Chapter 9

Argument structure and linking

Stephen Wechsler
The University of Texas

Jean-Pierre Koenig
University at Buffalo

Anthony Davis
Southern Oregon University

> What we cannot speak about, we must pass over in silence.

1 Introduction

When a verb or other predicator is composed with the phrases or pronominal affixes expressing its semantic arguments, the grammar must specify the mapping between the semantic participant roles and syntactic dependents of that verb. For example, the grammar of English indicates that the subject of *eat* fills the eater role and the object of *eat* fills the role of the thing eaten. In HPSG this mapping is usually broken down into two simpler mappings by positing an intermediate representation called ARG-ST ('argument structure'). The first mapping connects the participant roles within the semantic CONTENT with the elements of the ARG-ST feature; here we will call the theory of this mapping *linking theory* (see Section 4). The second mapping connects those ARG-ST list elements to the elements of the VALlists, namely COMPS ('complements') and SUBJ ('subject'; or SPR, 'specifier'); we will refer to this second mapping as *argument realization* (see Section 2).[1] These two mappings are illustrated with the simplified lexical sign

---

[1]Some linguists, such as Levin & Rappaport Hovav (2005), use the term 'argument realization' more broadly, to encompass linking as well.

for the verb *eat* in (1).

(1) Lexical sign for the verb *eat*

$$
\begin{bmatrix}
\text{PHON} & \langle \text{EAT} \rangle \\
\text{VALENCE} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{2} \rangle \end{bmatrix} \\
\text{ARG-ST} & \langle \boxed{1}\text{NP}_i , \boxed{2}\text{NP}_j \rangle \\
\text{CONTENT} & \textit{eat}(i, j)
\end{bmatrix}
$$

In (1), 'NP' abbreviates a feature structure representing syntactic and semantic information about a nominal phrase. The variables i and j are the referential indices for the eater and eaten arguments, respectively, of the **eat** relation. The semantic information in NP_i semantically restricts the value or referent of i.

The ARG-ST feature plays an important role in HPSG grammatical theory. In addition to regulating the mapping from semantic arguments to grammatical relations, ARG-ST is the locus of the theories of anaphoric binding and other construal relations such as control and raising. (This chapter focuses on the function of ARG-ST in semantic mapping, with some discussion of binding and other construal relations only insofar as they interact with that mapping. A more detailed look at binding is presented in Chapter **??**.)

In HPSG, verb diathesis alternations, voice alternations, and derivational processes such as category conversions are all captured within the lexicon (see Section 5 and Chapter **??**). The different variants of a word are grammatically related either through lexical rules or by means of the lexical type hierarchy. HPSG grammars explicitly capture paradigmatic relations between word variants, making HPSG a *lexical approach to argument structure*, in the sense of Müller & Wechsler (2014a). This fundamental property of lexicalist theories contrasts with many transformational approaches, where such relationships are treated as syntagmatically related through operations on phrasal structures representing sentences and other syntactic constituents. Arguments for the lexical approach are reviewed in Section 8.

Within the HPSG framework presented here, we will formulate and address a number of empirical questions:

- We know that a verb's meaning influences its valence requirements, (via the ARG-ST list, on this theory). What are the principles governing the mapping from CONTENT to ARG-ST? Are some aspects of ARG-ST idiosyncratically stipulated for individual verbs? What aspects of the semantic CONTENT bear on the value of ARG-ST, and what aspects do not? (For example, what is the role of modality?)

- How are argument alternations defined with respect to our formal system? For each alternation we may ask which of the following it involves: a shuffling of the ARG-ST list; a change in the mapping from ARG-ST to VAL; or a change in the CONTENT, with a concomitant change in the ARG-ST?

These questions will be addressed below in the course of presenting the theory. We begin by considering ARG-ST itself (Section 2), followed by the mapping from ARG-ST to VAL(Section 3) and the mapping from CONTENT to ARG-ST (Section 4). The remaining sections address further issues relating to argument structure: the nature of argument alternations, extending the ARG-ST attribute to include additional elements, whether ARG-ST is a universal feature of languages, and a comparison of the lexicalist view of argument structure presented here with phrasal approaches.

2 The representation of argument structure in HPSG

In the earliest versions of HPSG, the selection of dependent phrases was specified in the SUBCAT feature of the head word (Pollard & Sag (1987), Pollard & Sag (1994: ch. 1–8)). The value of SUBCAT is a list of items, each of which corresponds to the SYNSEM value of a complement or subject. Following are SUBCAT features for an intransitive verb, a transitive verb, and a transitive verb with obligatory PP complement:

(2)   a. *laugh*: [SUBCAT ⟨ NP ⟩]
      b. *eat*: [SUBCAT ⟨ NP, NP ⟩]
      c. *put*: [SUBCAT ⟨ NP, NP, PP ⟩]

Phrase structure rules in the form of immediate dominance schemata identify a certain daughter node as the head daughter (HEAD-DTR) and others as complement daughters (COMP-DTRS). In keeping with the *Subcategorization Principle*, here paraphrased from Pollard & Sag (1994: 34), list items are effectively 'cancelled' from the SUBCAT list as complement phrases are joined with the selecting head:

(3)   Subcategorization Principle: In a headed phrase, the SUBCAT value of the HEAD-DTR ('head daughter') is the concatenation of the phrase's SUBCAT list with the list of SYNSEM values of the COMPS-DTRS ('complement daughters').

Phrasal positions are distinguished by their saturation level: 'VP' is defined as a verbal projection whose SUBCAT list contains a single item, corresponding to the subject; and 'S' is defined as a verbal projection whose SUBCAT list is empty.

The 'subject' of a verb, a distinguished dependent with respect to construal processes such as binding, control, and raising, was then defined as the first item in the SUBCAT list, hence the last item with which the verb combines. However, defining 'subject' as the last item to combine with the head proved inadequate (Pollard & Sag 1994: Ch. 9). There are many cases where the dependent displaying subject properties need not be the last item added to the head projection. For example, in German the construal subject is a nominal in nominative case (Reis 1982), but the language allows subjectless clauses containing only a dative or genitive non-subject NP. If that oblique NP is the only NP dependent to combine with the verb then it is *ipso facto* the last NP to combine.

Consequently, the SUBCAT list was split into two valence lists, a SUBJ list of length zero or one for subjects, and a COMPS list for complements. Nonetheless, certain grammatical phenomena, such as binding and other construal processes, must still be defined on a single list comprising both subject and complements (Manning & Sag 1999). Additionally, some syntactic arguments are unexpressed or realized by affixal pronouns, rather than as subject or complement phrases. The new list containing all the syntactic arguments of a predicator was named ARG-ST ('argument structure').

In clauses without implicit or affixal arguments, the ARG-ST is the concatenation of SUBJ and COMPS respectively. For example, the SUBCAT list for *put* in (2c) is replaced with the following:

$$(4) \quad \begin{bmatrix} \text{PHON} & \langle \text{PUT} \rangle \\ \text{VALENCE} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{2}, \boxed{3} \rangle \end{bmatrix} \\ \text{ARG-ST} & \langle \boxed{1}\text{NP} , \boxed{2}\text{NP}, \boxed{3}\text{PP} \rangle \end{bmatrix}$$

The idealization according to which ARG-ST is the concatenation of SUBJ and COMPS is canonized as the *Argument Realization Principle* (ARP) (Sag et al. 2003: PAGE). Exceptions to the ARP, that is, dissociations between VALENCE and ARG-ST, are discussed in Section 3.2 below.

