

(8) a. \[ \text{af} \ [ ]_v \] \[ . \]
    b. \[ \text{door} \ [ ]_v \] \[ . \]
    c. \[ \text{op} \ [ ]_v \] \[ . \]

This suggestion comes with the claim that particles cannot be fronted. However, this claim is wrong for languages like Dutch and German. On Dutch see Hoeksema, 1991, p. 19, on German Müller, 2002a,b, 2003b, 2007c. Some more fundamental remarks on introspection and corpus data with relation to particle verbs can also be found in Müller, 2007c; Meurers and Müller, 2009. A German example is given in (9), several pages of attested examples can be found in the cited references and some more complex examples will also be discussed in Section 12 on page 62.

(9) \text{Los} \text{damit} \text{geht} \text{es} \text{schon} \text{am} \text{15. April} \[ . \]
\text{PART there.with} \text{went it} \text{already at.the} \text{15. April} \\
\text{‘It already started on April the 15th.’}

Particle verbs are mini-idioms. So the conclusion is that idiomatic expressions that allow for a certain flexibility in order should not be expressed in phrasal configurations describing adjacent elements. For some idioms, a lexical analysis along the lines of Sag, 2007 seems to be required.\footnote{taz, 01.03.2002, p. 8, see also Müller, 2005b, p. 313.}

\footnote{Note also that the German example is best described as a clause with a complex internally structured constituent in front of the finite verb and it is doubtful whether linearization-based...}
4 The pendulum of lexical and phrasal approaches

4.1 Some historical notes

Our historical section starts with Categorial Grammar (Ajdukiewicz, 1935; Steedman, 2000). Categorial Grammar (CG) is the prototype for a lexical model: every word (every functor) comes with descriptions of its arguments and the rules that combine functors with their arguments are very general and few in number. For instance an English transitive verb like know is assigned the lexical entry (s\np)/np. This means that knows takes an NP to its right and an NP to its left. The rules for combination do not contain any part of speech information. For instance the rule that combines a verb like know with its object has the form X/Y * Y = X. Such general combinatory rules have a component for semantic combination (for instance, functional application or composition).

Another branch of theoretical linguistics assumed phrase structure rules as base component in a transformational setting (Chomsky, 1957). While the rules of CG are binary branching and rather abstract, the early phrase structure rules were not. There were rules for VPs with ditransitive verbs that had three daughters (for examples see Chomsky, 1965, p. 72, 96, 107). On some analyses phrase structure rules introduced rich semantic features directly into the phrase structure, such as CAUSE for causation (Chomsky, 1970), an approach greatly expanded in the Generative Semantics school (Lakoff, 1969).

There were different answers on the question of how to integrate semantics into Generative Grammar: Transformational Grammar started out assigning semantics on the level of Deep Structure but problems quickly became apparent, which led to modifications of the framework and to interpretation rules that took into account Surface Structure as well (see Bach, 1976 for an overview). An alternative to the prevalent view in Transformational Grammar was proposed by Montague (1973), who assumed that interpretation is combined with the rules of syntactic combination. Bach (1976, p. 184) called this the rule-to-rule assumption. Also in the 1970s other non-transformational theories like TAG (Joshi, Levy and Takahashi, 1975), LFG (Bresnan and Kaplan, 1982), and GPSG (Gazdar, Klein, Pullum and Sag, 1985) were developed and some of them came with detailed semantic representations. For instance Gazdar (1982) and Gazdar, Klein, Pullum and Sag (1985, Chapter 10) are very explicit about the semantic representations and the combination rules for GPSG. They allow for rule-specific semantic interpretation and in fact propose a quite specific composition rule for passivized proposals like the ones in Kathol, 1995, p. 244–248 or Wetta, 2011 can capture this. The issue of particle verbs will be taken up in Section 12 again, where we discuss evidence for/against phrasal analyses from neuroscience. Resultative and caused motion constructions will be discussed in Section 5.4.
sentences (p. 219). That is they share the rule-to-rule assumption.

While Montague’s proposal was in the spirit of Categorial Grammar and assumed binary branching structures, GPSG was not: The authors of GPSG assume classical context free phrase structure rules, for example a VP rule with a verb and two objects on the right-hand side. Uszkoreit (1987) assumes (derived) rules for clauses in German that licences a verb together with all of its arguments, and while no interpretation rules are given in this book, it is clear that this rule is combined with a semantic representation in a fully worked out version of the theory.

The GPSG of the 1980s resembled some current versions of Construction Grammar in its adoption of what we call a plugging proposal: a verb that is semantically compatible with a certain phrasal construction is plugged into this construction. Valence information is not represented as part of lexical items in GPSG. Instead lexical items had a number assigned to them and could be inserted into phrasal rules that had the same number. It is only in interaction between rules and these numbers that lexical items are paired with certain arguments. For instance laugh is of category 2 so it can form a VP if used with rule (10a) and love, of category 3, can form a VP with rule (10b).

\[
\begin{align*}
(10) & \quad \text{a. } \text{VP } \rightarrow \text{H}[2] \\
& \quad \text{b. } \text{VP } \rightarrow \text{H}[3], \text{NP}
\end{align*}
\]

(The H stands for head, that is, for the verb in (10)). On this model lexical rules are impossible because the verb has no valence feature to which lexical rules could apply. Alternations like the passive, for example, were captured entirely within the phrase structure component, through meta-rules that expanded the stock of phrase structure rules.

In the next subsections we look at the problems that this proposal faced in order to understand why it was finally given up and replaced by theories that assume a lexical representation of valence information. We will look at two phenomena here: morphological derivation, and partial frontings.

### 4.2 Morphological Derivation

The first problem with the GPSG model is that there are morphological processes that are sensitive to valence (Müller, 2010a, p. 129). For instance -able derivation (and German -bar derivation) is possible with transitive verbs only:

\[
\begin{align*}
(11) & \quad \text{a. } \text{lößbar} \quad \text{(nominative, accusative)} \\
& \quad \text{solveable} \\
& \quad \text{b. } \text{vergleichbar} \quad \text{(nominative, accusative, PP[mit])} \\
& \quad \text{comparable}
\end{align*}
\]
4.3 Partial Frontings

The verbs have to have at least a nominative and an accusative argument (11a,b), intransitive verbs like *sleep or help do not allow for the -bar derivation.

Moreover, it will not work to say that -bar derivation applies only to verbs with a certain number. For example, *lösen (‘to solve’) and *vergleichen (‘to compare’) have different valence frames. This means that a GPSG rule for -bar derivation would have to mention several numbers that correspond to different valence frames that allow for -bar derivation. Since the numbers by themselves do not contain any information about the presence of a direct object, such a formulation of the -bar derivation rule would amount to stipulating a seemingly arbitrary set of numbers, and thereby miss an important generalization. This should be compared to models that assume a lexical representation of valence: the -bar suffix can be specified to attach to verbs whose valence list starts with a nominative and an accusative and hence the generalization is captured easily in such models.

4.3 Partial Frontings

Another reason for needing valence information is to allow for variation in where in the sentence structure the arguments are discharged. For example, German allows for partial frontings like (12):

(12) a. [Erzählen] wird er seiner Tochter ein Märchen können.
    tell will he.NOM his daughter.DAT a fairy.tale.ACC can
    ‘He will be able to tell his daughter a fairy tale.’
b. [Ein Märchen erzählen] wird er seiner Tochter können.
    a fairy.tale.ACC tell will he his daughter.DAT can
c. [Seiner Tochter ein Märchen erzählen] wird er können.
    his daughter.DAT a fairy.tale.ACC tell will he.NOM can

The non-finite verb erzählen may be realized together with all its complements (12c) or with proper subsets of its complements (12a,b) in the so-called prefield to the left of the finite verb (subjects can also be fronted with non-finite verbs, but this is rather restricted). The problem for GPSG-like approaches is that the arguments are licensed by a certain phrase structure rule. To be able to analyze (12a) and (12b) one needs phrase structure rules that license the verb without any argument and with a single argument, respectively. In addition it has to be ensured that the arguments that are missing in the prefield are realized in the remainder of...
The pendulum of lexical and phrasal approaches

the clause. It is not legitimate to omit obligatory arguments or realize arguments with other properties like a different case, as the examples in (13) show:

(13)   a. Verschlungen hat er es nicht.
       devoured has he it\text{ACC} not
       ‘He did not devour it.’

       b. * Verschlungen hat er nicht.
           devoured has he not

       c. * Verschlungen hat er ihm nicht.
           devoured has he him\text{DAT} not

The obvious generalization is that the fronted and unfronted arguments must add up to the total set belonging to the verb.

Nerbonne (1986) and Johnson (1986) suggest GPSG analyses that can deal with the data. However, they assume a valence representation that uses binary features like NPacc and NPdat. This makes it possible to represent the fact that the accusative object is realized in the prefield in (12b) and may not be realized in the remainder of the clause (in the so-called middle field). Similarly the dative object in (12b) is realized in the middle field and hence may not be realized in the prefield. As both authors state clearly, this incorporates ideas from Categorial Grammar into GPSG. Theories like HPSG (Pollard and Sag, 1987, 1994) that were developed after GPSG, explicitly borrow from CG and hence take over the CG solution to the partial fronting problem that was developed by Geach (1970). See for instance (Pollard, 1996; Meurers, 2000; Müller, 1996, 2002a; Kathol, 2001) and also (Nerbonne, 1994).

If one does not want to go with the lexical specification of valence frames, there seem to be just two alternatives: remnant-movement analysis as often assumed in the transformational literature (G. Müller, 1998) and linearization-based approaches that allow for discontinuous constituents (Reape, 1994). In remnant-movement-based approaches it is assumed that the prefield is filled by a VP. The elements that are not realized in the prefield are moved out of the VP before the (remnant of the) VP is fronted. Such movement-based analyses are usually not assumed in non-transformational frameworks, but apart from theoretical considerations there are also empirical facts that argue against remnant movement. See Haider, 1993, p. 281, De Kuthy, 2002, Chapter 4.2.5, De Kuthy and Meurers, 2001, Section 2, and Fanselow, 2002 for details.

The linearization-based variant seems to be what Goldberg (2006, p. 10) has in mind when she writes that argument structure constructions are fused with other constructions like a VP construction. To our knowledge the details of such an

---

6See Hinrichs and Nakazawa, 1994 for a notable exception. This work shows that it would be possible in principle to assume remnant movement analyses.
4.3 Partial Frontings

Analysis have not been worked out within the Construction Grammar setting, so we will discuss explicit proposals in other frameworks: The linearization proposal by Reape (1994) was criticized by Kathol (2001, Section 8.6), who argued on the basis of agreement, case assignment, and passive for a CG-like analysis of German verbal complexes. Reape assumed that a raising verb like *scheinen* (‘to seem’) embeds a full clause and allows for a discontinuous linearization of the parts of this clause. Similarly, verbs that allow for the formation of a verbal complex as for instance the control verb *verprechen* (‘to promise’) allowed the parts of its verbal argument to be serialized discontinuously. Kathol argued that such an approach fails to capture local agreement relations between the finite verb and the subject of a clause that is embedded under a raising verb. Consider his examples in (14):

(14) a. Den Mädchen scheint /* scheinen schlecht zu werden.
   ‘The girls seem to be getting ill to the stomach.’

b. Du scheinst /* scheint nicht zu verstehen.
   ‘You don’t seem to understand anything.’

The problem with a purely linearization-based account is that the inflection of the verb that is the head of the subject does not reflect any agreement, since it is an infinitive with *zu*. Instead we have agreement with the finite verb one level up. The example in (14a) shows that there does not have to be a subject at all. This is due to the fact that German allows for subjectless predicates and that raising verbs do not care whether the downstairs predicate selects for a subject or not. An approach that assumes that *du* (‘you’) is an argument of *scheinen* (‘seem’) can account for the agreement relation locally. Similarly, there are so-called remote passives in German. The object of a deeply embedded verb gets assigned nominative (Höhle, 1978, p. 175–176):

(15) weil der Wagen oft zu reparieren versucht wurde
   ‘because many attempts were made to repair the car’

This is explained by an analysis that assumes that *zu reparieren versucht* behaves like a complex word with respect to passive and hence the accusative object of *zu reparieren versucht* has to be realized as nominative. See (Kathol, 1994; Pollard, 1994; Müller 1999, Chapter 15.3.6; 2002a, Chapter 3.2.5) for argument composition analyses. Kathol’s criticism of linearization-based analyses was relativized later by the finding that the integration of data like (16) into the analysis of non-finite complementation requires some non-locality:
(16) a. [Ein Außenseiter gewonnen] hat hier noch nie.\(^7\)

an outsider.NOM won has here still never

‘No outsider has yet won here.’

To assign the proper case in the prefield and to account for agreement with the NP in the prefield this information has to be accessible to the finite verb. However, if gewonnen has the category s\np and ein Außenseiter gewonnen has the category s, the information about the np is not accessible any longer when ein Außenseiter gewonnen is combined with the finite verb. This shows that analyses that are based on the notion of argument composition alone are not sufficient (Meurers, 2000; Meurers and De Kuthy, 2001). This means that more like a basic Categorial Grammar analysis or an analysis under the assumptions of Pollard and Sag (1987) and Pollard and Sag (1994) is needed (Kathol, 2003, p. 27).

That the problem of partial fronting is non-trivial and that simply referring to fusion with other constructions is not sufficient is also indicated by data like the Danish example of partial fronting in (17a) that is due to Holmberg (1999), (17c) – which is taken from Müller and Ørsnes, 2012 – provides a corpus example:

(17) a. ?Læst har Bjarne den ikke.

read has Bjarne it not

‘Bjarne did not read it.’

b. *Læst har Bjarne ikke bogen

read has Bjarne not book.DEF

‘Bjarne did not read the book.’

c. men helt [udelukke] kan man [det] da ikke eller hvad\(^8\)

but wholly exclude can you it then not or what

‘but you cannot wholly exclude it, can you?’

Danish marginally allows for partial frontings but only if the element that is missing from the fronted VP is a pronoun as in (17a). Pronouns have to be serialized to the left of the negation (18a) and full NPs to the right of it as in (18b).

(18) a. Bjarne læser den ikke.

Bjarne reads it not

‘Bjarne is not reading it.’

b. Bjarne læser ikke bogen.

Bjarne reads not book.DEF

‘Bjarne is not reading the book.’

The point about the examples in (17a) and (17b) is that the linearization of the material with respect to each other and with respect to traditional topological fields

\(^7\)Haider, 1990, p. 96.

\(^8\)http://hope.pointblog.dk/svaert-at-vide-.html, 26.3.2012.
for Danish (Diderichsen, 1957; Ørsnes, 2009) is correct, but still (17a) is possible, but (17b) is not. Any proposal that tries to account for the difference between (17a) and (17b) has to explain why the VP can be discontinuous if the object is a pronoun but not when the object is a full NP. Müller and Ørsnes (To appear, 2012) present a fully worked out approach that uses argument composition, but it is difficult to see how a linearization- or fusion-based approach could account for the data in a non-stipulative way.

Concluding this section, we can say that phrasal models have been proposed earlier in the last century and that the development runs in waves. There have been arguments against GPSG-like analyses that are still valid and the problems are not addressed by current phrasal approaches. On the other hand there is evidence that purely lexical approaches in the spirit of basic Categorial Grammar without any complex valence representations are not sufficient either.

5 Valence versus plugging

We have seen how the pendulum has swung between lexical and phrasal approaches. Of the theories on offer, the best place for that pendulum to come to rest, in our view, is at a theory in which words are equipped with valence information that is subject to the effects of lexical rules formulated as unary branching trees (see Section 2.2 above). The previous section reviewed earlier arguments for needing lexical representations of valence. In this section we present more detailed arguments specifically directed against the claim that lexical valence representations (i.e. predicate argument structures) can or should be replaced by what we call a plugging proposal, that is, a system in which a verb or other predicator is plugged into a meaningful construction.

As noted in Section 2.1 above, we believe that grammars include meaningful phrasal constructions. Our purpose is not to argue against their existence, but rather to argue that they cannot replace lexical valence representations. In the next section we begin by examining claims to a purported advantage of Construction Grammar over lexical rules. Then we turn to positive arguments for lexical rules.

5.1 Usage-based theories

For many practitioners of Construction Grammar, their approach to syntax is deeply rooted in the ontological strictures of usage-based theories of language (Langacker, 1987; Goldberg, 1995; Croft, 2001; Tomasello, 2003). Usage-based theorists oppose the notion of ‘linguistic rules conceived of as algebraic procedures for combining symbols that do not themselves contribute to meaning.’ (Tomasello, 2003, p.99) All linguistic entities are symbolic of things in the realm
of denotations; ‘all have communicative significance because they all derive directly from language use.’ (ibid) Although the formatives of language may be rather abstract, they can never be divorced from their functional origin as a tool of communication. The usage-based view of constructions is summed up well in the following quote:

The most important point is that constructions are nothing more or less than patterns of usage, which may therefore become relatively abstract if these patterns include many different kinds of specific linguistic symbols. But never are they empty rules devoid of semantic content or communicative function. (Tomasello, 2003, p. 100)

Thus constructions are said to differ from grammatical rules in two ways: they must carry meaning; and they reflect the actual ‘patterns of usage’ fairly directly.

Consider first the constraint that every element of the grammar must carry meaning, which we call the semiotic dictum. Do lexical or phrasal theories hew the most closely to this dictum? Categorial Grammar, the paradigm of a lexical theory (recall Section 4), is a strong contender: it consists of meaningful words, with only a few very general combinatorial rules such as X/Y * Y = X. Given the rule-to-rule assumption those combinatorial rules specify the meaning of the whole as a function of the parts; whether that counts as meaningful in itself is not clear.

What does seem clear is that the combinatorial rules of Construction Grammar, such as Goldberg’s Correspondence Principle for combining a verb with a construction (1995, p. 50), have the same status:

(19) The Correspondence Principle: Each participant that is lexically profiled and expressed must be fused with a profiled argument role of the construction. If a verb has three profiled participant roles, then one of them may be fused with a non-profiled argument role of a construction.

Both verbs and constructions are specified for participant roles, some of which are profiled. Argument profiling for verbs is ‘lexically determined and highly conventionalized’ (Goldberg, 1995, p. 46). Profiled argument roles of a construction are mapped to direct grammatical functions, i.e., SUBJ, OBJ, or OBJ2. By (19) the lexically profiled argument roles must be direct, unless there are three of them, in which case one may be indirect.9 With respect to the semiotic dictum, the Correspondence Principle has the same status as the Categorial Grammar combinatorial rules: a meaningless algebraic rule that specifies the way to combine meaningful items.

9We assume that the second sentence of (19) provides for exceptions to the first sentence.
5.1 Usage-based theories

Turning now to the lexicalist syntax we favor, some elements abide by the semiotic dictum while others do not. Lexical valence structures clearly carry meaning since they are associated with particular verbs. Also, the lexical rule that adds a benefactive recipient argument to a verb adds meaning. But phrase structure rules for intransitive and transitive VPs (or the HPSG ID schemata) do not.

Which structures have meaning is an empirical question for us. For example, Wechsler (1991, p. 111ff; 1995, p. 88ff) proposed a constructional analysis of English ditransitives that prefigures that of Goldberg (1995). Inspired by Kiparsky’s (1987; 1988) notion of a *thetically restricted positional linker*, together with cross-Germanic comparison, the idea was that the inner object position (the phrasal position of the first of two NP objects in a ditransitive) is semantically restricted to expressing the role of ‘intended recipient’: hence *He carved her a toy* entails that he carved a toy with the intention that she receive it. Verbs like *carve* were treated as 2-place, verbs like *give* as 3-place, and the ‘recipient’ semantics is contributed by the structure itself, redundantly in the case of *give*. Wechsler (1991, Chapter 3; 1995, Chapter 3) treats this object function as a ‘covert dative’ in the Germanic languages that lost dative case, including English. On the other hand, no special meaning was assumed for monotransitive or intransitive structures.

By contrast, in Construction Grammar meaning is assumed *a priori* for all constructions. But while the ditransitive construction plausibly adds meaning, no meaning has yet been discovered for either the intransitive or (mono)transitive constructions. In short, the constructionist’s evidence for the meaningfulness of certain constructions such as the ditransitive does not constitute evidence that all phrasal constructions have meaning.

Now consider the second usage-based dictum, that the elements of the grammar directly reflect patterns of usage, which we call the transparency dictum. The Construction Grammar literature often presents their constructions informally in ways that suggest that they represent surface word order patterns: the transitive construction is ‘X VERB Y’ (Tomasetto) or ‘Subj V Obj’ (Goldberg, 1995, 2006); the passive construction is ‘X was VERBed by Y’ (Tomasetto, 2003, p. 100) or ‘Subj aux Vpp (PPhy)’ (Goldberg, 2006, p. 5). But a theory in which constructions consist of surface patterns was considered in detail and rejected by (Müller, 2006, Section 2), and does not accurately reflect Goldberg’s actual theory.\footnote{Goldberg et al. (2004, p. 300) report about a language acquisition experiment that involves an SOV pattern. The SOV order is mentioned explicitly and seen as part of the construction.}

\footnote{This applies to argument structure constructions only. In some of her papers Goldberg assumes that very specific phrase structural configurations are part of the constructions. For instance in her paper on complex predicates in Persian (Goldberg, 2003) she assigns V⁰ and V categories.}

The more detailed discussions present argument structure constructions,
which are more abstract and rather like the lexicalists’ grammatical elements (or perhaps an LFG f-structure): the transitive construction resembles a transitive valence structure (minus the verb itself); the passive construction resembles the passive lexical rule.

With respect to fulfilling the desiderata of usage-based theorists, the difference between the non-lexical and lexical approaches is somewhat exaggerated, in our opinion. However, a difference can be seen by considering novel uses such as *He smiled her his answer* or even *He carved her a toy.* On the constructional approach, the extra argument(s) of the verbs *smile* and *carve* arise only in the course of constructing the sentence and combining the verbs with the nominals that fill those argument positions. There is no sense in which the verbs have three arguments prior to generating the sentence. Nor is there any substructure of the sentence that corresponds (phonologically) to just the verb’s pronunciation, but encodes the existence of the three arguments. Goldberg’s theory, for example, sedulously avoids positing such entities. This is favorable to the transparency dictum: assuming the locution is novel, then the speaker has never been exposed to a VP of the form *[carve NP NP]*; but she has been exposed to the verb *carve* (with just one object) and to the ditransitive construction. But on the lexical approach there is a unary branching node with the phonology *carve* and a 3-place valence structure. In Section 5.3 we provide evidence for the lexical approach, after first reviewing a lexical account of such novel utterances in Section 5.2. This evidence undercuts a usage-based theory that entails the opposite conclusion.

5.2 Coercion

As noted already, researchers working with plugging proposals usually take coercion as showing the usefulness of phrasal constructions. For instance, Anatol Stefanowitsch (Lecture in the lecture series *Algorithmen und Muster — Strukturen in der Sprache*, 2009) discussed the example in (20):

(20) Das Tor zur Welt Hrnglb öffnete sich ohne Vorwarnung und verschlang [sie] … die Welt Hrnglb wird von Magiern erschaffen, die Träume zu Realität formen können, aber nicht in der Lage sind zu träumen. Haltet aus, Freunde. Und ihr da draußen, bitte träumt ihnen ein Tor.

The crucial part is *ihr träumt ihnen ein Tor* (‘Dream a gate for them’). In this fantasy context the word *träumen*, which is intransitive, if forced into the ditransitive construction and therefore gets a certain meaning. This forcing of a verb

---

12 The former examples are from Tomasello, 2003, p. 160.

corresponds to overwriting properties of the verb properties by the phrasal construction.

In cases in which the plugging proposals assume that information is overwritten or extended, lexical approaches assume mediating lexical rules. Briscoe and Copestake (1999, Section 4) have worked out a lexical approach in detail.\textsuperscript{14} They discuss the ditransitive sentences in (21), which either correspond to the prototypical ditransitive construction (21a) or deviate from it in various ways.

\begin{equation}
(21) \quad \begin{array}{ll}
\text{a.} & \text{Mary gave Joe a present.} \\
\text{b.} & \text{Joe painted Sally a picture.} \\
\text{c.} & \text{Mary promised Joe a new car.} \\
\text{d.} & \text{He tipped Bill two pounds.} \\
\text{e.} & \text{The medicine brought him relief.} \\
\text{f.} & \text{The music lent the party a festive air.} \\
\text{g.} & \text{Jo gave Bob a punch.} \\
\text{h.} & \text{He blew his wife a kiss.} \\
\text{i.} & \text{She smiled herself an upgrade.} \\
\end{array}
\end{equation}

For the non-canonical examples they assume lexical rules that relate transitive and intransitive verbs to ditransitive ones and contribute the respective semantic information or the respective metaphorical extension. The example in (21i) is rather similar to the träumen example discussed above and is also analyzed with a lexical rule (page 509). Briscoe and Copestake note that this lexical rule is much more restricted in productivity than other lexical rules that were suggested by them. They take this as motivation for developing a representational format in which lexical items (including those that are derived by lexical rules) are associated with probabilities, so that differences in productivity of various patterns can be captured.

Looking narrowly at such cases, it is hard to see any rational grounds for choosing between the phrasal analysis and the lexical rule. But if we broaden our view, the lexical rule approach can be seen to have much wider application. Coercion is a very general pragmatic process, occurring in many contexts where no construction seems to be responsible (Nunberg, 1995). Nunberg cites many cases such as the restaurant waiter asking Who is the ham sandwich? (Nunberg, 1995, p. 115). Copestake and Briscoe (1992, p. 116) discuss the conversion of animal terms to mass nouns. Example (22) is about a substance, not about a cute bunny.

\begin{equation}
(22) \quad \text{After several lorries had run over the body, there was rabbit splattered all over the road.}
\end{equation}

\textsuperscript{14}Kay (2005), working in the framework of CxG, also suggests unary constructions.
The authors suggest a lexical rule that maps a count noun onto a mass noun. This analysis is also assumed by Fillmore (1999, p. 114–115). Such coercion can occur without any syntactic context: one can answer the question *What's that stuff on the road?* or *What are you eating?* with the one-word utterance *Rabbit*. Some coercion happens to affect the complement structure of a verb, but this is simply a special case of a more general phenomenon that has been analyzed by rules of systematic polysemy.

### 5.3 Valence and Coordination

On the lexical account, the verb *paint* in (21b), for example, is lexically a 2-argument verb, while the unary branching node immediately dominating it is effectively a 3-argument verb. On the constructional view there is no such predicate seeking three arguments that dominates only the verb. Coordination provides evidence for the lexical account.

A generalization about coordination is that two constituents that have compatible syntactic properties can be coordinated and that the result of the coordination is an object that has the syntactic properties of each of the conjuncts. This is reflected by the Categorial Grammar analysis which assumes the category \((X\backslash X)/X\) for the conjunction: The conjunction takes an \(X\) to the right, an \(X\) to the left and the result is an \(X\).

For example, in (23a) we have a case of the coordination of two lexical verbs. The coordination *know and like* behaves like the coordinated simplex verbs: It takes a subject and an object. Similarly, two sentences with a missing object are coordinated in (23b) and the result is a sentence with a missing object.

(23)  

a. I know and like this record.  
b. Bagels, I like and Ellison hates.

The German examples in (24) show that the case requirement of the involved verbs have to be observed. In (24b,c) the coordinated verbs require accusative and dative respectively and since the case requirements are incompatible with unambiguously case marked nouns both of these examples are out.

(24)  

a. Ich kenne und unterstütze diesen Mann.  
   I know and support this man.ACC  
b. *Ich kenne und helfe diesen Mann.  
   I know and help this man.ACC  
c. *Ich kenne und helfe diesem Mann.  
   I know and help this man.DAT

Interestingly, it is possible to coordinate basic ditransitive verbs with verbs that have additional arguments licensed by the lexical rule. (25)–(26) provide some examples:

(25)  
(a) My sisters just baked and gave me a nutella cupcake with mint chocolate chip ice-cream in the middle and milk chocolate frosting on top.\(^{15}\)
(b) Randy baked and gave me a loaf of bread!!!\(^{16}\)

(26)  
(a) I’ve never had a problem if I want extra sauce or veggie rice, in fact the last time I was in, they offered and made me a special sauce for my crab rangoon.\(^{17}\)
(b) She then offered and made me a wonderful espresso — nice.\(^{18}\)
(c) I always passed and greeted a woman working in her front yard along the way, and one day in October, when it was absolutely disgustingly hot and humid outside, she offered and made me a glass of iced tea.\(^{19}\)
(d) I offered and made him a coffee, white with one and he set to work.\(^{20}\)

Note that in (25) the monotransitive verb precedes the ditransitive verb, while in (26) they appear in the opposite order. These sentences show that both verbs are 3-argument verbs at the \(V_0\) level, since they involve \(V_0\) coordination:

(27)  
(a) \([baked and gave]_{V_0} [me]_{DP} [a cake]_{DP}\)
(b) \([offered and made]_{V_0} [me]_{DP} [a special sauce]_{DP}\)

This is expected under the lexical rule analysis but not the non-lexical constructional one.

One might wonder whether these sentences could be instances of Right Node Raising (RNR) out of coordinated VPs:

(28)  
(a) \([baked \_\_]\) and \([gave me \_\_]\) a cake
(b) \([offered \_\_]\) and \([made me \_\_]\) a special sauce

But there are problems with assuming RNR analyses of all such cases, to the exclusion of \(V_0\) coordination. First, note that under such an analysis the first verb has been used without a benefactive or recipient object. But the verb \textit{offer}...
without the recipient ("She offered a special sauce") is somewhat more awkward than the sentences in (26). Also, extraction can be used to distinguish RNR from $V^0$ coordination. By way of background, note that extrapoosed constituents are islands to extraction (Ross, 1974; Bresnan, 1976, 1994):

(29)  
  a. I located a picture of you in the police files.
  b. I located __ in the police files a picture of you.
  c. I consider arguing with women who ride motorcycles silly.
  d. I consider __ silly arguing with women who ride motorcycles.

(30)  
  a. Guess who I located a picture of __ in the police files?
  b. * Guess who I located in the police files a picture of __?
  c. It’s you who(m) I consider arguing with __ silly.
  d. * It’s you who(m) I consider silly arguing with __.

(Examples from Bresnan, 1994, p. 88, examples (49)–(50).) The acceptable extraction in (31b) shows that (31a–b) has a non-RNR analysis. We assume the coordinate $V^0$ offered and drew simply takes two objects.

(31)  
  a. I offered and drew her a picture of an elephant dancing in a tutu.
  b. Guess what I offered and drew her a picture of __?
  c. I offered her __, and Bill drew her __, a picture of an elephant dancing in a tutu.
  d. * Guess what I offered her and Bill drew her a picture of __?

By contrast, (31c) has only a RNR analysis so extraction is not possible. In addition, it is not possible to do RNR with an unstressed pronoun, eliminating such an analysis of examples like the following:

(32)  
Someone make and give me it.\(^{21}\)

Finally, a RNR analysis is not available for the German counterpart in (33):

(33)  
ich hab ihr jetzt diese Ladung Muffins mit den Herzchen drauf
    I have her now this load Muffins with the little.heart there.on gebacken und gegeben.\(^{22}\)
    backed and given
    ‘I have now baked and given her this load of Muffins with the little heart on top.’