A predicator's VALENCE lists indicate its requirements for syntactic concatenation with phrasal dependents (Section 3). ARG-ST, meanwhile, provides syntactic information about the expression of semantic roles and is related, via linking theory, to the lexical semantics of the word (Section 3.2). The ARG-ST list contains specifications for the union of the verb's syntactic local phrasal dependents (the

subject and complements, whether they are semantic arguments, raised phrases, or expletives) and its syntactic arguments that are not realized locally, whether they are unbounded dependents, affixal, or unexpressed.

Figure 1 provides a schematic representation of linking and argument realization in HPSG. Linking principles govern the mapping of of participant roles in a predicator's CONTENT to syntactic arguments on ARG-ST. Argument realization is shown in this figure only for mapping to VAL; affixal and null arguments are not depicted. Here, the semantic roles are just arbitrary labels, but we discuss in Section (4) how they can be systematically related to lexical entailments of predicators. The ARG-ST and VALlists in this figure contain only arguments linked to participant roles, but in Section ?? we note motivations for extending ARG-ST to include additional elements. And in Section (3), we examine cases where the relationship between ARG-ST and VALviolates the ARP.

Semantics (CONTENT)

$$\begin{bmatrix} \textit{sem-rel-1} \\ \text{SEM-ROLE-1} \quad \boxed{1} \\ \text{SEM-ROLE-2} \quad \boxed{2} \\ \text{SOA} \quad \begin{bmatrix} \textit{sem-rel-2} \\ \text{SEM-ROLE-3} \quad \boxed{3} \\ \text{SEM-ROLE-4} \quad \boxed{4} \end{bmatrix} \end{bmatrix}$$

Structured representation of semantic relations, their roles and the participants filling them

⇕ *Linking principles*

Argument structure (ARG-ST)

$$\begin{bmatrix} \text{ARG-ST} \quad \langle \text{XP}_1, \text{XP}_2, ...\text{XP}_n \rangle \end{bmatrix}$$

List of syntactic arguments

⇕ *Argument Realization Principle*

Syntax (VALENCE)

$$\begin{bmatrix} \text{SPR} \quad \langle \, \rangle \\ \text{SUBJ} \quad \langle \text{XP}_1 \rangle \\ \text{COMPS} \quad \langle \text{XP}_2, ...\text{XP}_n \rangle \end{bmatrix}$$

Lists of locally realized dependents

Figure 1: How linking works in HPSG

3 Argument realization: The mapping from ARG-ST to VALlists

3.1 Variation in the expression of arguments

The VALfeature is responsible for composing a verb with its phrasal dependents, but this is just one of the ways that semantic arguments of a verb are expressed in natural language. Semantic arguments can be expressed in various linguistic forms: as local syntactic dependents (SUBJ and COMPS), as affixes, or displaced in unbounded dependency constructions (SLASH).

Affixal arguments can be illustrated with the first person singular Spanish verb *hablo* 'speak.1SG', as in 5.

(5) a. Habl-o   español.
         speak-1SG Spanish

         'I speak Spanish.'

  b. *hablo* 'speak.1SG':

$$
\begin{bmatrix}
\text{PHON} & \langle \text{HABLO} \rangle \\[2pt]
\text{VALENCE} & \begin{bmatrix} \text{SUBJ} & \langle\ \rangle \\ \text{COMPS} & \langle \boxed{2} \rangle \end{bmatrix} \\[10pt]
\text{ARG-ST} & \left\langle \text{NP:} \begin{bmatrix} ppro \\ \text{INDEX} \begin{bmatrix} \text{PERS} & 1st \\ \text{NUM} & sg \end{bmatrix} \end{bmatrix}, \boxed{2}\text{NP} \right\rangle
\end{bmatrix}
$$

The *-o* suffix contributes the first person singular pronominal subject content to the verb form (the morphological process is not shown here; see Chapter MORPHOLOGY-CHAPTER). The pronominal subject appears on the ARG-ST list and hence is subject to the binding theory. But it does not appear in SUBJ, if no subject NP appears in construction with the verb.

A lexical sign whose ARG-ST list that is just the concatenation of its SUBJ and COMPS lists conforms to the Argument Realization Principle (ARP); such signs are called *canonical signs* by Bouma et al. (2001). Non-canonical signs, which violate the ARP, have been approached in two ways. In one approach, a lexical rule takes as input a canonical entry and derives a non-canonical one by removing items from the VALlists, while adding an affix or designating an item as an unbounded dependent by placement on the SLASH list. In the other approach, a feature of each ARG-ST list item specifies whether the item is subject to the ARP (hence mapped to a VALlist), or ignored by it (hence expressed in some other way). See

the chapter on the lexicon for more detail and Miller & Sag (1997) for a treatment of French clitics as affixes.

A final case to consider is null anaphora, in which a semantic argument is simply left unexpressed and receives a definite pronoun-like interpretation. Japanese *mi-* 'see' is transitive but the object NP can be omitted as in (6).

(6) Naoki-ga   mi-ta.
     Naoki-NOM see-PAST
     'Naoki saw it/him/her/*himself.'

Null anaphors of this kind typically arise in discourse contexts similar to those that license ordinary weak pronouns, and the unexpressed object often has the (Principle B) obviation effects characteristic of overt pronouns, as shown in (6). But HPSG eschews the use of silent formatives like 'small *pro*' when there is no evidence for such items, such as local interactions with the phrase structure. Instead, null anaphors of this kind are present in ARG-ST but absent from VALlists. ARG-ST is directly linked to the semantic CONTENT and is the locus of binding theory, so the presence of a syntactic argument on the ARG-ST list but not a VALlist accounts for null anaphora. To account for obviation, the ARG-ST list item, when unexpressed, receives the binding feature of ordinary (non-reflexive) pronouns, usually *ppro*. This language-specific option can be captured in a general way by VALand ARG-ST defaults in the lexical hierarchy for verbs.

3.2 The syntax of ARG-ST and its relation to VALENCE

The ARG-ST ordering represents a preliminary syntactic structuring of the set of argument roles. In that sense it functions as an interface between the lexical semantics of the verb, and the expressions of dependents as described in Section 3. Its role thus bears some relation to the initial stratum in Relational Grammar, *argument structure* (including intrinsic classifications) in LFG Lexical Mapping Theory, D-structure in Government/Binding theory, and the Merge positions of arguments in Minimalism, assuming in the last two cases the Uniform Thematic Alignment Hypothesis (Baker 1988) or something similar. However, it also differs from all of those in important ways.

Semantic constraints on ARG-ST are explored in Section 4 below. But ARG-ST is not only structured by semantic distinctions between the arguments but also by syntactic ones. Specifically, the list ordering represents relative syntactic *obliqueness* of arguments. The least oblique argument is the subject (SUBJ), followed by

the complements (COMPS). Following Manning (1996) term arguments (direct arguments, i.e. objects) are assumed to be less oblique than 'oblique' arguments (adpositional and oblique case marked phrases), followed finally by predicate and clausal complements. The transitive ordering relation on the ARG-ST list is called *o-command* ('obliqueness command'): the subject list item o-commands those of the complements; an object list item o-commands those of any obliques; and so on.