\(^{21}\)http://www.penmai.com/forums/vegetarian-recipes/1559-aviyal.html
\(^{22}\)http://www.musiker-board.de/diverses-ot/35977-die-liebe-637-print.html. 08.06.2012

5.3 Valence and Coordination

Summarizing the coordination argument: If it is the verbs that are coordinated and if the coordinated items must have compatible syntactic properties like valence, this means that *gebacken* (‘baked’) and *gegeben* (‘given’) have the same valence properties. This is accounted for in a lexical approach since in the lexical approach the transitive version of the creation verb *bake* licenses a ditransitive verb which can be coordinated with *give*. In the phrasal approach however, the verb *bake* has two argument roles and is not compatible with the verb *give*, which has three argument roles. In the phrasal model, *bake* can only realize three arguments when it enters the ditransitive argument structure construction, but in sentences like (33) it is not *bake* alone that enters the phrasal syntax, but rather the combination of *baked* and *given*. But the verbs are incompatible as far as the semantic roles are concerned.

To fix this under the phrasal approach, one could posit a mechanism such that the semantic roles that are required for the coordinate phrase *baked and given* are shared by each of its conjunct verbs and that they are therefore compatible. But this would amount in saying that there are several verb senses for *baked*, something that the anti-lexicalists claim to avoid, as discussed in the next section.

The coordination facts illustrate a more general point: the output of a lexical rule is just a word (an $X^0$), so it has the same syntactic distribution as an underived word with the same category and valence feature. This important generalization follows from the lexical account while on the phrasal view it is at best mysterious. The point can be shown with any of the lexical rules that the anti-lexicalists are so keen to eliminate in favor of phrasal constructions. For example, the active and passive verbs can be coordinated, as long as they have the same valence properties, as in these Swedish examples:

(34) Han **dog och glömde-s bort.**
    He died and forget-PASS away
    ‘He died and was forgotten.’

(35) Vi **beställde och serverade-s en bra Cheers chowder till att börja med,**
    we ordered and served-PASS a good Cheers chowder to start with,
    och sedan en stor hummer varje.\(^23\)
    and then a big lobster each
    ‘We ordered and were served a good Cheers chowder to start with, and then a big lobster each.’

(English works the same way, as shown by the grammatical translation lines, but the periphrastic form of the passive complicates issues slightly.) In (34) the active and passive are both intransitive, while in (35) the passive of the ditransitive

verb servera ‘serve’ retains one object, so it is effectively transitive and can be coordinated with the active transitive beställa ‘order’.

Moreover, the English passive verb form, being a participle, can feed a second lexical rule deriving adjectives from verbs (see Figure 3 above). All categories of English participles can be converted to adjectives:

(36) a. active present participles (cp. The leaf is falling): the falling leaf
    b. active past participles (cp. The leaf has fallen): the fallen leaf
    c. passive participles (The toy is being broken (by the child)): the broken toy

That the derived forms are adjectives, not verbs, is shown by a host of properties, including negative un- prefixation: unbroken means ‘not broken’, just as unkind means ‘not kind’, while the un- appearing on verbs indicates, not negation, but action reversal, as in untie (Bresnan, 2001, Chapter 3). Predicate adjectives preserve the subject of predication of the verb and for prenominal adjectives the role is simply that the role that would be assigned to the subject goes to the modified noun instead (The toy remained (un-|)broken; the broken toy). Being an A0, such a form can be coordinated with another A0, as in the following:

(37) a. The suspect should be considered [armed and dangerous].
    b. any [old, rotting, or broken] toys

In (37b), three adjectives are coordinated, one underived (old), one derived from a present participle (rotting), and one from a passive participle (broken). Such coordination is completely mundane on a lexical theory. Each A0 conjunct has a valence feature (in HPSG it would be the SPR feature for predicates or the MOD feature for the prenominal modifiers), which is shared with the mother node of the coordinate structure. But the point of the phrasal (or ASC) theory is to deny that words have such valence features.

The claim that lexical derivation of valence structure is distinct from phrasal combination is further supported with evidence from deverbal nominalization (Wechsler, 2008b). To derive nouns from verbs, -ing suffixation productively applies to all declinable verbs (the shooting of the prisoner), while morphological productivity is severely limited for various other suffixes such as -(a)tion (*the shootation of the prisoner). So forms such as destruction and distribution must be retrieved from memory while -ing nouns such as looting or growing could be (and in the case of rare verbs or neologisms, must be) derived from the verb or the root through the application of a rule (Zucchi, 1993). This difference explains why ing-nominals always retain the argument structure of the cognate verb, while other forms show some variation. A famous example is the lack of the agent argument for the noun growth versus its retention by the noun growing: *John’s
growth of tomatoes versus John's growing of tomatoes. However, on Marantz’s (1997) phrasal analysis, a phrasal construction (notated as vP) is responsible for assigning the agent role of growing. For him, none of the words directly select agents, but -ing forms are permitted to appear in the agent-licensing vP construction, while other nouns cannot. A problem for Marantz is that these two types of nouns can coordinate and share dependents (examples (38a,b) are from Wechsler, 2008b, Section 7):

(38)  

a. With nothing left after the soldier’s [destruction and looting] of their home, they reboarded their coach and set out for the port of Calais.  

b. Anyone with information in relation to the [growing or distribution] of cannabis is urged to contact Crime Stoppers on 1 800 333 0000.  

c. I believe it is time in the USA voting population to have the opportunity to vote on adding an amendment to the Bill of Rights to legalize the [use, growth and selling] of marijuana.

On the phrasal analysis, the nouns looting, growing, and selling occur in one type of syntactic environment (namely vP), while forms destruction, distribution, use, and growth occur in a different syntactic environment. This places contradictory demands on the structure of sentences like (38). As far as we know, neither this problem nor the others raised by Wechsler (2008b) have even been addressed by advocates of the phrasal theory of argument structure.

Consider one last example. In an influential phrasal analysis, Hale and Keyser (1993) derived denominal verbs like to saddle through noun incorporation out of a structure akin to [PUT a saddle ON x]. Again, verbs with this putative derivation routinely coordinate and share dependents with verbs of other types:

(39) Realizing the dire results of such a capture and that he was the only one to prevent it, he quickly [saddled and mounted] his trusted horse and with a grim determination began a journey that would become legendary.

As in all of these \(X^0\) coordination cases, under the phrasal analysis the two verbs place contradictory demands on a single phrase structure.

A lexical valence structure is an abstraction or generalization over various occurrences of the verb in syntactic contexts. To be sure, one key use of that valence structure is simply to indicate what sort of phrases the verb must (or can) combine with, and the result of semantic composition; if that were the whole story then the phrasal theory would be viable. But it is not. As it turns out, this lexical valence
structure, once abstracted, can alternatively be used in other ways: among other possibilities, the verb (crucially including its valence structure) can be coordinated with other verbs that have a similar valence structure; or it can serve as the input to lexical rules specifying a new word bearing a systematic relation to the input word. The coordination and lexical derivation facts follow from the lexical view, while the phrasal theory at best leaves these facts as mysterious and at worst leads to irreconcilable contradictions for the phrase structure.

5.4 Valence and Derivational Morphology

Goldberg and Jackendoff (2004), Alsina (1996), and Asudeh, Dalrymple and Toivonen (2008) suggest analyzing resultative constructions and/or caused motion constructions as phrasal constructions. As was argued in Müller, 2006 this is incompatible with the assumption of Lexical Integrity, that is, that word formation happens before syntax (Bresnan and Mchombo, 1995). Let us consider a concrete example, such as (40):

(40) a. Er tanzt die Schuhe blutig / in Stücke.
    he dances the shoes bloody into pieces

b. die in Stücke / blutig getanzt Schuhe
    the into pieces bloody danced shoes

c. * die getanzt Schuhe
    the danced shoes

The shoes are not a semantic argument of tanzt. Nevertheless the NP that is realized as accusative NP in (40a) is the element the adjectival participle in (40b) predicates over. Adjectival participles like the one in (40b) are derived from a passive participle of a verb that governs an accusative object. If the accusative object is licensed phrasally by configurations like the one in (40a) it cannot be explained why the participle getanzt can be formed despite the absence of an accusative object. See Müller, 2006, Section 5 for further examples of the interaction of resultatives and morphology. The conclusion, which was drawn in the late 70s and early 80s by Dowty (1978, p. 412) and Bresnan (1982c, p. 21), is that phenomena that feed morphology should be treated lexically. The natural analysis in frameworks like HPSG, CG, CxG, and LFG is therefore one that assumes a lexical rule for the licensing of resultative constructions. See Verspoor, 1997; Wechsler, 1997, Wechsler and Noh, 2001, Müller, 2002a, Kay, 2005, and Simpson, 1983 for lexical proposals in the some of these frameworks.

This argument is similar to the one that was discussed in conection with the GPSG representation of valence in Section 4.2: morphological processes have to be able to see the valence of the element they attach to. This is not the case if arguments are introduced by phrasal configurations after the morphology level.

5.5 Simplicity and polysemy

Much of the intuitive appeal of the plugging approach stems from its apparent simplicity relative to the use of lexical rules. But the claim to greater simplicity for Construction Grammar is based on misunderstandings of both lexical rules and Construction Grammar (specifically of Goldberg’s (1995, 2006) version). It draws the distinction in the wrong place and misses the real differences between these approaches. This argument from simplicity is often repeated and so it is important to understand why it is incorrect.

Tomasello (2003) presents the argument as follows. Discussing first the lexical rules approach, Tomasello (2003, p. 160) writes that

One implication of this view is that a verb must have listed in the lexicon a different meaning for virtually every different construction in which it participates [...]. For example, while the prototypical meaning of *cough* involves only one participant, the cougher, we may say such things as *He coughed her his cold*, in which there are three core participants. In the lexical rules approach, in order to produce this utterance the child’s lexicon must have as an entry a ditransitive meaning for the verb *cough*. (Tomasello, 2003, p. 160)

Tomasello (2003, p. 160) then contrasts a Construction Grammar approach, citing Fillmore et al. (1988), Goldberg (1995), and Croft (2001). He concludes as follows:

The main point is that if we grant that constructions may have meaning of their own, in relative independence of the lexical items involved, then we do not need to populate the lexicon with all kinds of implausible meanings for each of the verbs we use in everyday life. The construction grammar approach in which constructions have meanings is therefore both much simpler and much more plausible than the lexical rules approach. (Tomasello, 2003, p. 161)

This reflects a misunderstanding of lexical rules, as they are normally understood. There is no implausible sense populating the lexicon. The lexical rule approach to *He coughed her his cold* states that when the word *coughed* appears with two objects, the whole complex has a certain meaning. See Müller (2006, p. 876).

The simplicity argument also relies on a misunderstanding of a theory he advocates, namely the theory due to Goldberg (1995, 2006). For his argument to go through, Tomasello must tacitly assume that verbs can combine freely with constructions, that is, that the grammar does not place extrinsic constraints on such combinations. If it is necessary to also stipulate which verbs can appear in which...
constructions then the claim to greater simplicity collapses: each variant lexical item with its ‘implausible meaning’ under the lexical rule approach corresponds to a verb-plus-construction combination under the phrasal approach.

Passages such as the following may suggest that verbs and constructions are assumed to combine freely:

Constructions are combined freely to form actual expressions as long as they can be construed as not being in conflict (invoking the notion of construal is intended to allow for processes of accommodation or coercion). (Goldberg, 2006, p. 22)

Allowing constructions to combine freely as long as there are no conflicts, allows for the infinitely creative potential of language. [...] That is, a speaker is free to creatively combine constructions as long as constructions exist in the language that can be combined suitably to categorize the target message, given that there is no conflict among the constructions. (Goldberg, 2006, p. 22)

But in fact Goldberg does not assume free combination, but rather that a verb is ‘conventionally associated with a construction’ (Goldberg, 1995, p. 50): verbs specify their participant roles and which of those are obligatory direct arguments (profiled, in Goldberg’s terminology; see Section 3). In fact Goldberg herself (2006, p. 211) argues against Borer’s putative assumption of free combination 2003 on the grounds that Borer is unable to account for the difference between dine (intransitive), eat (optionally transitive), and devour (obligatorily transitive). Indeed, it is well-established that not all realizational patterns are determined by the meaning of the verb. In addition to lexically specified transitivity, preposition and oblique case selection is sometimes loosely semantically motivated but nonetheless arbitrarily associated with certain verbs: rely on/*in me versus trust in/*on me, and so on. Despite Tomasello’s comment above, Construction Grammar is no simpler than the lexical rules.

The resultative construction is often used to illustrate the simplicity argument. For example, Goldberg (1995, Chapter 7) assumes that the same lexical item for the verb sneeze is used in (41a) and (41b). It is simply inserted into different constructions:

(41) a. He sneezed.

The context of these quotes makes clear that ‘combining constructions’ includes combining a verb with an argument structure construction, since the verb itself is considered a construction. See Goldberg (2006, p. 21, ex. (2)).

Goldberg’s critique cites a 2001 presentation by Borer with the same title as Borer (2003). See Section 6.3 for more discussion of Borer’s theory.

b. He sneezed the napkin off the table.

The meaning of (41a) corresponds more or less to the verb meaning, since the verb is used in the Intransitive Construction. But the Caused-Motion Construction in (41b) contributes additional semantic information concerning the causation and movement: His sneezing caused the napkin to move off the table. *sneeze* is plugged into the Caused Motion Construction, which licenses the subject of *sneeze* and additionally provides two slots: one for the theme (*napkin*) and one for the goal (*off the table*).

In a nuanced comparison of the two approaches, Goldberg (1995, p. 139ff) considers again the added recipient argument in *Mary kicked Joe the ball*, where *kick* is lexically a 2-place verb. She notes that on the constructional view, ‘the composite fused structure involving both verb and construction is stored in memory’. Since this is recognized as a composite structure, the verb itself retains its original meaning as a 2-place verb, so that ‘we avoid implausible verb senses such as “to cause to receive by kicking”’; the idea seems to be that the lexical approach, in contrast, must countenance such implausible verb senses since a lexical rule adds a third argument.

But the lexical and constructional approaches are actually indistinguishable on this point. The lexical rule does not produce a verb with the ‘implausible sense’ in (42a), but rather the one in (42b):

(42)  
\begin{align*}
\text{a. } & \text{cause-to-receive-by-kicking}(x, y, z) \\
\text{b. } & \text{cause(kick}(x, y), (receive(y,z))
\end{align*}

In short, the same sort of ‘the composite fused structure’ is assumed under either view. The particular conception of lexical rules as unary branching structures actually encodes this in the phrase structure, but decomposition is assumed under other lexical approaches as well. With respect to the semantic structure, the number and plausibility of senses, and the polyadicity of the semantic relations, the two theories are identical. They differ in one respect: on the lexical theory, the derived valence structure, which includes places for three participants, is associated with the phonological string *kicked*. We saw evidence for that claim in Section 5.3 above.

### 6 Radical Underspecification: The End of Argument Structure?

In the last section we examined proposals that assume that verbs come with certain argument roles and are inserted into prespecified structures that may contribute additional arguments. While we showed that this is not without problems, there
are even more radical proposals that the construction adds all agent arguments, or even all arguments. The notion that the agent argument should be severed from its verbs is put forth by Marantz (1984, 1997), Kratzer (1996), Embick (2004) and others. Others suggest that no arguments are selected by the verb. Borer (2003) calls such proposals exoskeletal since the structure of the clause is not determined by the predicate, that is, the verb does not project an inner ‘skeleton’ of the clause. Counter to such proposals are endoskeletal approaches, in which the structure of the clause is determined by the predicate, that is, lexical proposals. The radical exoskeletal proposals are mainly proposed in Mainstream Generative Grammar (Borer, 1994, 2003, 2005; Schein, 1993; Hale and Keyser, 1997; Lohndal, 2012) but can also be found in HPSG (Haugereid, 2009). We will not discuss these proposals in detail here, but we review the main issues insofar as they relate to the question of lexical argument structure. We conclude that the available empirical evidence favors the lexical argument structure approach over such alternatives.