Voice alternations like the passive, which are defined on the ARG-ST list, illustrate the ordering of terms before obliques on the ARG-ST list. Passivization alters the syntactic properties of ARG-ST list items: the initial item of the active, normally mapped to SUBJ of the active, is an oblique (*by* phrase) or unexpressed argument in the passive. Given that terms precede obliques in the list order, any term arguments must o-command the passive oblique, so passive effectively reorders the initial item in ARG-ST to a list position following any terms.

(7) a. Susan gave Mary a book.
     b. Mary was given a book by Susan.

(8) a. *give* (active): [ARG-ST $< \text{NP}_i, \text{NP}_j, \text{NP}_k >$]
     b. *given* (passive): [ARG-ST $< \text{NP}_j, \text{NP}_k, \text{PP}[by]_i >$]

Relative obliqueness conditions a number of syntactic processes and phenomena, including anaphoric binding. The o-command relation replaces the c-command in the Principles A, B, and C of Chomsky's (1981) configurational theory of binding. For example, HPSG's Principle B states that an ordinary pronoun cannot be o-commanded by its coargument antecedent, which accounts for the pronoun obviation observed in the English sentence *Naoki$_i$ saw him$_{*i/j}$*, and also accounts for obviation in the Japanese sentence (6) above.

Relative obliqueness also conditions the accessibility hierarchy of Keenan & Comrie (1977), according to which a language allowing relativization of some type of dependent also allows relativization of any less oblique than it. Hence if a language has relative clauses at all, it has subject relatives; if it allows obliques to relativize then it also allows subject and object relatives; and so on. Similar implicational universals apply to verb agreement with subjects, objects, and obliques Greenberg (1966).

Returning now to argument realization, we saw above that the rules for the selection of the subject from among the verb's arguments are also stated on the ARG-ST list. In a 'canonical' realization the subject is the first list item, o-commanding all of its coarguments.

3.3 Syntactic ergativity

The autonomy of ARG-ST from the VALlists is further illustrated by cross-linguistic ▮
variation in the mapping between them. As just noted, in English and many other
languages the initial item in ARG-ST maps to the subject. However, languages
with so-called *syntactically ergative* clauses have been analyzed as following a
different mapping rule. Crucially, the ARG-ST ordering in those languages is still
supported by independent evidence from properties such as binding and NP ver-
sus PP categorial status of arguments. Balinese (Austronesian), as analyzed by
Wechsler & Arka (1998), is such a language. In the morphologically unmarked,
and most common voice, called 'Objective voice' (OV), the subject is any term
except the ARG-ST-initial one.

Balinese canonically has SVO order, regardless of the verb's voice form (Artawa ▮
1994; Wechsler & Arka 1998). The preverbal NPs in (9) are the surface subjects and
the postverbal ones are complements. When the verb appears in the unmarked
OV verb, a non-initial term is the subject, as in (9a). But verbs in 'Agentive Voice'
(AV) select as their subject the ARG-ST-initial item, as in (9b).

(9) a. Bawi adol   ida.
         pig   OV.sell 3sg
         'He/She sold a pig.'

  b. Ida ng-adol bawi.
       3sg AV-sell pig
       'He/She sold a pig.'

A ditransitive verb such as the benefactive applied form of *beli* 'buy' in (10), has
three term arguments on its ARG-ST list. The subject can be either term that is
non-initial in ARG-ST:

(10) a. Potlote   ento beli-ang=a   I   Wayan.
          pencil-DEF that OV.buy-APPL=3 Art Wayan
          '(s)he bought Wayan the pencil.'

  b. I   Wayan beli-ang=a   potlote   ento.
       Art Wayan OV.buy-APPL=3 pencil-DEF that
       '(s)he bought Wayan the pencil.'

Unlike the passive voice, which reorders the ARG-ST list, the Balinese OV does
not affect ARG-ST list order. Thus the agent argument can bind a coargument
reflexive pronoun (but not vice versa), regardless of whether the verb is in OV or
AV form:

(11) a. Ida ny-ingakin ragan idane.
          3sg AV-see   self

          '(s)he saw himself/herself'

      b. Ragan idane cingakin ida.
          self   OV.see   3sg

          '(s)he saw himself/herself'

The 'seer' argument o-commands the 'seen', with the AV versus OV voice forms regulating subject selection:

(12) Agentive Voice form of 'see':

$$
\begin{bmatrix}
\text{PHON} & \langle nyinkagin \rangle \\
\text{VALENCE} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{2} \rangle \end{bmatrix} \\
\text{ARG-ST} & \langle \boxed{1}\text{NP}_i , \boxed{2}\text{NP}_j \rangle \\
\text{CONTENT} & \begin{bmatrix} see\text{-}rel \\ \text{SEER} & i \\ \text{SEEN} & j \end{bmatrix}
\end{bmatrix}
$$

(13) Objective Voice form of 'see':

$$
\begin{bmatrix}
\text{PHON} & \langle cinkagin \rangle \\
\text{VALENCE} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{2} \rangle \\ \text{COMPS} & \langle \boxed{1} \rangle \end{bmatrix} \\
\text{ARG-ST} & \langle \boxed{1}\text{NP}_i , \boxed{2}\text{NP}_j \rangle \\
\text{CONTENT} & \begin{bmatrix} see\text{-}rel \\ \text{SEER} & i \\ \text{SEEN} & j \end{bmatrix}
\end{bmatrix}
$$

Languages like Balinese illustrate the autonomy of ARG-ST. Although the agent binds the patient in both (11)a and b, the binding conditions cannot be stated directly on the thematic hierarchy. For example, in HPSG a raised argument appears on the ARG-ST list of the raising verb even though that verb assigns no thematic role to that list item. But a raised subject can bind a coargument reflexive in Balinese (this is comparable to English *John seems to himself to be ugly*). Anaphoric binding in Balinese raising constructions thus behaves as predicted by the ARG-ST based theory (Wechsler 1999). In conclusion, neither VAL nor CONTENT provides the right representation for defining binding conditions, but ARG-ST fills the bill.

Syntactically ergative languages that have been analyzed as using an alternative mapping between ARG-ST and VALinclude Tagalog, Inuit, some Mayan languages, Chukchi, Toba Batak, Tsimshian languages, and Nadëb (Manning 1996; Manning & Sag 1999).

Interestingly, while the GB/Minimalist configurational binding theory may be defined on analogues of VALor CONTENT, those theories lack any analogue of ARG-ST. This leads to special problems for such theories in accounting for binding in many Austronesian languages like Balinese. In transformational theories since Chomsky (1981), anaphoric binding conditions are usually stated with respect to the A-positions ('argument positions'). A-positions are analogous to HPSG VALlist items, with relative c-command in the configurational structure corresponding to relative list ordering in HPSG, in the simplest cases. Meanwhile, to account for data similar to (11), where agents asymmetrically bind patients, Austronesian languages like Balinese were said to define binding on the 'thematic structure' encoded in d-structure or Merge positions, where agents asymmetrically c-command patients regardless of their surface positions (Guilfoyle et al. 1992). But the interaction with raising shows that neither of those levels is appropriate as the locus of binding theory (Wechsler 1999).[2]

3.4 Symmetrical objects

We have thus far tacitly assumed a total ordering of elements on the ARG-ST list, but Ackerman et al. (2018), Ackerman et al. (2017) propose a partial ordering for certain so-called 'symmetrical object' languages. In Moro (Kordofanian), the two term complements of a ditransitive verb have exactly the same object properties. Relative linear order of the theme and goal arguments is free, as shown by the two translations of (14) (from Ackerman et al. 2017: 9):

(14) é-g-a-natʃ-ó   óráŋ   ɲeɾá
      1SB.SM-CLg-MAIN-give-PFV CLg.man CLɪŋ.girl

      'I gave the girl to the man.' / 'I gave the man to the girl.'

More generally, the two objects have identical object properties with respect to occurrence in post-predicate position, case marking, realization by an object marker, and ability to undergo passivization (Ackerman et al. 2017: 9).

---

[2]To account for (11b) under the configurational binding theory, the subject position must be an A-bar position; but to account for binding by a raised subject, it must be an A-position. See Wechsler (1999).

[Ackerman et al. (2017)](#) propose that the two objects are unordered on the ARG-ST list. This allows for two different mappings to the COMPS list, as shown here:

(15)   a.   Goal argument as primary object:

$$
\begin{bmatrix}
\text{VALENCE} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{2}, \boxed{3} \rangle \end{bmatrix} \\
\text{ARG-ST} & \langle \boxed{1}\text{NP}_i, \{\boxed{2}\text{NP}_j, \boxed{3}\text{NP}_k\} \rangle \\
\text{CONTENT} & \begin{bmatrix} \textit{give-rel} \\ \text{AGENT} & i \\ \text{GOAL} & j \\ \text{THEME} & k \end{bmatrix}
\end{bmatrix}
$$

      b.   Theme argument as primary object:

$$
\begin{bmatrix}
\text{VALENCE} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{3}, \boxed{2} \rangle \end{bmatrix} \\
\text{ARG-ST} & \langle \boxed{1}\text{NP}_i, \{\boxed{2}\text{NP}_j, \boxed{3}\text{NP}_k\} \rangle \\
\text{CONTENT} & \begin{bmatrix} \textit{give-rel} \\ \text{AGENT} & i \\ \text{GOAL} & j \\ \text{THEME} & k \end{bmatrix}
\end{bmatrix}
$$

The primary object properties, which are associated with the initial term argument of COMPS, can go with either the goal or theme argument.

To summarize this section, while the relationship between ARG-ST, SUBJ, and COMPS lists was originally conceived as a straightforward one, enabling binding principles to maintain their simple form by defining ARG-ST as the concatenation of the other two, the relationship was soon loosened. Non-canonical relationships between ARG-ST and the VALlists are invoked in accounts of several core syntactic phenomena. Arguments not realized overtly in their canonical positions, due to extraction, cliticization, or pro-drop (null anaphora), appear on ARG-ST but not in any VALlist. Accounts of syntactic ergativity in HPSG involve variations in the mapping between ARG-ST and VALlists; in particular, the element of SUBJ is not, in such languages, the first element of ARG-ST. Modifications of ARG-ST play a role in treatments of passivization, where its expected first element is suppressed, and in languages with multiple, symmetric objects, where a partial rather than total ordering of ARG-ST elements has been postulated. Thus ARG-ST has now acquired an autonomous status within HPSG, and is not merely a predictable rearrangement of information present elsewhere in lexical entries.

4 Linking: the mapping between semantics and ARG-ST

4.1 HPSG approaches to linking

The term *linking* refers to the mapping specified in a lexical entry between participant roles in the semantics and their syntactic representations on the ARG-ST list. Early HPSG grammars stipulated the linking of each verb: semantic CONTENT values with predicator-specific attributes like DEVOURER and DEVOURED were mapped to the subject and object, respectively, of the verb *devour*. But linking observes prevailing patterns, e.g. if one argument of a transitive verb in active voice has an agentive role, it will map to the subject, not the object. Thus these early accounts were unsatisfying, as they lead to purely stipulative accounts of linking, specified verb by verb. Beginning with Wechsler (1995b) and Davis (1996), researchers formulated linking principles stated on more general semantic properties holding across verbs.

Within the history of linguistics there have been three general approaches to modeling the lexico-semantic side of linking: thematic role types (Pāṇini ca 400 B.C., Fillmore 1968); lexical decomposition (Foley & Van Valin 1984; Rappaport Hovav & Levin 1998); and the proto-roles approach (Dowty 1991). In developing linking theories within the HPSG framework Wechsler (1995b) and Davis (1996) employed a kind of lexical decomposition that also incorporated some elements of the proto-roles approach. The reasons for preferring this over the alternatives are discussed in Section 4.4 below.

Wechsler's 1995b linking theory constrains the relative order of pairs arguments on the ARG-ST list according to semantic relations entailed between them. For example, his *notion rule* states that if one participant in an event is entailed to have a mental notion of another, then the first must precede the second on the ARG-ST list. The *conceive-pred* type is defined by the following type declaration (based on Wechsler (1995b: 127), with formal details adjusted for consistency with current usage):

(16) *conceive-pred:*
$$\begin{bmatrix} \text{ARG-ST} & \langle \text{NP}_i, \text{NP}_j \rangle \\ \text{CONTENT} & \begin{bmatrix} \textit{conceive-rel} \\ \text{CONCEIVER} & i \\ \text{CONCEIVED} & j \end{bmatrix} \end{bmatrix}$$

This accounts for a host of linking facts in verbs as varied as *like, enjoy, invent, claim,* and *murder,* assuming these verbs belong to the type *conceive-pred.* It explains the well known contrast between experiencer-subject *fear* and experiencer-object *frighten* verbs: *fear* entails that its subject has some notion of its object, so

The tourists feared the lumberjacks entails that the tourists are aware of the lumberjacks. But the object of *frighten* need not have a notion of its subject: in *The lumberjacks frightened the tourists (by cutting down a large tree that crashed right in front of them*, the tourists may not be aware of the lumberjacks' existence.

Two other linking rules appeared in Wechsler (1995). One stated that 'affected themes,' that is, participants that are entailed to undergo a change, map to the object, rather than subject, of a transitive verb. Another pertained to stative transitive verbs entailing a part-whole relation between the two participants, such as *include* and *contain*: the whole maps to the subject and the part to the object.

The linking constraints do not rely on a total ordering of thematic roles, nor on an exhaustive asssignment of thematic role types to every semantic role in a predicator. Instead, a small set of partial orderings of semantic roles, based on lexical entailments, suffices to account for the linking patterns of a wide range of verbs. This insight was adopted in a slightly different guise in work by Davis (1996; 2001) and Davis & Koenig (2000), who develop a more elaborated representation of lexical semantics, with which simple linking constraints can be stated. The essence of this approach is to posit a small number of dyadic semantic relations such as *act-und-rel* ('actor-undergoer relation') with attributes ACT(OR) and UND(ERGOER) that serve as intermediaries between semantic roles and syntactic arguments (akin to the notion of Generalized Semantic Roles discussed in Van Valin 1999).

What are the truth conditions of *act-und-rel*? Following Fillmore (1977), Dowty (1991), and Wechsler (1995b), Davis and Koenig note that many of the pertinent lexical entailments come in related pairs. For instance, one of Dowty's entailments is that one participant causally affects another, and of course the other is entailed to be causally affected. Another involves the entailments in Wechsler's notion rule (16); one participant is entailed to have a notion of another. These entailments of paired participant types characterize classes of verbs (or other predicators), and can then be naturally represented as dyadic relations in CONTENT. Collecting those entailments we arrive at a disjunctive statement of truth conditions:

(17)   **act-und-rel**(x, y) is true iff x causes a change in y, or x has a notion of y.