Davidson (1967) argued for an event variable in the logical form of action sentences (43a). Dowty (1989) coined the term Neodavidsonian for the variant in (43b), in which the verb translates to a property of events, and the subject and complement dependents are translated as arguments of secondary predicates such as agent and theme. (Dowty (1989) called the system in 43a an ordered argument system.) Kratzer (1996) further noted the possibility of mixed accounts such as (43c), in which the agent (subject) argument is severed from the kill relation, but the theme (object) remains an argument of the kill relation.

(43)  
  a. kill: $\lambda y \lambda x \exists e [\text{kill}(e, x, y)] \quad \text{(Davidsonian)}$
  b. kill: $\lambda y \lambda x \exists e [\text{kill}(e) \land \text{agent}(e, x) \land \text{theme}(e, y)]$
      \quad \text{(Neodavidsonian)}
  c. kill: $\lambda y \lambda x \exists e [\text{kill}(e, y) \land \text{agent}(e, x)] \quad \text{(mixed)}$

Kratzer (1996) observed that a distinction between Davidsonian, Neodavidsonian and mixed can be made either ‘in the syntax’ or ‘in the conceptual structure’ (Kratzer, 1996, p. 110-1). For example, on a lexical approach of the sort we advocate here, any of the three alternatives in 43 could be posited as the semantic content of the verb kill. A lexical entry for kill on the mixed model appears in 44.

(44)  
\[
\begin{array}{c}
\text{PHON} \\
\text{ARG-ST} \\
\text{CONTENT}
\end{array}
\begin{array}{c}
\langle \text{kill} \rangle \\
\langle \text{NP}_{\text{NP}} , \text{NP}_{\text{NP}} \rangle \\
\text{kill}(e, y) \land \text{agent}(e, x)
\end{array}
\]

In other words, the lexical approach is neutral on this question of the ‘conceptual structure’ of eventualities, as noted already in a different connection in Section 5.5. For that reason, certain semantic arguments for the Neodavidsonian approach,
6.1 Little v and idiom asymmetries

such as those put forth by Schein (1993, Chapter 4) and Lohndal (2012), do not directly bear upon the issue of lexicalism, as far as we can tell.

But Kratzer (1996), among others, has gone further and argued for an account that is neodavidsonian (or rather, mixed) ‘in the syntax’. Kratzer’s claim is that the verb specifies only the internal argument(s), as in (45a) or (45b), while the agent (external argument) role is assigned by the phrasal structure. On the ‘neodavidsonian in the syntax’ view, the lexical representation of the verb has no arguments at all, except the event variable, as shown in as in (45c).

(45)  

a. \( \text{kill}: \lambda y \exists e[\text{kill}(e, y)] \) (agent is severed)  
b. \( \text{kill}: \lambda y \exists e[\text{kill}(e) \wedge \text{theme}(e, y)] \) (agent is severed)  
c. \( \exists e[\text{kill}(e, y)] \) (all arguments severed)

On such accounts, the remaining dependents of the verb receive their semantic roles from silent secondary predicates, which are usually assumed to occupy the positions of functional heads in the phrase structure. A standard term for the agent-assigning silent predicate is ‘little \( v \)’. These extra-lexical dependents are the analogues of the ones contributed by the constructions in Construction Grammar.

In the following subsections we address arguments that have put forth in favor of the ‘little \( v \)’ hypothesis, from idiom asymmetries (Section 6.1) and deverbal nominals (Section 6.2). We argue that the evidence actually favors the lexical view. Then we turn to problems for exoskeletal approaches, from idiosyncratic syntactic selection (Section 6.3) and expletives (Section 6.4). We conclude with a look at the exoskeletal alternative (Section 6.5), and a summary (Section 6.6).

6.1 Little v and idiom asymmetries

Marantz (1984) and Kratzer (1996) argued for severing the agent from the argument structure as in 45a, on the basis of putative idiom asymmetries. Marantz (1984) observed that while English has many idioms and specialized meanings for verbs in which the internal argument is the fixed part of the idiom and the external argument is free, the reverse situation is considerably rarer. To put it differently, the nature of the role played by the subject argument often depends on the filler of the object position, but not vice versa. To take Kratzer’s examples (Kratzer, 1996, p. 114):

(46)  

a. kill a cockroach  
b. kill a conversation  
c. kill an evening watching TV  
d. kill a bottle (i.e. empty it)  
e. kill an audience (i.e., wow them)

On the other hand, one does not often find special meanings of a verb associated with the choice of subject, leaving the object position open (examples from Marantz (1984, p. 26)):

(47)  
   a. Harry killed NP.  
   b. Everyone is always killing NP.  
   c. The drunk refused to kill NP.  
   d. Silence certainly can kill NP.

Kratzer observes that a mixed representation of \textit{kill} as in (48a) allows us to specify varying meanings that depend upon its sole NP argument.

(48)  
   a. \textit{kill} : \lambda y \exists e [\text{kill}(e, y)]  
   b. If \( a \) is a time interval, then \( \text{kill}(a, e) = \text{truth} \) if \( e \) is an event of wasting \( a \)  
      If \( a \) is animate, then \( \text{kill}(a, e) = \text{truth} \) if \( e \) is an event in which \( a \) dies  
      ... etc.

On the polyadic (Davidsonian) theory, the meaning could similarly be made to depend upon the filler of the agent role. On the polyadic view, ‘there is no technical obstacle’ (Kratzer, 1996, p. 116) to conditions like those in 48b, except reversed, so that it is the filler of the agent role instead of the theme role that affects the meaning. But, she writes, this could not be done if the agent is not an argument of the verb. According to Kratzer, the agent-severed representation (such as 48a) disallows similar constraints on the meaning that depend upon the agent, thereby capturing the idiom asymmetry.

But as noted by Wechsler (2005), ‘there is no technical obstacle’ to specifying agent-dependent meanings even if the Agent has been severed from the verb as Kratzer proposes. It is true that there is no variable for the agent in 48a. But there is an event variable \( e \), and the language user must be able to identify the agent of \( e \) in order to interpret the sentence. So one could replace the variable \( a \) with ‘the agent of \( e \)’ in the expressions in 48b, and thereby create verbs that violate the idiom asymmetry.

While this may seem to be a narrow technical or even pedantic point, it is nonetheless crucial. For Kratzer’s argument to go through, it requires an additional assumption: that modulations in the meaning of a polysemous verb can only depend upon arguments of the relation denoted by that verb, and not on other participants in the event. But under that additional assumption, it makes no difference whether the agent is severed from the lexical entry or not. For example, consider the following (mixed) neodavidsonian representation of the semantic content in the lexical entry of \textit{kill}:

(49) \textit{kill} : \lambda y \lambda x \exists e [\text{kill}(e, y) \land \text{agent}(e, x)]

Assuming that sense modulations can only be affected by arguments of the \textit{kill}(e,y) relation, we derive the idiom asymmetry, even if (49) is the lexical entry for \textit{kill}.

Moreover, recasting Kratzer’s account in lexicalist terms allows for verbs to vary. This is an important advantage, because the putative asymmetry is only a tendency. Following are examples in which the subject is a fixed part of the idiom and there are open slots for non-subjects:

\begin{enumerate}
  \item \(a\). A little bird told X that S.  
  \(\text{‘X heard the rumor that S’}\) (Nunberg et al., 1994, p. 526)
  \item \(b\). The cat’s got x’s tongue.  
  \(\text{‘X cannot speak.’} \) (Bresnan, 1982a, p. 349–350)
  \item \(c\). What’s eating x?  
  \(\text{‘Why is X so galled?’} \) (Bresnan, 1982a, p. 349–350)
\end{enumerate}

Further data and discussion of subject idioms in English and German can be found in the Appendix below.

The tendency towards a subject-object asymmetry plausibly has an independent explanation. Nunberg et al. (1994) argue that the subject-object asymmetry is a side-effect of an animacy asymmetry. The open positions of idioms tend to be animate while the fixed positions tend to be inanimate. Nunberg et al. (1994) derive these animacy generalizations from the figurative and proverbial nature of the metaphorical transfers that give rise to idioms. If there is an independent explanation for this tendency, then a lexicalist grammar successfully encodes those patterns, perhaps with a mixed neodavidsonian lexical decomposition, as explained above (see Wechsler (2005) for such a lexical account of the verbs \textit{buy} and \textit{sell}). But the ‘little v’ hypothesis rigidly predicts this asymmetry for all agentive verbs, and that prediction is not borne out.

\section{Deverbal nominals}

An influential argument against lexical argument structure involves English deverbal nominals and the causative alternation. It originates from a mention in Chomsky (1970), and is developed in detail by Marantz (1997); see also Pesetsky (1996) and Harley and Noyer (2000). The argument is often repeated, but it turns out that the empirical basis of the argument is incorrect, and the actual facts point in the opposite direction, in favor of lexical argument structure (Wechsler, 2008a,b).

Certain English causative alternation verbs allow optional omission of the agent argument (51), while the cognate nominal disallows expression of the agent (52):

\begin{enumerate}
  \item \(a\). that John grows tomatoes
\end{enumerate}
b. that tomatoes grow

(52) a. *John’s growth of tomatoes
   b. the tomatoes’ growth, the growth of the tomatoes

In contrast, nominals derived from obligatorily transitive verbs such as destroy allow expression of the agent, as shown in (54a):

(53) a. that the army destroyed the city
    b. *that the city destroyed

(54) a. the army’s destruction of the city
    b. the city’s destruction

Following a suggestion by Chomsky (1970), Marantz (1997) argued on the basis of these data that the agent role is lacking from lexical entries. In verbal projections (51) and (53) the agent role is assigned in the syntax by little v. Nominal projections like (52) and (54) lack little v. Instead, pragmatics takes over to determine which agents can be expressed by the possessive phrase: the possessive can express ‘the sort of agent implied by an event with an external rather than an internal cause’ because only the former can be ‘easily reconstructed’ (quoted from Marantz (1997); see also Harley and Noyer (2000)). The destruction of a city has a cause external to the city, while the growth of tomatoes is internally caused by the tomatoes themselves (Smith, 1970). Marantz points out that this explanation is unavailable if the noun is derived from a verb with an argument structure specifying its agent, since the deverbal nominal would inherit the agent of a causative alternation verb.

The empirical basis for this argument is the putative mismatch between the allowability of agent arguments, across some verb-noun cognate pairs: e.g. grow allows the agent but growth does not. But it turns out that the grow/growth pattern is rare. Most deverbal nominals precisely parallel the cognate verb: if the verb has an agent, so does the noun. Moreover, there is a ready explanation for the exceptional cases that exhibit the grow/growth pattern (Wechsler, 2008b). First consider non-alternating theme-only intransitives (‘unaccusatives’), as in (55) and non-alternating transitives as in (56). The pattern is clear: if the verb is agentless, then so is the noun:

(55) arriv(al), disappear(ance), fall, etc.:
   a. A letter arrived.
   b. the arrival of the letter
   c. * The mailman arrived a letter.
   d. * the mailman’s arrival of the letter

6.2 Deverbal nominals

(56)  

\textit{destroy/destruction, construct(ion), creat(ion), assign(ment) etc.}:

\begin{enumerate}
\item The army is destroying the city.
\item the army’s destruction of the city
\end{enumerate}

This favors the view that the noun inherits the lexical argument structure of the verb. For the anti-lexicalist, the badness of (55c) and (55d), respectively, would have to receive independent explanations. For example, on Harley and Noyer’s 2000 proposal, (55c) is disallowed because a feature of the root ARRIVE prevents it from appearing in the context of v, but the badness of (55d) is instead ruled out because the cause of an event of arrival cannot be easily reconstructed from world knowledge. This exact duplication in two separate components of the linguistic system would have to be replicated across all non-alternating intransitive and transitive verbs, a situation that is highly implausible.

Turning to causative alternation verbs, Marantz’s argument is based on the implicit generalization that noun cognates of causative alternation verbs (typically) lack the agent argument. But apart from the one example of \textit{grow/growth}, there do not seem to be any clear cases of this pattern. Besides \textit{grow(th)}, Chomsky 1970, examples 7c and 8c cited two experiencer predicates, \textit{amuse} and \textit{interest}: \textit{John amused (interested) the children with his stories} versus \textbf{*John’s amusement (interest) of the children with his stories}. But this was later shown by Rappaport 1983 and Dowty 1989 to have an independent aspectual explanation. Deverbal experiencer nouns like \textit{amusement} and \textit{interest} typically denote a mental state, where the corresponding verb denotes an event in which such a mental state comes about or is caused. These result nominals lack not only the agent but all the eventive arguments of the verb, because they do not refer to events. Exactly to the extent that such nouns can be construed as representing events, expression of the agent becomes acceptable.

In a response to Chomsky 1970, Carlota Smith (1972) surveyed Webster’s dictionary and found no support for Chomsky’s claim that deverbal nominals do not inherit agent arguments from causative alternation verbs. She listed many counterexamples, including ‘\textit{explode, divide, accelerate, expand, repeat, neutralize, conclude, unify}, and so on at length.’ (Smith, 1972, p. 137). Harley and Noyer (2000) also noted many so-called ‘exceptions’: \textit{explode, accumulate, separate, unify, disperse, transform, dissolve/dissolution, detach(ment), disengage-(ment),…}. The simple fact is that these are not exceptions because there is no generalization to which they can be exceptions. These long lists of verbs represent the norm, especially for suffix-derived nominals (in \textit{-tion, -ment}, etc.). Many zero-derived nominals from alternating verbs also allow the agent, such as \textit{change, release, and use: My constant change of mentors from 1992-1997. The frequent release of the prisoners by the governor. The frequent use of sharp tools by underage children}. (examples from Borer (2003, fn. 13)). Pesetsky (1996, p. 79, ex.
assigns a star to the thief’s return of the money, but it is acceptable to many speakers, the Oxford English Dictionary lists a transitive sense for the noun return (definition 11a), and corpus examples like her return of the spoils are not hard to find.

Like the experiencer nouns mentioned above, many zero-derived nominals lack event readings. Some reject all the arguments of the corresponding eventive verb, not just the agent: *the freeze of the water, *the break of the window, and so on. In the judgment of the second author, his drop of the ball is slightly odd, but the drop of the ball has exactly the same degree of oddness. The locution a drop in temperature matches the verbal one The temperature dropped, and both verbal and nominal forms disallow the agent: *The storm dropped the temperature. *the storm’s drop of the temperature. In short, the facts seem to point in exactly the opposite direction from what has been assumed in this oft-repeated argument against lexical valence. Apart from the one isolated case of grow(th), event-denoting de-verbal nominals match their cognate verbs in their argument patterns.

Turning to grow(th), we find a simple explanation for its unusual behavior. When the noun growth entered the English language, causative (transitive) grow did not exist. The OED provides these dates of the earliest attestations of grow and growth:

(57) a. intransitive grow: c725 ‘be verdant’ ... ‘increase’ (intransitive)
    b. the noun growth: 1587 ‘increase’ (intransitive)
    c. transitive grow: 1774 ‘cultivate (crops)’

Thus growth entered the language at a time when transitive grow did not exist. The argument structure and meaning were inherited by the noun from its source verb, and then preserved into present-day English. This makes perfect sense if, as we claim, words have predicate argument structures. Nominalization by -th suffixation is not productive in English, so growth is listed in the lexicon. To explain why growth lacks the agent we need only assume that a lexical entry’s predicate argument structure dictates whether it takes an agent argument or not. So even this one word provides evidence for lexical argument structure.

### 6.3 Idiosyncratic Syntactic Selections

As was mentioned at the beginning of this section, proponents of so-called neo-constructivist approaches assume that roots are stored in the lexicon and connected to encyclopedic knowledge that helps to determine which arguments may be or have to be present. The arguments are licenced by functional projections that may contribute meaning to the core meaning contributed by the root or in Haugereid’s
proposal by binary branching ID schemata that licence an argument that fills one of five argument roles.