We can designate the x participant in the pair as the value of ACTOR (or ACT) and y as the value of UNDERGOER (or UND), in a relation of type *act-und-rel*. Semantic arguments that are ACTOR or UNDERGOER will then bear at least one of the entailments characteristic of ACTORS or UNDERGOERS (Davis & Koenig 2000: 72). This then simplifies the statement of linking constraints for all of these paired

Draft of 9th September 2019, 12:08

participant types. Davis (1996) and Koenig & Davis (2001) argue that this obviates counting the relative number of proto-agent and proto-patient entailments, as advocated by Dowty (1991).

The linking constraints 18 and 19 state that a verb whose semantic CONTENT is of type *act-und-rel* will be constrained to link the ACT participant to the the first element of the verb's ARG-ST list (its subject), and the UND participant to the second element of the verb's ARG-ST list (this is analogous to Wechsler's constraints based on partial orderings).

These linking constraints can be viewed as parts of the definition of lexical types, as in Davis (2001), where (18) defines a particular class of lexemes (or words).[3]

(18)
$$\begin{bmatrix} \text{CONTENT|KEY} & \begin{bmatrix} \text{ACTOR} & \boxed{1} \end{bmatrix} \\ \text{ARG-ST} & \left\langle \text{NP}_{\boxed{1}}, ... \right\rangle \end{bmatrix}$$

(19)
$$\begin{bmatrix} \text{CONTENT|KEY} & \begin{bmatrix} \text{UNDERGOER} & \boxed{2} \end{bmatrix} \\ \text{ARG-ST} & \left\langle ..., \text{NP}_{\boxed{2}}, ... \right\rangle \end{bmatrix}$$

(20)
$$\begin{bmatrix} \text{CONTENT|KEY} & \begin{bmatrix} \textit{cause-possess-rel} \\ \text{SOA} & \begin{bmatrix} \text{ACTOR} & \boxed{3} \end{bmatrix} \end{bmatrix} \\ \text{ARG-ST} & \left\langle \textit{SYNSEM} \right\rangle \oplus \left\langle \text{NP}_{\boxed{3}}, ... \right\rangle \end{bmatrix}$$

The first constraint, in (18), links the value of ACT (when not embedded within another attribute) to the first element of ARG-ST. The second, in (19), merely links the value of UND (again, when not embedded within another attribute) to some NP on ARG-ST. Given this understanding of how the values of ACTOR and UNDERGOER are determined, these constraints cover the linking patterns of a wide range of transitive verbs: *throw* (ACT causes motion of UND), *slice* (ACT causes change of state in UND), *frighten* (ACT causes emotion in UND), *imagine* (ACT has a notion of UND), *traverse* (ACT "measures out" UND as an incrememtnal theme), and *outnumber* (ACT is superior to UND on a scale).

---

[3]Alternatively, (18) (and other linking constraints) can be recast as implicational constraints on lexemes or words (Koenig & Davis 2003). (i) is an implicational constraint indicating that a word whose semantic content includes an ACTOR role must map that role to the initial item in the ARG-ST list.

(i)
$$\begin{bmatrix} \text{CONTENT|KEY} & \begin{bmatrix} \text{ACTOR} & \boxed{1} \end{bmatrix} \end{bmatrix} \Rightarrow \begin{bmatrix} \text{ARG-ST} & \left\langle \text{NP}_{\boxed{1}}, ... \right\rangle \end{bmatrix}$$

The third constraint, in (20), links the value of an ACT attribute embedded within a SOA attribute to an NP that is second on ARG-ST. This constraint accounts for the linking of the (primary) object of ditransitives. In English, these verbs (*give*, *hand*, *send*, *earn*, *owe*, etc.) involve (prospective) causing of possession (Pinker 1989; Goldberg 1995), and the possessor is represented as the value of the embedded ACT in (20). There could be additional constraints of a similar form in langages with a wider range of ditransitive constructions; conversely, such a constraint might be absent in languages that lack ditransitives entirely. As mentioned earlier in this section, the range of subcategorization options varies somewhat from one language to another.

The KEY attribute in (18) – (20) also requires explanation. The formulation of linking constraints here employs the architecture used in Koenig & Davis (2006), in which the semantics represented in CONTENT values is expressed as a set of *elementary predications*, formalized within Minimal Recursion Semantics (Copestake et al. 2001; 2005). Each elementary predication is a simple relation, but the relationships among them may be left unspecified. For linking, one of the elementary predications is designated the KEY, and it serves as the locus of linking. This allows us to indicate the linking of participants that play multiple roles in the denoted situation. The KEY selects one relation as the "focal point," and the other elementary predications are then irrelevant as far as linking is concerned. The choice of KEY then becomes an issue demanding consideration; we will see in the discussion of argument alternations in Section 5 how this choice might account for some alternation phenomena.

Note too that these linking constraints are treated as constraints on classes in the lexical hierarchy (see Chapter **??**). One consequence of this fact merits brief mention. Constraint (19), which links the value of UND to some NP on ARG-ST, is a specification of one class of verbs. Not all verbs (and certainly not all other predicators, such as nominalizations) with a CONTENT value containing an UND value realize it as an NP. Verbs obeying this constraint include the transitive verbs noted above, and intransitive "unaccusative" verbs such as *fall* and *persist*. But some verbs with both ACT and UND attributes in their CONTENT are intransitive, such as *impinge (on)*, *prevail (on*, and *tinker (with)*. Interactions with other constraints, such as the requirement that verbs (in English, at least) have an NP subject, determine the range of observed linking patterns.

These linking constraints also assume that the proto-role attributes ACTOR, UNDERGOER, and SOA are appropriately matched to entailments, as described above. Other formulations are possible, such as that of Koenig & Davis (2003), where the participant roles pertinent to each lexical entailment are represented in CONTENT

by corresponding, distinct attributes.

In addition to the linking constraints, there may be some very general well-formedness conditions on linking. We rarely find verbs that obligatorily map one semantic role to two distinct members of the ARG-ST list that are both expressed overtly. A verb meaning 'eat', but with that disallowed property, could appear in a ditransitive sentence like (4.1), with the meaning that Pat ate dinner, and his dinner was a large steak.

(21)   *Pat ate dinner a large steak.

Typically semantic arguments map to at most one (overtly expressed) ARG-ST list item (Davis 2001: 262-268).

4.2 Linking oblique arguments

In this section we discuss linking of oblique arguments, that is, PP's and oblique case marked NP's. In some instances, a verb's selection of a particular preposition appears at least partly arbitrary; it is hard to explain why we *hanker after* and *yearn for*, but we don't **yearn after*. In these cases, the choice of preposition may be stipulated by the individual lexical entry. But as Gawron (1986) and Wechsler (1995a) have shown, many prepositions are semantically meaningful. *For* in the above-mentioned cases, and in *look for*, *wait for*, and *aim for* is surely not a lexical accident. And in the cases like *cut with*, *with* is used in an instrumental sense, denoting a *use-rel* relation, as with verbs that either allow (*eat*) or require (*cut*) an instrument. Davis (1996; 2001) adopts the position of Gawron and Wechsler in his treatment of linking to PPs As an example of this kind of account, the linking type in (22) characterizes a verb selecting a *with*-PP. The PP argument is linked from the RELS list rather the KEY.