The notion of lexical valence structure immediately explains why the argument realization patterns are strongly correlated with the particular lexical heads selecting those arguments. It is not sufficient to have general lexical items without valency information and let the syntax and world knowledge decide about argument realizations, because not all realizational patterns are determined by the meaning. The form of the preposition of a prepositional object is sometimes loosely semantically motivated but in other cases arbitrary. For example, the valence structure of the English verb *depend* captures the fact that it selects an *on*-PP to express one of its semantic arguments:

(58) a. John depends on Mary. (*counts, relies, etc.*)
    b. John trusts (*on) Mary.
    c. \[
       \begin{array}{c}
         \text{PHON} & \langle \text{depend} \rangle \\
         \text{ARG-ST} & \langle \text{NP} \text{on} , \text{PP}[\text{on}] \rangle \\
         \text{CONTENT} & \text{depend}(x,y)
       \end{array}
    \]

Such idiosyncratic lexical selection is utterly pervasive in human language. The verb or other predicator often determines the choice between direct and oblique morphology, and for obliques, it determines the choice of adposition or oblique case. In some languages such as Icelandic even the subject case can be selected by the verb (Zaenen et al., 1985).

Compare the German translation of the English example in (59a):

(59) a. I am waiting for my man.
     b. Ich warte auf meinen Mann.
     I wait on my man.ACC

While the English verb *wait* requires the preposition *for* (German *für*) the preposition *auf* (‘on’) with accusative is used. A learner has to acquire that *warten* has to be used with *auf* + accusative and not with other prepositions or other case.

Similarly it is often impossible to find semantic motivation for case. For instance there is a tendency to avoided the genitive in German in favour of the dative: Instead of the genitive as in (60a) one also finds examples with the dative as in (60b):

(60) a. dass der Opfer gedacht werde
     that the victims.GEN remembered was
     ‘that the victims would be remembered’
Verbs like helfen (‘to help’) and unterstützen (‘to support’) are rather similar semantically, but govern different cases:

(61)  
a. Er unterstützt ihn.
     he supports him.ACC  
b. Er hilft ihm.
     he helps him.DAT

In order to avoid that the verb helfen appears in the syntactic environment that licences (61a) and that the verb unterstützen appears in the construction that licences (61b), one has to specify the case that the respective verbs require in the lexical items of the verbs.\(^30\)

Pollard and Sag (1987, p. 126) discuss a similar example with the two verbs treffen and begegnen (‘to meet’):

(62)  
a. Er traf den Mann.
     he.NOM met the man.ACC  
b. Er begegnete dem Mann.
     he.NOM met the man.DAT

Without any semantic motivation one verb takes an accusative object and the other one takes a dative.

Proponents of neo-constructionist approaches either admit that they cannot account for such contrasts, or make proposals that are difficult to distinguish from lexical valence structures (see Section 6.5 below). In a phrasal constructionist approach one would have to assume phrasal patterns into which warten can be inserted and a representation for warten, in which it is represented that warten takes a prepositional object with auf and an accusative NP (see Kroch and Joshi, 1985, Section 5.2 for such a proposal in the framework of TAG). Since generalizations regarding verbs with such valence representations can be found one would be forced to have two inheritance hierarchies: one for lexical entries with their valency properties and another one for specific phrasal patterns that are needed for the specific constructions in which these lexical items can be used.

\(^{29}\)Frankfurter Rundschau, 07.11.1997, S. 6.  
\(^{30}\)Or at least mark the fact that unterstützen takes an object with the default case for objects and helfen takes a dative object in German. See Haider, 1985, Heinz and Matiasek, 1994, and Müller, 2001 for structural and lexical case.
A radical variant of the plugging approach is suggested by Haugereid (2009). Haugereid (pages 12–13) assumes that the syntax combines a verb with an arbitrary combination of a subset of five different argument roles. Which arguments can be combined with a verb is not restricted by the lexical item of the verb.31 Such radically underspecified proposals cannot explain the infelicity of (63):

(63) # Ich verspreche dir, das nicht zu borgen. (Max, 4;9)

Intended: ‘I promise to you not to lend it to somebody.’

(63) was uttered in a situation where it was discussed whether a toy squirrel could be taken to the kindergarten or not. The promise was to not lend the toy to anybody else. The German verb *borgen* is ambiguous in an interesting way: it has the two translations *borrow* and *lend*, which basically are two different perspectives on the same event. See Kunze, 1991, 1993 for an extensive discussion of verbs of exchange of possession. Similar pairs are *give* and *take*, *buy* and *sell*. The interesting fact to note about *borgen* is that the dative object is obligatory with the *borrow*’ reading:

(64) a. Ich borge ihm das Eichhörnchen.
    I lend him the squirrel
    ‘I lend the squirrel to him.’

b. Ich borge (mir) das Eichhörnchen.
    I borrow me the squirrel
    ‘I borrow the squirrel.’

If we omit it, we get the *lend*’ reading. So, instead of (63), Max should have uttered (65a) or (65b):

(65) a. Ich verspreche dir, das niemandem zu borgen.
    I promise you it nobody to lend

b. Ich verspreche dir, das nicht zu verborgen.
    I promise you it not to lend out

It follows that all theories have to have a place where it is fixed that certain arguments are necessary for a certain verb meaning or a certain perspective on an event, respectively.

Webelhuth (1995, p. 34) pointed out the nice minimal triplet *dine*, *devour*, and *eat*. *Dine* is obligatorily intransitive, *devour* is transitive, and *eat* can be used both intransitively and transitively. Here, we have another example of different valency frames, without there being any possibility to reduce this to a semantic contrast.

---

31Haugereid has the possibility to impose valence restrictions on verbs, but he claims that he uses this possibility just in order to get a more efficient processing of his computer implementation (p. 13).

with regard to the core meaning of the involved predicates: all three involve an eating frame.

The problem of the argument realization in the triplet *dine*, *devour*, and *eat* and other examples by Levin and Rappaport Hovav (2005) that show that certain arguments are obligatory is sometimes noted in the literature, but they are simply ignored. For instance Lohndal (2012) writes:

> An unanswered question on this story is how we ensure that the functional heads occur together with the relevant lexical items or roots. This is a general problem for the view that Case is assigned by functional heads, and I do not have anything to say about this issue here. (Lohndal, 2012)

We think that this view is illegitimate at the current state of linguistics. Getting case assignment right in simple sentences without vast overgeneration of ill-formed word sequences is a minimal requirement for a linguistic theory that is asked to be taken seriously.

### 6.4 Expletives

A final example for the irreducibility of valence to semantics are verbs that select for expletives and reflexive arguments of inherently reflexive verbs in German:

(66) a. weil *es* regnet
    because *it* rains

   b. weil *(es)* mir *(vor der Prüfung)* graut
    because *EXPL me* before the exam dreads

   c. weil *(es)* bis zum Professor bringt
    because he *EXPL until to the professor* brings
    ‘because he made it to professor’

   d. weil *(es)* sich um den Montag handelt
    because *EXPL REFL around the Monday* trades
    ‘It is about the Monday.’

   e. weil *(ich mich (jetzt)) erhole
    because *I myself now recreate*
    ‘because I recreate’

The lexical heads in (66) need to contain information about the expletive subjects/objects and/or reflexive pronouns that do not fill semantic roles. Note that German allows for subjectless predicates and hence the presence of expletive subjects cannot be claimed to follow from general principles. (66c) is an example with an expletive object. Explanations referring to the obligatory presence of a
subject would fail on such examples in any case. Furthermore it has to be ensured that erholen is not realized in the [Sbj IntrVerb] construction for intransitive verbs or respective functional categories in a Minimalist setting although the relation erholen’ (recreate) is a one-place predicate and hence erholen is semantically compatible with the construction.

6.5 Is there an alternative?

The question for theories denying the existence of valence structure is what replaces it to explain idiosyncratic lexical selection. In her exoskeletal approach, Borer (2005) explicitly rejects lexical valence structures. But she posits post-syntactic interpretive rules that are difficult to distinguish from them. To explain the correlation of depend with an on-PP, she posits the following interpretive rule Borer (2005, Vol. II, p. 29):

\[(67)\text{MEANING} \leftrightarrow \pi_9 + [(\text{on})]
\]

Borer refers to all such cases of idiosyncratic selection as idioms. In rule such as this one, ‘MEANING is whatever the relevant idiom means.’ (Borer, 2005, Vol. II, p. 27) In this rule, \(\pi_9\) is the ‘phonological index’ of the verb depend and \(e^{on}\) ‘corresponds to an open value that must be assigned range by the f-morph on’ (ibid), where f-morphs are function words or morphemes. Hence this rule brings together much the same information as the lexical valence structure in (58c). Discussing such ‘idiom’ rules, Borer writes

Although by assumption a listeme cannot be associated with any grammatical properties, one device used in this work has allowed us to get around the formidable restrictions placed on the grammar by such a constraint—the formation of idioms. …Potentially, then, within the system developed here, any syntactic or morphological property which does not reduce directly to some formal computational principle is to be captured by classifying the relevant item as an idiom—a partial representation of a phonological index with some functional value. …Such idiomatic specification could be utilized, potentially, not just for arrive and depend on, but also for obligatorily transitive verbs . . . , for verbs such as put, with their obligatory locative, and for verbs which require a sentential complement.

The reader may object that subcategorization, of sorts, is introduced here through the back door, with the introduction, in lieu of lexical syntactic annotation, of an articulated listed structure, called an idiom, which accomplishes, de facto, the same task. The objection of
course has some validity, and at the present state of the art, the introduction of idioms may represent somewhat of a concession . . . On the positive side, we note that to the extent that the existence of idioms is costly, we have attempted to put in place here a system which at least potentially extricates from the costly component of language all properties of listemes which are otherwise derivable from the structure.


Borer goes on to pose various questions for future research, related to constraining the class of possible idioms. With regard to that research program, as well as the last sentence in the above quote, it should be noted that a major focus of lexicalist research has been narrowing the class of subcategorization and extricating derivable properties from idiosyncratic subcategorization. Those are the functions of HPSG lexical hierarchies, for example. Whether future research within the exoskeletal approach can improve upon the past research in the lexical approach remains to be seen.

### 6.6 Summary

The discussion in the Sections 14–6.4 shows that the questions which arguments have to be realized in a sentence cannot be reduced to semantics and world knowledge or to general facts about subjects. The consequence is that valence information has to be connected to lexical items. One therefore has to assume a connection between a lexical item and a certain phrasal configuration as in Croft’s approach (2003) and in LTAG or assume our lexical variant. In a Minimalist setting the right set of features would have to be specified lexically to ensure the presence of the right case assigning functional heads. This is basically equivalent to what we are proposing here.

The discussion also showed that information about subjects is to some extent lexeme specific and hence has to be part of the specification of argument structure. This shows that proposals like the one by Goldberg and Jackendoff (2004) that only refers to phrasal VP constructions without mentioning the subject are not sufficient. Of course one could represent the fact that a certain head requires a subject lexically and assume a phrasal approach for the VP. This would be a hybrid approach and it seems unclear why one should not assume the lexical approach in total given that subjects are also involved in valence alternations as in (68):

(68)  

a. The soup cooked. 

b. He cooked the soup.
7 Relations between constructions

On the lexical rules approach, word forms are related by lexical rules: a verb stem can be related to a verb with finite inflection and to a passive verb form; verbs can be converted to adjectives or nouns; and so on. The lexical argument structure accompanies the word and can be manipulated by the lexical rule. In this section we consider what can replace such rules within a phrasal or ASC approach.

7.1 Inheritance hierarchies for constructions

We have seen that in a lexical approach one can relate variants of a construction by lexical rules. In a phrasal approach one would assume that there are alloconstructions for active and passive (for instance Bergen and Chang, 2005; Kallmeyer and Osswald, To appear). If one wants to capture the underlying nature of passive one would have to find some way to relate phrasal configurations. This is necessary since a phrasal approach must otherwise stipulate a great many phrasal constructions that correspond to observable valency patterns. For instance, one would assume the ones in (69) for German:

(69)  a. Nom V
   b. Nom Acc V
   c. Nom Dat V
   d. Nom Dat Acc V

The passive patterns corresponding to (69) would be:

(70)  a. V V-Aux
   b. Nom V V-Aux
   c. Dat V V-Aux
   d. Dat Nom V V-Aux

Researchers working in various frameworks tried to develop inheritance-based analyses that could capture the relation between the valency patterns in (69) and (70) (see for instance Kay and Fillmore, 1999, p. 12; Koenig, 1999, Chapter 3; Michaelis and Ruppenhofer, 2001, Chapter 4; Candito, 1996; Clément and Kinyon, 2003, p. 188; Koenig, 1999; ?; Kordoni, 2001 for proposals in CxG, TAG, and HPSG), but this enterprise was not successful. For technical problems with the original Berkeley CxG proposal see Müller, 2006, Section 2.4, Section 3.

As was also pointed out in Müller, 2006, Section 4, valency changing processes in general may not be modeled by inheritance since processes like causativization can be applied several times. An example is Turkish, which allows double and even triple causation (Lewis, 1967, p. 146):

7.2 Mappings between Different Levels of Representations

Culicover and Jackendoff (2005, Chapter 6.3) suggest that passive should be analyzed as one of several possible mappings from the Grammatical Function tier to the surface realization of arguments, that is, an NP in a certain case, with certain agreement properties, or in a certain position. While such analysis that work by mapping elements with different properties onto different representations are common in theories like LFG and HPSG (Koenig, 1999; Bouma, Malouf and Sag, 2001), a general property of these analyses is that one needs one level of representation per interaction of phenomena (ARG-ST, SEM-ARG, ADD-ARG in Koenig’s proposal, ARG-ST, DEPS, SPR, COMPS in Bouma, Malouf, and Sag’s proposal). This was discussed extensively in Müller, 2008, Section 7.5.2.2 with respect to extensions that would be needed for Koenig’s analysis.

Since Culicover and Jackendoff argue for a phrasal model, we will discuss
7.2 Mappings between Different Levels of Representations

their proposal here. The authors write:

we wish to formulate the passive not as an operation that deletes or alters part of the argument structure, but rather as a piece of structure in its own right that can be unified with the other independent pieces of the sentence. The result of the unification is an alternative licensing relation between syntax and semantics. (Culicover and Jackendoff, 2005, p. 203)

They suggest the following representation of the passive:

\[ GF_i > [GF \ldots]_k \Leftrightarrow [ \ldots V_k + \text{pass} \ldots (\text{by NP}_i) \ldots ]_k \]

The italicized parts are the normal structure of the sentence and the non-italicized parts are an overlay on the normal structure, that is additional constraints that have to hold in passive sentences. GF stands for Grammatical Function. Culicover and Jackendoff (2005, p. 204) explicitly avoid names like Subject and Object since this is crucial for their analysis of the passive to work. They assume that the first GF following a bracket is the subject (p. 195–196) and hence has to be mapped to the appropriate tree position in English. Note that this view on grammatical functions and obliqueness does not account for subjectless sentences that are possible in some languages, for instance in German.32

Although Culicover and Jackendoff emphasize the similarity between their approach and Relational Grammar (Perlmutter, 1983), there is an important difference: In Relational Grammar additional levels (strata) can be stipulated if additional remappings are needed. In Culicover and Jackendoff’s proposal there is no additional level. This causes problems for the analysis of languages like Lithuanian (Timberlake, 1982, Section 5), Irish Noonan, 1994, and Turkish (Özkaragöz, 1986), which allow for double passivization. We will use Özkaragöz’s Turkish examples in (73) for illustration (1986, p. 77):

\begin{align*}
(73) & \quad \text{a. Bu şato-da boğ-ul-un-ur.} \\
& \quad \text{this chateau-LOC strangle-PASS-PASS-AOR} \\
& \quad \text{‘One is strangled (by one) in this chateau.’} \\
& \quad \text{b. Bu oda-da döv-ul-ün-ür.} \\
& \quad \text{this room-LOC hit-PASS-PASS-AOR} \\
& \quad \text{‘One is beaten (by one) in this room.’} \\
& \quad \text{c. Harp-te vur-ul-un-ur.} \\
& \quad \text{war-LOC shoot-PASS-PASS-AOR} \\
& \quad \text{‘One is shot (by one) in war.’}
\end{align*}

32 Of course one could assume empty expletive subjects, as was suggested by Grewendorf (1993, p. 1311), but empty elements and especially those without meaning are generally avoided in the constructionist literature. See Müller, 2010a, Section 3.4, Section 11.1.1.3 for further discussion.
-In, -n, and -Il are allomorphs of the passive morpheme. According to Özkara göz the data is best captured by an analysis that assumes that the passive applies to a passivized transitive verb and hence results in an impersonal passive.