(22)

$$
\begin{bmatrix}
\text{CONTENT} & \begin{bmatrix} \text{KEY} & \boxed{1} \\ \text{RELATIONS} & \left\langle \boxed{1}, \boxed{2} \begin{bmatrix} \textit{use-rel} \\ \text{ACT} & a \\ \text{UND} & u \\ \text{SOA} & s \end{bmatrix}, ... \right\rangle \end{bmatrix} \\
\text{ARG-ST} & \left\langle ..., \text{PP}_{with}\text{:}\boxed{2} ... \right\rangle
\end{bmatrix}
$$

Apart from the details of individual linking constraints, we have endeavored here to describe how linking can be modeled in HPSG using the same kinds of constraints used ubiquitously in the framework. Within the hierarchical lexicon, constraints between semantically defined classes and syntactically defined

ones, can furnish an account of linking patterns, and there is no resort to additional mechanisms such as a thematic hierarchy or numerical comparison of entailments.

4.3 To what extent does meaning predict linking?

The framework outlined above allows us to address the following question: how much of linking is strictly determined by semantic factors, and how much is left open to lexically arbitrary subcategorization specifications, or perhaps subject to other factors?

Subcategorization— the position and nature of ARG-ST elements, in HPSG terms— is evidently driven to a great extent by semantics, but debate continues about how much, and which components of semantics are involved. Views have ranged from the strict, highly constrained relationship in which lexical semantics essentially determines syntactic argument structure, to a looser one in which some elements of subcategorization may be stipulated. Among the first camp are those who espouse the Uniformity of Theta Assignment Hypothesis proposed in Baker (1988: 46) or Baker (1997), which maintains that "identical thematic relationships between items are represented by identical structural relationships" in the syntax. With regard to the source of diathesis alternations, Levin (1993: 12-13) notes that "studies of these properties suggest that argument structures might in turn be derivable to a large extent from the meaning of words", and accordingly "pursues the hypothesis of semantic determinism seriously to see just how far it can be taken."

Others, including Pollard & Sag (1987) (Section 5.3) and Davis (2001) (Section 5.1), have expressed caution, pointing out cases where subcategorization and diathesis alternations seem to be at least partly arbitrary. Pollard & Sag note contrasts like these:

(23)   a. Sandy spared/*deprived Kim a second helping.
        b. Sandy *spared/deprived Kim of a second helping. (Pollard & Sag 1987: ex. 214–215)

And Davis provides these pairs of semantically similar verbs with differing subcategorization requirements:

(24)   a. Few passengers waited for/awaited the train.
        b. Homer opted for/chose a chocolate frosted donut.
        c. The music grated on/irritated the critics. (Davis 2001: ex. 5.4)

Other cases where argument structure seems not to mirror semantics precisely include raising constructions, in which one of a verb's direct arguments bears no semantic role to it at all. Similarly, overt expletive arguments cannot be seen as deriving from some participant role in a predicator's semantics. Like the examples above, these phenomena suggest that some aspects of subcategorization are specified independently of semantics.

Another point against strict semantic determination of argument structure comes from cross-linguistic observations of subcategorization possibilities. It is evident, for example, that not all languages display the same range of direct argument mappings. Some lack ditransitive constructions entirely (Halkomelem), some allow them across a limited semantic range (English), some quite generally (Georgian), and a few permit tritransitives (Kinyarwanda and Moro). Gerdts (1992) surveys about twenty languages and describes consistent patterns like these. The range of phenomena such as causative and applicative formation in a language is constrained by what she terms its "relational profile;" this includes, in HPSG terms, the number of direct NP arguments permitted on its ARG-ST lists. Again, it is unclear that underlying semantic differences across languages in the semantics of verbs meaning 'give' or 'write' would be responsible for these general patterns.

Summarizing, there is much evidence tempting us to derive the contents of ARG-ST solely from lexical semantics. If this ultimately proves feasible, then ARG-ST serves more as a convenient interface notion with little possibility of independently expressing strictly syntactic aspects of subcategorization. This view, however satisfying it might be, does not accord with our current best understanding of the syntactic and semantic evidence. In the following sections we delve into some of the nuances that make linking more than a simple rendering of lexical semantics. We begin by noting a point on which HPSG accounts of linking differ from many others— the absence of traditional thematic roles.

4.4 HPSG and thematic roles

The ARG-ST list constitutes the syntactic side of the mapping between semantic roles and syntactic dependents. As ARG-ST is merely an ordered list of arguments, without any semantic "labels," it contains no counterparts to thematic roles, such as AGENT, PATIENT, THEME, or GOAL. Thematic roles like these, however, have been a mainstay of linking in generative grammar since Fillmore (1968) and have antecedents going back to (Pāṇini. Ranking them in a *thematic hierarchy*, and labeling each of a predicator's semantic roles with a unique thematic role, then yields an ordering of roles analogous to the ordering on the ARG-ST list. Indeed, it

would not be difficult to import this kind of system into HPSG, as a means of determining the order of elements on the ARG-ST list. However, HPSG researchers have generally avoided using a thematic hierarchy, for reasons we now briefly set out.

Fillmore (1968) and many others thereafter have posited a small set of disjoint, thematic roles, with each semantic role of a predicator assigned exactly one thematic role. Thematic hierarchies depend on these properties for a consistent linking theory. But they do not hold up well to formal scrutiny. Jackendoff (1987) and Dowty (1991) note (from somewhat different perspectives) that numerous verbs have arguments not easily assigned a thematic role from the typically posited inventory (e.g., the objects of *risk*, *blame*, and *avoid*), that more than one argument might sensibly be assigned the same role (e.g., the subjects and objects of *resemble*, *border*, and some alternants of commercial transaction verbs), and that multiple roles can be sensibly assigned to a single argument (the subjects of verbs of volitional motion are like both an AGENT and a THEME). In addition, consensus on the inventory of thematic roles has proven elusive, and some, notoriously THEME, have resisted clear definition. Work in formal semantics, including Ladusaw & Dowty (1988), Dowty (1989), Landman (2000), and Schein (2002), casts doubt on the prospects of assigning formally defined thematic roles to all of a predicator's arguments, at least in a manner that would allow them to play a crucial part in linking. Thematic role types seem to pose problems, and there are alternatives that avoids those problems. As Carlson (1998) notes about thematic roles, "It is easy to conceive of how to write a lexicon, a syntax, a morphology, a semantics, or a pragmatics without them."

4.5 CONTENT decomposition and ARG-ST

Instead of thematic role types, lexical decomposition is typically used in HPSG to model the semantic side of the linking relation. The word meaning represented by the CONTENT value is decomposed into a set of elementary predications that share arguments, as described in Section 4 above. Lexical decompositions cannot be directly observed, but the decompositions are justified indirectly by the roles they play in the grammar. Decompositions play a role in at least the following processes:

- *Linking.* As described in Section 4, linking constraints are stated on semantic relations like *act-und-rel* ('actor-undergoer relation'), so those relations must be called out in the CONTENT field.

- *Sublexical scope.* Certain modifiers can scope over a part of the situation denoted by a verb (Dowty 1979).

  (25) John sold the car, and then he bought it again.

  In this sentence the adverb *again* either adds the presupposition that John bought it before, or, in the more probable interpretation, it adds the presupposition that *the result of buying the car* obtained previously. The result of buying a car is owning it, so this sentence presupposes that John previously owned the car. Thus the decomposition of the verb *buy* in (36) below includes the *possess-rel* ('possession relation') holding between the buyer and the goods. This is available for modification by adverbials like *again*.

- *Argument alternations.* Some argument alternations can be modeled as highlighting of different portions of a single lexical decomposition. See Section 5.