Approaches that assume that the personal passive is the unification of a general structure with a passive-specific structure will not be able to capture this, since they committed to a certain structure too early. The problem for approaches that state syntactic structure for the passive is that such a structure, once stated, cannot be modified. Culicover and Jackendoff’s proposal works in this respect since there are no strong constraints in the right-hand side of their constraint in (72). But there is a different problem: When passivization is applied the second time, it has to apply to the innermost bracket, that is, the result of applying (72) should be:

\( (74) \ [GF_i > [GF_j \ldots ]]_k \Leftrightarrow [\ldots V_k + \text{pass} \ldots (\text{by NP}_i) \ldots (\text{by NP}_j) \ldots ]_k \)

This cannot be done with unification, since unification checks for compatibility and since the first application of passive was possible it would be possible for the second time as well. Dots in representations are always dangerous and in the example at hand one would have to make sure that NP\(_i\) and NP\(_j\) are distinct, since the statement in (72) just says there has to be a by PP somewhere. What is needed instead of unification would be something that takes a GF representation and searches for the outermost bracket and then places a bracket to the left of the next GF. But this is basically a rule that maps one representation onto another one, just like lexical rules do.

If Culicover and Jackendoff want to stick to a mapping analysis, the only option to analyze the data seems to be to assume an additional level for impersonal passives from which the mapping to phrase structure is done. In the case of sentences like (75), which is a personal passive, the mapping to this level would be the identity function and in the case of double passivization the correct mappings would be implemented by two mappings between the three levels that finally result in a mapping as the one that is seen in (73b).

\( (75) \ \text{Arkada-¸sım bu oda-da döv-ül-dü.} \)

friend-my this room-LOC hit-PASS-AOR

‘My friend is beaten (by one) in this room.’

Note that the double passivization is also problematic for purely inheritance based approaches. What all these approaches can suggest though is that they just stipulate three different relations between argument structure and phrase structure: active, passive, double passive. But this misses the fact that (73b) is a further passivization of (75).

In comparison a lexical rule-based approach as the one suggested by Müller (2003a) does not have any problems with double passivization: The first appli-
7.3 Is there an alternative to lexical rules?

In this section we have presented the attempts to replace lexical rules with methods of relating constructions. We believe those attempts have been unsuccessful because they fail to capture the derivational character of relations between word forms.

8 Arguments from Language Acquisition

Tomasello (2003) argues for a surface-oriented, pattern-based view on language acquisition. According to him a child hears sentences like (76) and realizes that certain slots can be filled freely (see also Dąbrowska, 2001 for analogous proposals in Cognitive Grammar).

(76)  a. Do you want more juice/milk?
     b. Mommy is gone.

From such utterances so-called pivot schemata are derived. Such schemata contain open slots into which words can be inserted. Examples of schemata that are abstracted from utterances like (76) are shown in (77):

(77)  a. more ___ → more juice/milk
     b. ___ gone → mommy/juice gone

At this stage (22 months) children do not generalize over such schemata. The schemata are construction islands and do not have syntax (Tomasello et al., 1997). English children acquire the capability to use novel verbs with a subject and an object in the SVO order slowly in their third or fourth year of life (Tomasello, 2003, p. 128–129). More abstract syntactic and semantic generalizations only emerge in the course of time: after a sufficient amount of encounters of the transitive construction, the child can generalize over the patterns:

(78)  a. [S [NP The man/the woman] sees [NP the dog/the rabbit/it]].
     b. [S [NP The man/the woman] likes [NP the dog/the rabbit/it]].
     c. [S [NP The man/the woman] kicks [NP the dog/the rabbit/it]].

According to Tomasello (2003, p. 107) the abstraction of these patterns is [Sbj TrVerb Obj], the so-called transitive construction. For Tomasello, children acquire the item-specific patterns such as (77)—the ones involving specific verbs are called verb islands—and they eventually also acquire phrasal constructions such as the transitive construction.

Verb islands would seem to be perfect precursors to the acquisition of the lexicalist’s valence structures, and indeed Tomasello comments that ‘The lexical rules approach would seem to be better adapted to children’s verb island constructions and other item-based constructions—which are defined by particular verbs or other words—than to their totally abstract constructions (Tomasello, 2003, p. 193). Constructions such as the transitive construction nonetheless carry meaning, according to Tomasello.

What people crucially never acquire, according to Tomasello, are ‘meaningless’ phrase structure rules or lexical rules that relate one variant of a word to another. Constructions seem to take the place of both of these at once, and can plausibly be acquired by children by generalizing directly over the phrasal patterns. The constructions at various levels of generality are related to one another by inheritance hierarchies (Langacker, 1987; Goldberg, 1995, Chapter 3; Croft, 2001, p. 26; Tomasello, 2003, p. 106–107). In language production a number of such constructions combine to form a sentence (Goldberg, 2006, p. 10).

However, this attractive picture depends upon the phrasal view of constructions, which has been shown to be untenable. In addition to the problems already noted, a purely pattern-based approach faces difficulties from discontinuities in the realization of a head and its arguments. For instance, adjuncts can be serialized between the subject and the verb. Bergen and Chang (2005, p. 170), who implement the phrasal approach, suggest an active-ditransitive construction with the pattern in (79):

(79) [RefExpr Verb RefExpr RefExpr]

RefExpr stands for referential expression. Their formalization allows a discontinuity between the first referential expression and the verb. This makes it possible to analyze (80a,b), but excludes (80c), since in (80c) the adverb intervenes between verb and the first object:

(80) a. Marry tossed me a drink.
   b. Marry happily tossed me a drink.
   c. * Marry tossed happily me a drink.

However, by enforcing the adjacency between verb and object the analysis of coordinations like (81) becomes impossible.

(81) Marry tossed me a juice and Peter a water.
One part of the meaning of this sentence is contributed by the ditransitive construction for *Marry tossed Peter a water*. However, *tossed* and *Peter* are discontinuous. Similarly, one can construct examples with a discontinuity between the two objects of the ditransitive construction:

(82) He showed me and bought for Mary the book that was recommended in the Guardian last week.

(me is not adjacent to the book … although both phrases are part of the ditransitive construction. If one does not use empty elements and dislocation, one cannot maintain the claim that the items of the ditransitive construction have to be continuous. The point here is that it is not a certain fixed configuration that has to be acquired but rather the fact that there is a certain dependency between material in a clause. If material is realized together in a certain syntactic environment, a certain meaning can be observed.

Note also that a purely pattern-based approach is weakened by the existence of examples like (83):

(83) a. John tried to sleep.
    b. John tried to be loved.

Although no argument of *sleep* is present in the phrase *to sleep* and neither a subject nor an object is realized in the phrase *to be loved*, both phrases are recognized as phrases containing an intransitive and a transitive verb, respectively.\textsuperscript{33} The same applies to arguments that are supposed to be introduced/licensed by a phrasal construction: in (84) the resultative construction is passivized and then embedded under a control verb, resulting in a situation in which only the result predicate and the matrix verb are realized overtly.

(84) Der kranke Mann wünschte sich, tot geschossen zu werden.\textsuperscript{34}  
    the ill man wished SELF dead shot to be  
    ‘The ill man wanted to be shot dead.’

Of course passivization and control are responsible for these occurrences, but the important point here is that arguments can remain unexpressed or implicit and nevertheless a meaning that is usually connected to some overt realization of arguments is present (Müller, 2007a, Section 4). So, what has to be acquired by the language learner is that when a result predicate and a main verb are realized together, they contribute the resultative meaning. An additional example that shows that the NP arguments that are usually realized in active resultative constructions may remain implicit are nominalizations like the ones in (85):

\textsuperscript{33}Constructionist theories do not assume empty elements. Of course, in the GB framework the subject would be realized by an empty element. So it would be in the structure, although invisible.

\textsuperscript{34}Müller, 2007a, p. 387.
9 Lexical Items, Licensed Trees, Classes of Trees

Proponents of the phrasal approach criticize lexical approaches as being inadequate for capturing language acquisition facts. However, by looking at the properties of lexical approaches carefully, it becomes evident that lexical approaches fulfill the desiderata that are required by the language acquisition facts. As we saw in the previous section, the fact that certain arguments of a head have to be realized in a certain environment have to be captured by a linguistic theory. Discontinuities between heads and their arguments have to be allowed. To see that lexical theories can capture the facts consider Categorial Grammar (Ajdukiewicz, 1935; Steedman, 2000). Tomasello’s Transitive Construction [Sbj TrVerb Obj] corresponds to the following lexical item in Categorial Grammar (s\np)/np. This lexical entry for, let’s say, likes expresses the fact that likes takes an np to its right (marked by the direction of the slash '/') and an np to its left (marked by the direction of the slash '\'). The lexical item licences structures like the one that is displayed as a tree in Figure 4. The combinations are licensed by combinatorial rules that combine a functor with an argument. So all lexical items that are assigned to the category (s\np)/np can appear in configurations like the one shown in Figure 4. And of course if we observe unknown words in the position of the verb, we can infer that the unknown words have to belong into the same lexical class as likes. However, the structure in Figure 4 is not the only one that is possible for items of the category (s\np)/np. For instance an adjunct of the category (s\np)/(s\np) may intervene between the subject and the combination of verb and

---

object. This is shown in Figure 5. The adjunct probably takes a VP (s/np) to its right and the result of the combination is a VP again.

Similarly, lexical items like likes can appear in coordination structures of the kind discussed above. See Steedman, 1991 for details on coordination.

The bracketing in (s\np)/np ensures that the rightmost np in the expression is combined with the verb first and then the combination with the second np takes place. This results in the SVO order that can be observed for languages like English. For languages with a freer constituent structure Steedman and Baldridge (2006) suggest a generalized representation. The arguments are represented in a set and for the combination of a head with an argument, one element from this set is chosen. This results in different orders. For a head finial language the lexical item of a transitive verb would be s \{ \np, \np \}. See also Hoffman (1995, Section 3.1) for a similar proposal for Turkish. Such a lexical item stands for trees in which the two np arguments precede their head in any order. Such an approach to
constituent order was also suggested by Gunji (1986) in the framework of HPSG and by Fanselow (2001) in the framework of Minimalism.

10 Arguments from Psycholinguistics

Goldberg (1995, Section 1.4.5) uses evidence from psycholinguistic experiments to argue against lexical approaches that use lexical rules to account for argument structure alternations: Carlson and Tanenhaus (1988) showed that sentences with true lexical ambiguity like those in (86) and sentences with two verbs with the same core meaning have different processing times.

(86) a. Bill set the alarm clock onto the shelf.
    b. Bill set the alarm clock for six.

(87) a. Bill loaded the truck onto the ship.
    b. Bill loaded the truck with bricks.

Errors due to lexical ambiguity cause a bigger increase in processing time than errors in the use of the same verb. Experiments showed that there was a bigger difference in processing times for the sentences in (86) than for the sentences in (87). The difference in processing times between (87a) and (87b) would be explained by different preferences for phrasal constructions. In a lexicon-based approach one could explain the difference by assuming that one lexical item is more basic, that is, stored in the mental dictionary and the other is derived from the stored one. The application of lexical rules would be time consuming, but since the lexical items are related, the overall time consumption is smaller than the time that is needed to process two unrelated items (Müller, 2002a, p. 405).

Alternatively one could assume that the lexical items for both valency patterns are the result of lexical rule applications. As with the phrasal constructions, the lexical rules would have different preferences. This shows that the lexical approach can explain the experimental results as well, so that they do not force us to prefer phrasal approaches.

Goldberg (1995, p. 18) claims that lexical approaches have to assume two variants of load with different meaning and that this would predict that load alternations would behave like two verbs that really have absolutely different meanings. The experiments discussed above show that such predictions are wrong and hence lexical analyses would be falsified. However, as was shown in Müller, 2010a, Section 11.11.8.2, the argumentation contains two flaws: Let’s assume that the construction meaning of the construction that licences (87a) is $C_1$ and the construction meaning of the construction that licences (87b) is $C_2$. Under such assumptions the semantic contribution of the two lexical items in the lexical analysis
would be (88). load(…) is the contribution of the verb that would be assumed in phrasal analyses.

(88)   a. load (onto): C1 ∧ load(…)
   b. load (with): C2 ∧ load(…)

(88) shows that the lexical items partly share their semantic contribution. We hence predict that the processing of the dispreferred argument realization of load is simpler than the dispreferred meaning of set: in the latter case a completely new verb has to be activated while in the first case parts of the meaning are activated already. (See also Croft, 2003, p. 64–65 for a brief rejection of Goldberg’s interpretation of the experiment that corresponds to what is said here)

Goldberg (1995, p. 107) argues against lexical rule-based approaches for locative inversions like (89), since according to her such approaches have to assume that one of the verb forms has to be the more basic form.

(89)   a. He loaded hay onto the wagon.
   b. He loaded the wagon with hay.

She remarks that this is problematic since we do not have clear intuitions on what is the basic and what the derived form. She argues that the advantage of phrasal approaches is that various constructions can be related to each other without necessitating the assumption that one of the constructions is more basic than the other. There are two phrasal patterns and the verb is used in one of the two patterns. This criticism can be addressed in two ways: First one could introduce two lexical types (for instance onto-verb and with-verb) into a type hierarchy. The two types correspond to two valency frames that are needed for the analysis of (89a) and (89b). These types can have a common supertype (onto-with-verb) which is relevant for all spray/load verbs. One of the subtypes or the respective lexical item of the verb is the preferred one. This corresponds to a disjunction in the lexicon, while the phrasal approach assumes a disjunction in the set of constructions.

A variant of this approach is to assume that the lexical description of load just contains the supertyp, that describes all spray/load verbs. Since Modell Theoretic Approaches assume that all structures that are models of utterances contain only maximally specific types (see for instance King, 1999 and Pollard and Sag, 1994, p. 21), it is sufficient to say about verbs like load that they are of type onto-with-verb. Since this type has exactly two subtypes, load has to be either onto-verb or with-verb in an actual model.37

A second option is to stick with lexical rules and to assume a single representation for the root of a verb that is listed in the lexicon. In addition one assumes

---

37 This analysis does not allow one to specify verb specific preferences for one of the realization patterns since the lexicon contains the general type only.
two lexical rules that map this basic lexical item onto other items that can be used
in syntax after being inflected. The two lexical rules can be described by types
that are part of a type hierarchy and that have a common supertype. This would
capture commonalities between the lexical rules. We therefore have the same
situation as with phrasal constructions (two lexical rules vs. two phrasal construc-
tions). The only difference is that the action is one level deeper in the lexical
approach, namely in the lexicon (Müller, 2002a, p. 405–406).

The argumentation with regard to the processing of resultative constructions
like (89c) is parallel:

(90)  a. He drinks.
     b. He drinks the milk.
     c. He drinks the pub empty.

When humans parse a sentence they build up structure incrementally. If one hears
a word that is incompatible with the current hypothesis, the parsing process breaks
down or the current hypothesis is revised. In (90c) the pub does not correspond
to the normal transitive use of drink, so the respective hypothesis has to be re-
vised. In the phrasal approach the resultative construction would have to be used
instead of the transitive construction. In the lexical analysis the lexical item that
is licensed by the resultative lexical rule would have to be used rather than the bi-
valent one. The building of syntactic structure and lexicon access in general place
different demands on our processing capacities. However, when (90c) is parsed,
the lexical items for drink are active already, we only have to use a different one.
It is currently unclear to us whether psycholinguistic experiments can differentiate
between the two approaches, but it seems to be unlikely.

11 Arguments from Statistical Distribution of Material

In this section, we want to look at arguments from statistics that were claimed
to support a phrasal view. We first look at data-oriented parsing, a technique
that was successfully used by Bod (2009b) to model language acquisition and
then we turn to the collostructional analysis of Stefanowitsch and Gries (2009)
and discuss whether this distributional analysis can decide the question whether
argument structure constructions are phrasal or lexical.

11.1 Unsupervised Data-Oriented Parsing

Rens Bod (2009b) demonstrated that a simple statistical procedure can learn quite
elaborated linguistic structures that correspond roughly to what linguists would
assume. In particular he showed that such a technique can learn auxiliary inversion and gets the inversion facts right even for complex examples containing relative clauses with auxiliaries and even if this type of clause was not in the data that was used for learning. Chomsky (1971, p. 29–33) has used (and is still using\textsuperscript{38}) auxiliary inversion as his key example of a Poverty of the Stimulus in the language acquisition debate, but Bod has shown that six examples are sufficient to acquire complex auxiliary inversion structures. The examples that are needed do not include the data that Chomsky considered crucial for a language acquisition device that relies on input alone. Bod’s procedure works as follows: An utterance is partitioned into (binary branching) trees. It is then checked how likely each of the subtrees is, that is, it is checked whether an identical subtree occurred in previous utterances. If this is the case, this renders the subtree under consideration more likely. To take an example, consider the corpus in (91). Figure 6 shows the unlabeled trees for the two sentences.

(91)  
\begin{enumerate}
\item Watch the dog.
\item The dog barks.
\end{enumerate}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{binary_trees.png}
\caption{Possible binary branching structures for Watch the dog and The dog barks.}
\end{figure}

Some of these trees do not correspond to structures that linguists would assume, but the good news is, that the subtree for the dog appear more often than for instance watch the and this renders the correct structures for watch the dog and the dog barks more likely in Bod’s procedure.

If one assumes that language acquisition is based on input alone, the structures that Bod’s procedure extracts from the input must be those that are acquired by children (part of speech information, meaning, and context are currently not

\textsuperscript{38}Talk in Stuttgart at 24th of March, 2010

included in Bod’s computations, but of course this can be done). By the reverse argument of the poverty of the stimulus the structures that are learned should be the ones that linguists assume. Since Bod did not have sufficient data to do the computations with flat structures and arbitrary branchings, he arbitrarily restricted the system to binary branching structures (p. 760). This means that his experiments do not answer the question if rules should licence flat structures or binary branching ones. But we probably can expect interesting results in the future.

What distributional analyses can not determine is where the meaning is represented in a tree. Bod (2009a, p. 132) claims that his procedure is a testable realization of CxG in Goldberg’s sense, but the trees that he constructs do not help deciding between phrasal and lexical analyses or analyses that involve empty elements. The alternative proposals for the analysis of (92) are depicted in Figure 7.

\[(92) \text{[dass] er den Teich leer fischt} \]

that he the pond empty fishes

‘that he fishes the pond empty’

\[ \text{er ihn leer fischt} \]

\[ \text{er ihn leer fischt} \]

\[ \text{er ihn leer fischt} \]

Figure 7: Three possible analyses for resultative constructions: Holistic construction, lexical rule, empty head

The first picture stands for a complex construction that contributes the resultative meaning as a whole. The second picture stands for the analysis with a lexical rule, and the third one for an analysis with an empty head. Empty heads are often the choice in mainstream generative grammar, but as shown in Müller, 2010a, Section 11.10, some of them can be converted into lexical rules by known techniques.
of grammar conversion (Bar-Hillel, Perles and Shamir, 1961). A distributional analysis cannot differentiate between these proposals. The distribution is computed with regard to the words. The meaning of the words is not considered. One can observe that the utterance contains the word *fischt* (‘fishes’), but one cannot see whether this word contributes the resultative semantics or not. Similarly the distributional analysis cannot distinguish between analyses with and without an empty head. The empty head is naturally not recognizable in the signal. The empty head is a theoretical construct and as was mentioned above an analysis with an empty head of the kind in Figure 7 can be automatically converted into one with a lexical rule. The argumentation for one of the latter analyses is purely theory internal.

Concluding this subsection, we contend that Bod’s paper is a milestone in the Poverty of the Stimulus debate, but it does not and cannot show that a particular version of constructionist theories, namely the phrasal one, is correct.

### 11.2 Collostructions

Stefanowitsch and Gries (2009, Section 5) assume a plugging analysis: *words occur in (slots provided by) a given construction if their meaning matches that of the construction*. The authors claim that their *collostructional analysis has confirmed [the plugging analysis] from various perspectives*. Stefanowitsch and Gries are able to show that certain verbs occur more often than not in particular constructions, while other verbs never occur in the respective constructions. For instance, *give, tell, send, offer* and *show* are attracted by the Ditransitive Construction, while *make* and *do* are repelled by this construction. Regarding this distribution the authors write:

> These results are typical for collexeme analysis in that they show two things. First, there are indeed significant associations between lexical items and grammatical structures. Second, these associations provide clear evidence for semantic coherence: the strongly attracted collexemes all involve a notion of ‘transfer’, either literally or metaphorically, which is the meaning typically posited for the ditransitive. This kind of result is typical enough to warrant a general claim that collostructional analysis can in fact be used to identify the meaning of a grammatical construction in the first place. (Stefanowitsch and Gries, 2009, p. 943)

We hope that the preceding discussion made clear that the distribution of words in a corpus cannot be seen as evidence for a phrasal analysis. The corpus study shows that *give* usually is used with three arguments in a certain pattern that is typical for English (Subject Verb Object1 Object2) and that this verb forms a cluster with
other verbs that have a transfer component in their meaning. The corpus data does not show whether this meaning is contributed by a phrasal pattern or by lexical entries that are used in a certain configuration.

12 Arguments from Neurolinguistics

Pulvermüller, Cappelle and Shtyrov (To appear) discuss neurolinguistic facts and relate them to the CxG view on grammar theory. One important finding is that deviant words (lexical items) cause brain responses that differ in polarity from brain responses on incorrect strings of words, that is, syntactic combinations. This suggests that there is indeed an empirical basis for deciding the issue.

Concerning the standard example of the caused motion construction in (93) the authors write the following:

(93) She sneezed the foam off the cappuccino. 39

this constellation of brain activities may initially lead to the co-activation of the verb sneeze with the DCNAs for blow and thus to the sentence mentioned. Ultimately, such co-activation of a one-place verb and DCNAs associated with other verbs may result in the former one-place verb being subsumed into a three-place verb category and DCNA set, a process which arguably has been accomplished for the verb laugh as used in the sequence laugh NP off the stage. (Pulvermüller, Cappelle and Shtyrov, To appear)

A DCNA is a discrete combinatorial neuronal assembly. Regarding the specifics of DCNAs the authors write that

Apart from linking categories together, typical DCNAs establish a temporal order between the category members they bind to. DCNAs that do not impose temporal order (thus acting, in principle, as AND units for two constituents) are thought to join together constituents whose sequential order is free or allow for scrambling. (Pulvermüller, Cappelle and Shtyrov, To appear)

We believe that this view is entirely compatible with the lexical view outlined above: the lexical item or DCNA requires certain arguments to be present. A lexical rule that relates an intransitive verb to one that can be used in the caused motion construction is an explicit representation of what it means to activate the valence frame of blow.

39Goldberg, 2006, p. 42

The authors cite earlier work (Cappelle, Shtyrov and Pulvermüller, 2010) and argue that particle verbs are lexical objects, admitting for a discontinuous realization of particle verbs despite their lexical status (p. 21). They restrict their claim to frequently occurring particle verbs. This claim is of course compatible with our assumptions here, but the differences in brain behaviour are interesting when it comes to fully productive uses of particle verbs. For instance any mono-valent verb in German can be combined with the aspectual particle los: losstanzen (‘start to dance’), loslachen (‘start to laugh’), losingen (‘start to sing’), . . . Similarly, the combination of mono-valent verbs with the particle an with the reading directed-towards is also productive, anfahren (‘drive towards’), anlachen (‘laugh in the direction of’), ansegeln (‘sail towards’), . . . (see Stiebels, 1996 on various productive patterns). The interesting question is how particle verbs behave that follow these patterns but occur with low frequency. This is still an open question as far as the experimental evidence is concerned, but as we argue below lexical proposals to particle verbs as the one suggested by Müller (2003b) are compatible with both possible outcomes.

Summarizing the discussion so far, lexical approaches are compatible with the accumulated neurobiological evidence and as far as particle verbs are concerned they seem to be better suited than the phrasal proposals by Booij (2002, Section 2) and Blom (2005) (See Section 3.2 for discussion). However, in general it remains an open question what it means to be a discontinuous lexical item. The idea of discontinuous words is pretty old (Wells, 1947), but there have not been many formal accounts of this idea. Nunberg, Sag and Wasow (1994) suggest a representation in a linearization-based framework of the kind that was proposed by Reape (1994) and Kathol (1995, p. 244–248) and Crysmann (2002) worked out such analyses in detail. Kathol’s lexical item for aufwachen (‘to wake up’) is given in (94):

\[
\text{aufwachen} \quad \text{(following Kathol (1995, p. 246))):
\]

\[
\begin{align*}
\text{DOM} & \left( \langle \text{wachen} \rangle \cdots \text{HEAD} \varnothing \text{ verb} \cdots \text{VCOMP} \langle \rangle \right) \bigcirc \left( \langle \text{auf} \rangle \text{ SYNSEM} \varnothing \cdots \text{HEAD} \begin{bmatrix} \text{FLIP} - \text{sepref} \end{bmatrix} \right) \\
\end{align*}
\]

The lexical representation contains the list-valued feature DOM that contains a description of the main verb and the particle. The dependency between the particle and the main verb was characterized by the value of the VCOMP feature, which is a valence feature for the selection of arguments that form a complex predicate with their head. The shuffle operator \( \bigcirc \) concatenates two lists without specifying an order between the elements of the two lists, that is, both wachen, auf and auf,
wachen are possible. The little marking vc is an assignment to a topological field in the clause.

Müller (2007a) criticized such linearization-based proposals since it is unclear how analyses that claim that the particle is just linearized in the domain of its verb can account for sentence like (95), in which complex syntactic structures are involved. German is a V2 language and the fronting of a constituent into the position before the finite verb is usually described as some sort of nonlocal dependency, that is, even authors who assume linearization-based analyses do not assume that the initial position is filled by simple reordering of material (Kathol, 2001; Müller, 1999, 2002a; Bjerre, 2006).

(95) a. \[[v \text{mf} \text{Den Atem} ]\ [v \text{cn} \text{an} ]\] hielt die ganze Judenheit.\(^{41}\)
    the breath PART held the whole Jewish community
    ‘The whole Jewish community held their breath.’

b. \[[v \text{mf} \text{Wieder} ]\ [v \text{cn} \text{an} ]\] treten auch die beiden Sozialdemokraten.\(^{42}\)
    again PART kick also the two social democrats
    ‘The two Social Democrats are also running for office again.’

\(^{40}\)Kathol (1995, Section 6.3) suggested such an analysis, but later changed his view. Wetta (2011) also assumes a purely linearization-based approach. He assumes that sentences in which multiple constituents are fronted (Müller, 2003b) are analyzed in such a way that more than one linearization object are inserted as one single object into the position before the finite verb. This fails to account for multiple frontings that cross a clause boundary as in the examples in (i) discussed by Fanselow (1993, p. 67):

(i) a. Der Maria das Buch wenn du denkst daß du geben darfst bist du schön
    the Maria the book if you think that you give be.allowed.to are you pretty,
    stupid
    ‘You are pretty stupid if you think you are allowed to give Maria the book.’

b. Der Maria einen Ring glaub ich nicht daß er je schenken wird.
    the Maria a ring believe I not that he ever give.as.present will
    ‘I do not believe that he will ever give Maria a ring as a present.’

If such sentences are to be analyzed as verb second sentences involving a dislocation mechanism – as also assumed by Wetta for non-local extraction (p. 265), there has to be a connection between a single element in the embedded clause and the fronted constituents der Maria das Buch and der Maria einen Ring. However, no such single projection exists in linearization-based proposals, since der Maria is the dative object and das Buch is the accusative object and the two do not form a constituent on the level of structure that is relevant for extraction. In earlier work Müller (2000) suggested two nonlocal dependencies for the analysis of multiple frontings, but this was revised later and superseded by an analysis that integrates nicely into the rest of German grammar (Müller, 2005a,b).

\(^{41}\)Lion Feuchtwanger, Jud Süß, p. 276, quoted from Grubačić, 1965, p. 56.
\(^{42}\)taz, bremen, 24.05.2004, p. 21.
The conclusion that has to be drawn from examples like (95) is that particles interact in complex ways with the syntax of sentences. This is captured by the lexical treatment that was suggested in Müller, 2002a, Chapter 6 and Müller, 2003b: The main verb selects for the verbal particle. By assuming that wachen selects for auf the tight connection between verb and particle is represented. Such a lexical analysis provides an easy way to account for fully intransparent particle verbs like anfangen (‘to begin’). However, we also argued for a lexical treatment of transparent particle verbs like losfahren (‘to start to drive’) and jemanden/etwas anfahren (‘drive directed towards somebody/something’). The analysis involves a lexical rule that licences a verbal item that selects for an adjunct particle. The particles an and los can modify verbs and contribute arguments (in the case of an) and their semantics. This analysis can be shown to be compatible with the neuro-mechanical findings: if it is the case that even transparent particle verb combinations with low frequency are stored than the rather general lexical rule that was suggested by Müller is the generalization of the relation between a large amount of lexical particle verb items and their respective main verb. The individual particle verbs would be special instantiations that have the form of the particle specified as it is also the case for non-transparent particle verbs like anfangen. If it should turn out that productive particle verb combinations with particle verbs of low frequency cause syntactic reflexes in the brain, this could be explained as well: The lexical rule licences an item that selects for an adverbial element. This selection would then be seen as parallel to the relation between the determiner and the noun in the NP der Mut (‘the courage’), which (Cappelle et al., 2010, p. 191) discuss as an example of a syntactic combination. Note that Müller’s analysis is also compatible with another observation made by Shtyrov, Pihko and Pulvermüller (2005): Morphological affixes also cause the lexical reflexes. In Müller’s analysis the stem of the main verb is related to another stem that selects for a particle. This stem can be combined with (derivational and inflectional) morphological affixes causing the lexical activation pattern in the brain. After this combination the verb is combined with the particle and the dependency can be either a lexical or a syntactic one, depending on the results of the experiments to be carried out. The

---

\[ \text{c. } [\text{vf } [\text{vc Los} ] [\text{af dammit} ] \text{ geht es schon am 15. April.}^{43} \]

PART there with went it already at the 15th April

‘It already started on April the 15th.’