In general, sublexical decompositions are included in the CONTENT field only insofar as they are visible to the grammar for processes like these.

The ARG-ST feature lies at the syntax side of the linking relation. Much like the CONTENT field, the ARG-ST items are justified only insofar as they are visible to the syntax. Many ARG-ST list items are obviously justified by being explicitly expressed as subject and complement phrases or as affixal pronouns. Certain implicit arguments appear if they are subject to the binding theory as applied to the ARG-ST list (as discussed in Section 3.1 above).

Implicit arguments can also participate in the syntax, and therefore appear on the ARG-ST list, by acting as controllers of adjunct clauses. For example, English rationale clauses like the infinitival phrase in (26a) are controlled by the agent argument in the clause, *the hunter* in this example. The implicit agent of a short passive can likewise control the rationale clause as shown in(26b). But the middle in (26c) lacks an implicit agent that is capable of controlling, even though native speakers assume that some agent must have caused the gun to load. This contrast was observed by Keyser & Roeper (1984) and confirmed in experimental work by Mauner & Koenig (2000).

(26) a. The shotgun was loaded quietly by the hunter to avoid the possibility of frightening off the deer.
    b. The shotgun was loaded quietly to avoid the possibility of frightening off the deer.

> c. *The shotgun had loaded quietly to avoid the possibility of frightening off the deer.

If the syntax of control is specified such that the controller of the rationale clause is an (agent) argument on the ARG-ST list of the verb, then this contrast is captured by assuming that the agent appears on the ARG-ST list of the passive verb but not the middle.

4.6 Modal transparency

Another observation concerning lexical entailments and linking was developed by Koenig & Davis (2001), who point out that linking appears to ignore modal elements of lexical semantics, even when those elements invalidate entailments (expanding on an observation implicit in Goldberg 1995). For instance, there are various English verbs displaying linking patterns like the ditransitive verbs of possession transfer *give* and *hand*, but which denote situations in which the transfer need not, or does not, take place. Thus, *offer* describes a situation where the transferor is willing to effect the transfer, *owe* one in which the transferor should effect the transfer but has not yet, *promise* describes a situation where the transferor commits to effect the transfer, and *deny* one in which the transferor does not effect the contemplated transfer. Koenig and Davis argue that modal elements should be clearly separated in CONTENT values from the representations of predicators and their arguments. (27) exemplifies this factoring out of sublexical modal information from core situational information. This pattern of linking functioning independently of sublexical modal information applies not only to these ditransitive cases, but also to verbs involving possession (cf. *own* and *obtain*, vs. *lack*, *covet* and *lose*), perception (*see* vs. *ignore* and *overlook*), and carrying out an action (*manage* vs. *fail* and *try*). Whatever the role of lexical entailments in linking, then, the modal information should be factored out, since the entailments canonically driving, e.g., the ditransitive linking patterns of verbs like *give* and *hand*, do not hold of *offer*, *owe*, or *deny*. The constraints in (18)-(20) need only been minimally altered to target the value of SIT-CORE within the representation of relation.

(27)   The lexical semantic representation of *promise* (Koenig & Davis 2001: 101)

$$
\begin{bmatrix}
promise\text{-}sem \wedge cause\text{-}possess\text{-}sem \\[4pt]
\text{SIT-CORE} \quad \boxed{3}
\begin{bmatrix}
cause\text{-}possess\text{-}rel \\
\text{ACTOR} \quad \boxed{1} \\
\text{UNDERGOER} \quad \boxed{2} \\[4pt]
\text{SOA} \quad
\begin{bmatrix}
\text{sit-core} \; \boxed{5}
\begin{bmatrix}
have\text{-}rel \\
\text{ACTOR} \quad \boxed{2} \\
\text{UNDERGOER} \quad \boxed{4}
\end{bmatrix}
\end{bmatrix}
\end{bmatrix} \\[10pt]
\text{MODAL-BASE} \quad
\left\langle
\begin{bmatrix}
deontic\text{-}mb \wedge condit\text{-}satis\text{-}mb \\
\text{SOA} \quad \boxed{3}
\end{bmatrix}
\right\rangle
\end{bmatrix}
$$

5 The semantics and linking of argument alternations

A verb can often occur in varied syntactic contexts, as *find* does in (28); these are termed *valence alternations* or *diathesis alternations*, in reference to their different argument structuress. Levin (1993) lists around 50 kinds of alternations in English, and there are still more, including the alternation illustrated in (28).

(28) a. I found that the chair was comfortable
      b. I found the chair to be comfortable
      c. I found the chair comfortable

Another well studied alternation, the locative alternation, is exemplified by the two uses of *spray* in (29).

(29) a. *spray$_{loc}$*: Joan sprayed the paint onto the statue.
      b. *spray$_{with}$*: Joan sprayed the statue with paint.

It is typically assumed that these two different uses of *spray* in (29) have slightly different meanings, with the statue being in some sense more affected in the *with* alternant. This exemplifies the "holistic" effect of direct objecthood, which we will return to. Here, we will examine how semantic differences between alternant relate to their linking patterns. The semantic side of linking has often been devised with an eye to syntax (e.g., Pinker (1989), and see Koenig & Davis (2006) for more examples). There is a risk of stipulation here, without independent evidence for these semantic differences. In the case of locative alternations, though, the meaning difference between (29a) and (29b) is easily stated (and Pinker had the right intuition), as (29b) entails (29a), bot not conversely. Informally, (29a) describes a particular kind of caused motion situation, while (29b) describes a situation in which this kind of caused motion additionally results in a caused change of state. The difference is depicted in the two structures in (30).

(30)   a. CAUSE (JOAN, GO (PAINT, TO (STATUE)))

          b. ACT-ON (JOAN, STATUE, BY (CAUSE (JOAN, GO (PAINT, TO (STATUE)))))

This description of the semantic difference between sentences (29a) and (29b) provides a strong basis for predicting their different argument structures. But we still need to explain how linking principles give rise to this difference. Pinker's account rests on semantic structures like (30), in which depth of embedding reflects sequence of causation, with ordering on ARG-ST stemming from depth of semantic embedding, a strategy adopted in Davis (1996; 2001). This is one reasonable alternative, although the resulting complexity of some of the semantic rpresentations raises valid questions about what independent evidence supports them. An alternative appears in Koenig & Davis (2006), who borrow from Minimal Recursion Semantics (see the chapter on Semantics for an introduction to MRS). MRS "flattens" semantic relationss, rather than embedding them in one another, so the arrangment of these *elementary predications*, as they are temed, is of less import. They posit a RELATIONS (or RELS) attribute that collects a set of elementary predications, each representing some part of the predicator's semantics. A KEY attribute specifies a particular member of RELS as the relevant one for linking (of direct syntactic arguments). In the case of (29b) the KEY is the caused change of state description. These MRS-style representations of the two alternants of *spray*, with different KEY values, are shown in (31) and (32).