\[ \text{The main verb selects for the verbal particle. By assuming that wachen selects for auf the tight connection between verb and particle is represented.}^{44} \]

\[ \text{Such a lexical analysis provides an easy way to account for fully intransparent particle verbs like an-fangen (‘to begin’). However, we also argued for a lexical treatment of transparent particle verbs like losfahren (‘to start to drive’) and jemanden/etwas anfahren (‘drive directed towards somebody/something’). The analysis involves a lexical rule that licences a verbal item that selects for an adjunct particle. The particles an and los can modify verbs and contribute arguments (in the case of an) and their semantics. This analysis can be shown to be compatible with the neuro-mechanical findings: if it is the case that even transparent particle verb combinations with low frequency are stored than the rather general lexical rule that was suggested by Müller is the generalization of the relation between a large amount of lexical particle verb items and their respective main verb. The individual particle verbs would be special instantiations that have the form of the particle specified as it is also the case for non-transparent particle verbs like anfangen. If it should turn out that productive particle verb combinations with particle verbs of low frequency cause syntactic reflexes in the brain, this could be explained as well: The lexical rule licences an item that selects for an adverbial element. This selection would then be seen as parallel to the relation between the determiner and the noun in the NP der Mut (‘the courage’), which (Cappelle et al., 2010, p. 191) discuss as an example of a syntactic combination. Note that Müller’s analysis is also compatible with another observation made by Shtyrov, Pihko and Pulvermüller (2005): Morphological affixes also cause the lexical reflexes. In Müller’s analysis the stem of the main verb is related to another stem that selects for a particle. This stem can be combined with (derivational and inflectional) morphological affixes causing the lexical activation pattern in the brain. After this combination the verb is combined with the particle and the dependency can be either a lexical or a syntactic one, depending on the results of the experiments to be carried out. The

---

\[ \text{taz, 01.03.2002, p. 8.}^{43} \]

\[ \text{Cappelle et al. (2010, p. 197) write: the results provide neurophysiological evidence that phrasal verbs are lexical items. Indeed, the increased activation that we found for existing phrasal verbs, as compared to infelicitous combinations, suggests that a verb and its particle together form one single lexical representation, i. e. a single lexeme, and that a unified cortical memory circuit exists for it, similar to that encoding a single word We believe that Müller’s analysis is compatible with this statement.}^{44} \]
Note that Müller’s analysis allows the principle of lexical integrity to be maintained. We therefore do not follow (Cappelle et al., 2010, p. 198), who claim that they provide proof that potentially separable multi-word items can nonetheless be word-like themselves, and thus against the validity of a once well-established linguistic principle, the Lexical Integrity Principle. We agree that non-transparent particle verbs are multi-word lexemes, but the existence of multi-word lexemes does not show that syntax has access to the word internal morphological structure. The parallel between particle verbs and clearly phrasal idioms was discussed in Müller, 2002a,b and it was concluded that idiom-status is irrelevant for the question of wordhood. Since the interaction of clearly phrasal idioms with derivational morphology as evidenced by examples like (96) did not force grammarians to give up on Lexical Integrity, it can be argued that particle verbs are not convincing evidence for giving up the Lexical Integrity Principle either.\footnote{45}{However, see Booij, 2009 for some challenges to Lexical Integrity.}

\begin{enumerate}
\item a. Er hat ins Gras gebissen.
\hspace{1cm} ‘He bit the dust.’
\item b. „Heath Ledger“ kann ich nicht einmal schreiben, ohne dass mir sein Heath Ledger can I not even write without that me his ins Gras-Gebeiße wieder so wahnsinnig leid tut\footnote{46}{http://www.coffee2watch.at/egala. 23.03.2012} in.the grass.biting again so crazy sorrow does ‘I cannot even write “Heath Ledger” without being sad again about his repeated biting the dust.’
\end{enumerate}

The example in (96b) involves the discontinuous derivation with the circumfix Ge-\text_-e (Lüdeling, 2001, Section 3.4.3; Müller, 2002a, p. 324–327, 372–377; Müller, 2003b, Section 2.2.1, Section 5.2.1). Still the parts of the idiom ins Gras beiß- are present and with them the idiomatic reading. See Sag, 2007 for a lexical analysis of idioms that can explain examples like (96).

\section*{13 Conclusion}

We have shown in this paper that there are no compelling arguments for assuming phrasal argument structure constructions, but that there are several arguments against them. Assuming a lexical or — in the terminology of Goldberg (To appear) — template-based view solves all the problems that arise for phrasal approaches. Furthermore we showed that radically underspecified approaches in the sense of Borer (2005), Haugereid (2007), and Lohndal (2012) are not restricted enough.
The only way to establish the necessary restrictions is a lexical representation, since the information that has to be captured is in part lexeme dependent.

14 Appendix: subject idioms

A very frequent claim throughout the literature is that subjects have a special status, that subjects are not selected by their head (Chomsky, 1981, p. 26–28; Hoekstra, 1987, p. 33). However, Bresnan (1982a, Chapter 2) already showed that many of the claims that were made to support the non-selection of the subject are empirically wrong.

The claim that the subject is not selected is supported by the claim that there are no idioms with a subject as part of the idiom and an object slot that can be filled freely (Marantz, 1981, p. 50–51). Nunberg et al. (1994, p. 526) cited example (97a), among others, and Bresnan (1982a, p. 349–350) cited (97b) and (97c) as examples in which the subject is a fixed part of the idiom and there are open slots for non-subjects:

(97) a. A little bird told X that S.
   ‘X heard the rumor that S’

b. The cat’s got x’s tongue.
   ‘X cannot speak.’

c. What’s eating x?
   ‘Why is X so galled?’

Marantz (1984, p. 29) remarks that examples like (97b) are irrelevant for the discussion of his claims, since the free element is not the object. Regarding (97c), Marantz writes:

From the point of view of the present theory, it is important that this apparent subject idiom has no S-internal syntax, for it is precisely S-internal syntax that is at issue. What’s eating NP? is not a combination of subject and verb, forming a predicate on the object, but rather a combination of wh-question syntax, progressive aspect, plus subject and verb—that is, a complete sentence frame—with an open slot for an argument. (Marantz, 1984, p. 27)

Interestingly this points towards a Construction Grammar analysis that assumes that even complex linguistic objects can be stored. However, such approaches assume that complex linguistic objects do have internal structure. If one ignored this internal structure, it would not be possible to explain in what respect (97c) is similar to other English sentences and why (97c) is structured internally according to normal syntactic laws of English. Therefore, idiomatic phrases do get assigned...
an internal structure in frameworks like TAG (Abeillé and Schabes, 1989), HPSG (Krenn and Erbach, 1994; Sailer, 2000; Soehn and Sailer, 2008; Soehn, 2006; Sag, 2007; Richter and Sailer, 2009), and Simpler Syntax (Culicover and Jackendoff, 2005).

Marantz explains some of the examples that were provided by Bresnan away by assuming that the subjects are not subjects underlingly but objects since the verbs in the idioms are unaccusative verbs.

The claim that subjects are never part of idioms that have an open slot is widespread (see for instance den Dikken, 1995, p. 92, Kratzer, 1996, p. 112–116) and is also explicitly made for German for instance by Grewendorf (2002, p. 50). That the claim is not correct for German and hence cannot hold universally is shown by examples like (98) and (99), some of which were already discussed by Reis (1982, p. 178): 47

(98)  a. weil ihn der Hafer sticht
because him.ACC the oats.NOM pricks
‘because he is feeling his oats’

b. weil ihn der Teufel reitet
because him.ACC the devil.NOM rides
‘because he had the devil in him’

c. weil ihm alle Felle davonschwimmen
because him.DAT all furs.NOM away.swim
‘because all his hopes were dashed’

d. weil ihn der Schlag trifft
because him.ACC the stroke.NOM strikes
‘because he was flabbergasted’

e. weil ihn die Wut packt
because him.ACC the rage.NOM grabs
‘because he flies into a rage’

f. weil ihm der Kopf raucht
because him.DAT the head.NOM smokes
‘because his head is spinning’

g. weil kein Hahn danach kräht
because no cock there.after crows
‘because no one cares two hoots about that’

h. weil nach ihm kein Hahn kräht
because after him no cock.NOM crows
‘because no one cares two hoots about him’

47(98a), (98b), (98c), (99a) und (99b) are due to Reis (1982, p. 178). See also Haider (1993, p. 173) for examples that are similar to those in (98).
i. weil ihn der Esel im Galopp verloren hat
der donkey.in.the gallop lost has
‘He is out of the question.’ (according to Duden Redewendungen)

j. weil die Spatzen das vom Dach gepfiffen haben
because the sparrows.from the roof whistled have
‘because it was all over town’

(99) a. weil ihm der Geduldsfaden reißt
because him.DAT the patience.twist.NOM tears
‘because his patience gave out’

b. weil ihm der Kragen geplatzt ist
because him.DAT the collar is
‘because he blew his top’

c. weil ihm ein Stein vom Herzen fällt
because him a stone from the heart falls
‘because that’s a load off his mind’

d. weil bei ihm Hopfen und Malz verloren ist
because at him hop and malt lost is
‘because he is a hopeless case’

Marantz (1997, p. 208–209) formulates a variant of his earlier claim stating that
an agent may not be a fixed part of an idiom. The examples in (98b) and (98g)
show that even this more restricted version of the no-subject in idioms claim is
wrong.

Scherpenisse (1986, p. 89) working on German repeats Marantz’ claim for
English (1984) that idioms that contain subjects as fixed part can only be formed
with unaccusative verbs, that is, the subjects are claimed to be underlying objects.
If this claim were true there should not be passivizable idioms with fixed sub-
jects, since the passive is the suppression of the subject. If there is no underlying
subject, passive cannot apply. However, if one looks at the above examples on
notes that those idioms in (98) that contain verbs that can be passivized in their
literal meaning can be passivized in the idiom reading as well. It is well-known
from other idioms that passivization is sometimes marked. Many of the examples
contain inanimate subjects which also influences passivizability.48 Examples of
passivizations of the idiomatic variants are shown in (100) and (102a, b):

(100) a. Was schon von den Dächern gepfiffen wurde, jetzt ist es amtlich:

[...]

48 But see Müller, 2002a, Chapter 3.1.2 for passives with inanimate subjects.

49 http://witch.muensterland.org/2003/04/18.html. 05.01.2007.
b. Was vor wenigen Wochen nur vereinzelt von den Dächern gepfiffen wurde, nahm in der letzten Woche konkrete Formen an: […]  

c. In einer Pressemitteilung bestätigte die WWF am 23.03.2001 offiziell die Nachricht, die von allen Spatzen längst von den Dächern gepfiffen wurde.  

Example (101) shows a state passive of (98a), and the examples in (102) are passives of (98e), (102c) containing a prenominal participle with passive argument structure:

(101) Vom Hafer gestochen  

(102) a. Und doch kann sich kaum eine Frau davon freisprechen, daß sie von Empörung oder gar heiliger Wut gepackt wird, sobald sie eines Mannes auf ihrem Territorium ansichtig wird.  

b. […] Michaela Gerg, die nach einem verkorksten ersten Lauf von Wut gepackt wurde, im zweiten Durchgang Bestzeit fuhr und noch auf den siebten Rang kam.  


(103) and (104) show further passivizations and prenominal participles with passive argument structure:

(103) a. „Jordannis wird vom Teufel geritten“, witzelte Vangelis.  

b. Wird Gerda vom Teufel geritten?  

c. Er wurde verhext, oder sie wurde vom Teufel geritten.  

(104) a. viele vom Schlag getroffene und Lahme aber wurden geheilt.  

b. Jeffrey, der Sohn des vom Schlag Getroffenen, besucht seinen Vater im Krankenhaus.  

51 http://people.freenet.de/wwf-hp/WWFschlucktWCW.htm. 05.01.2007.  
52 taz, 06.04.2000, S. 20.  
54 taz, 29.01.1990, S. 13.  
This shows that idioms like *Die Spatzen pfeifen X von den Dächern*, *X packt die Wut, der Teufel reitet X* and *Y trifft der Schlag* are syntactically active. A theory about human language should capture that the structures for active sentences are related to the ones for passive sentences in the same way as this is captured for non-idiomatic sentences.

Sternefeld (1985, p. 435) notes that the nominatives in Marga Reis’ idiom examples behave like objects since they are adjacent to the verb and therefore have the same ordering properties as objects. Therefore Sternefeld assumes an incorporation analysis for such idioms, that is, he assumes that the verb and the nominative form a fixed unit. The examples in (98g) and (98j) show that idiom parts do not have to be adjacent to the verb. Apart from this, the idiom can be passivized, which shows that it would be inadequate to list the idiom as a fixed unchangable unit in the lexicon. *die Spatzen* (‘the sparrows’) is a syntactically normal subject that behaves normally in transformations like passivization.

We can conclude this discussion as follows:

- There exist idioms with the subject as part of the idiom and an open slot for the object. (accusative object, dative object/possessive dative, prepositional object, and probably other grammatical functions as well)
- Some of these idioms contain verbs that are not unaccusative.
- There are idioms with fixed subject that can be passivized or form adjectival passive participles.
- There are idioms with fixed subject whose parts (including the subject) can be reordered.

This shows that these idioms are not fixed entities with one open slot but rather units that have an active syntactic live. Hence they contradict Marantz’ claim in the quote above. Since this claim is falsified for one language it cannot hold for all languages. Since approaches to idioms that establish the idiomaticity of an expression after the syntactic analysis were shown to be inadequate (Keil, 1997), the discussion suggests that there has to be a selection relation between heads and their subjects in order to capture the idioms discussed above.

**References**


References


Calcagno, Mike and Pollard, Carl J. 1995. Lexical Rules in HPSG: What are they?,
Department of Linguistics, Ohio State University, Columbus, ms.
Candito, Marie-Hélène. 1996. A Principle-Based Hierarchical Representation of
Cappelle, Bert, Shtyrov, Yury and Pulvermüller, Friedemann. 2010. Heating up
or cooling up the Brain? MEG Evidence that Phrasal Verbs are Lexical Units.
Brain and Language 115, 189–201.
Carlson, Gregory N. and Tanenhaus, Michael K. 1988. Thematic Roles and Lan-
guage Comprehension. In Wendy Wilkins (ed.), Thematic Relations, Syntax
Chomsky, Noam. 1957. Syntactic Structures. Janua Linguarum / Series Minor,
Press.
Chomsky, Noam. 1970. Remarks on Nominalization. In Roderick A. Jacobs and
Peter S. Rosenbaum (eds.), Readings in English Transformational Grammar,
Chapter 12, pages 184–221, Waltham, MA/Toronto/London: Ginn and Com-
pany.
Publications.
LFG-TAG Grammars from a MetaGrammar. In Erhard Hinrichs and Dan Roth
(ed.), Proceedings of the 41st Annual Meeting of the Association for Compu-
tational Linguistics, pages 184–191, Sapporo, Japan: Association for Compu-
tational Linguistics.
Copestake, Ann. 1992. The Representation of Lexical Semantic Information. Cog-
nitive Science Research Papers 280, University of Sussex.
Based Framework. In James Pustejovsky and Sabine Bergler (eds.), Lexical
Semantics and Knowledge Representation, Lecture Notes in Artificial Intelli-
ac.uk/Research/NL/acquilex/papers.html, 18.08.2002.
Croft, William. 2001. Radical Construction Grammar. Syntactic Theory in Typo-
logical Perspective. Oxford University Press.
Hubert Cuyckens, Thomas Berg, René Dirven and Klaus-Uwe Panther (eds.),
Motivation in Language: Studies in honour of Günter Radden, pages 49–68,
Amsterdam: John Benjamins Publishing Co.
Crysmann, Berthold. 2002. Constraint-Based Coanalysis: Portuguese Cliticisa-
tion and Morphology-Syntax Interaction in HPSG. Saarbrücken Dissertations
in Computational Linguistics and Language Technology, No. 15, Saarbrücken:

Deutsches Forschungszentrum für Künstliche Intelligenz und Universität des Saarlandes.


Fillmore, Charles J. 1999. Inversion and Constructional Inheritance. In Gert We-


REFERENCES


Höhle, Tilman N. 1978. *Lexikalische Syntax: Die Aktiv-Passiv-Relation und an-
dere Infinitkonstruktionen im Deutschen. Linguistische Arbeiten, No. 67, Tübingen: Max Niemeyer Verlag.
REFERENCES

Max Niemeyer Verlag.


Müller, Stefan. 1996. Yet another Paper about Partial Verb Phrase Fronting in
Müller, Stefan (ed.). 2007b. *Proceedings of the International Conference on*


Language, No. 8, John Benjamins Publishing Co.

REFERENCES