(31)
$$\begin{bmatrix} \text{KEY} & \boxed{5} \\ \text{RELATIONS} & \left\langle \boxed{5}\begin{bmatrix} \textit{spray-ch-of-loc-rel} \\ \text{ACT} & \boxed{1} \\ \text{UND} & \boxed{4} \\ \text{SOA} & \begin{bmatrix} \textit{ch-of-loc-rel} \\ \text{FIG} & \boxed{4} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

(32)
$$\begin{bmatrix} \text{KEY} & \boxed{3}\begin{bmatrix} \textit{spray-ch-of-st-rel} \\ \text{ACT} & \boxed{1} \\ \text{UND} & \boxed{2} \\ \text{SOA} & \begin{bmatrix} \textit{ch-of-st-rel} \\ \text{UND} & \boxed{2} \end{bmatrix} \end{bmatrix} \\ \text{RELATIONS} & \left\langle \boxed{3}, \begin{bmatrix} \textit{use-rel} \\ \text{ACT} & \boxed{1} \\ \text{UND} & \boxed{4} \\ \text{SOA} & \boxed{3} \end{bmatrix}, \begin{bmatrix} \textit{spray-ch-of-loc-rel} \\ \text{ACT} & \boxed{1} \\ \text{UND} & \boxed{4} \\ \text{SOA} & \begin{bmatrix} \textit{ch-of-loc-rel} \\ \text{FIG} & \boxed{4} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

Generalizing from this example, one possible characterization of valence al-

ternations, implicit in Koenig & Davis (2006), is as systematic relations between two sets of lexical entries in which the RELS of any pair of related entries are in a subset/subset relation (a weaker version of that definition would merely require an overlap between the RELS values of the two entries). Consider another case; (33) illustrates the causative-inchoative alternation, where the intransitive alternant describes only the change of state, while the transitive one ascribes a explicit causing agent.

(33) a. John broke the window.
      b. The window broke.

Under a MRS representation, the change of state relation is a separate member of RELS; it is also included in the RELS of the transitive alternant, which contains a cause relation as well. Again, the RELS value of one member of each pair of related entries is a subset of the RELS value of the other.

Many other alternations linvolve one argument shifting from direct to oblique. Some English examples include conative, locative preposition drop, and *with* preposition drop alternations, as shown in (34):

(34) a. Rover clawed (at) Spot.
      b. Bill hiked (along) the Appalachian Trail.
      c. Burns debated (with) Smithers.

It is well known that the direct objects in these alternations seem to be "affected" more than their oblique counterparts. So if Rover clawed Spot, we infer that Spot was subjected to direct contact with Rover's claws and may have been injured by them, while if Rover merely clawed *at* Spot, no such inference can be made. Similarly, to say the one has hiked the Appalachian Trail suggests that one has hiked its entire length, not merely hiked along some portion of it. This holistic effect is not so evident in cases like (34c), though the direct object variant suggests that a formally organized debate took place, while the *with* variant could just describe an informal discussion. How might these varying intuitions related to "affectedness" relate to lexical semantic representations like those in (31) and (32)? Beavers (2010) provides one analytical advance in this direction, similar to the subset relationship between RELS values described above. He generalizes from affectedness to strength of entailments, where one semantic role's entailments are stronger than another's if and only if the set of entailments characterizing the second role also hold of the first. That is, what is true for any particpant that bears the first role will be true for any particpant that bears the

second, but not necessarily the converse. His *Morphosyntactic Alignment Principle* then relates this to linking, as stated in (35), where an "L-thematic role" is a linguistically relevant semantic role:

(35)   When participant *x* may be realized as either a direct or oblique argument of verb V, it bears L-thematic role *R* as a direct argument and L-thematic role $Q \subseteq_M R$ as an oblique. (Beavers 2010: 848)

Here, $Q \subseteq_M R$ means that Q is a "minimally weaker" role than *R*; in other words, there is no role *P* in the predicator such that $Q \subset P \subset R$. Thus, the substantive claim is essentially that the MAP rules out "verbs where the alternating participant has MORE lexical entailments as an oblique than the corresponding object realization" (Beavers 2010: 849).

The entailments Beavers employs differ somewhat from those we have discussed here, involving quantized change, nonquantized change, potential for change (where change can refer to change in location, possession, state, or something more abstract), furnishing the clear ordering by strength that is central to his proposal. But they do resemble entailments of semantic relations we have represented as elementary predications, such as incremental theme, change of state, and possession, along with the modal effects described in Koenig & Davis (2001). Thus the notion of a stronger role in Beavers' analysis has a rough analog in terms of whether a particular elementary predication is present in the semantics of a particular alternant. And only if an elementary predication is present, can it be designated as the KEY, and its roles linked directly. For example, in (31), there is nothing representing affectedness of the location, while in (32), there is, and it is designated as the KEY. As noted earlier, the semantics in (32) represents this additional entailment borne by the location argument. However, we are not aware of any simple, general way to represent Beavers' MAP within the EP-based model of Koenig & Davis (2006). Indeed, there is an aspect of Beavers' view that seems more in accord with numerical comparison approaches such as those of Dowty (1991) and Ackerman & Moore (2001), in that role strength is determined by the number of entailments that hold of it relative to others.
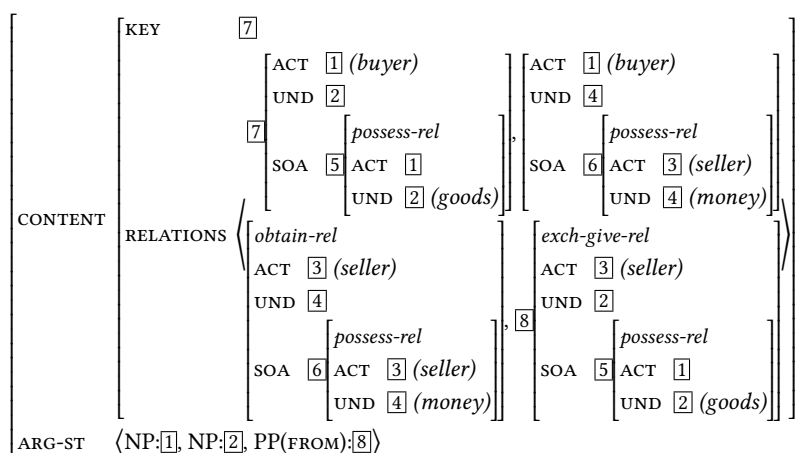
Having outlined the semantic basis of the different linking patterns of alternating verbs, we briefly take up three other issues. First is the question of how the alternants are related to one another. Second is how KEY selection has been used to account not just for alternants of the same verb, but for (nearly) synonymous verbs whose semantics contain the same set of elementary predications. Third is whether passives, which arguably do not differ semantically from their active counterparts, should be assimilated with other alternations or treated distinctly,

Draft of 9th September 2019, 12:08
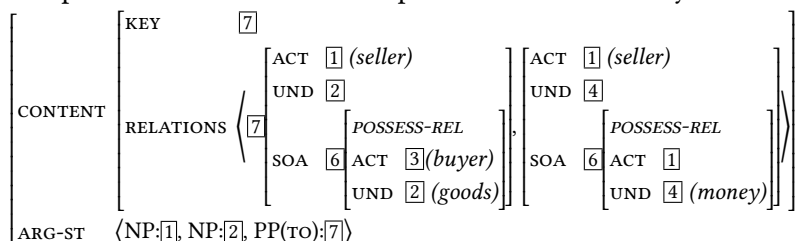
as a kind of non-canonical lexical item.

The hypothesis pursued in Davis (1996; 2001) is that most alternations are the consequence of classes of lexical entries having two related meanings. This follows researchers such as Pinker (1989) and Levin (1993) in modeling subcategorization alternations as underlyingly meaning alternations. This change in meaning is crucial in Koenig & Davis (2006) KEY shifts as well. In some cases, the value of the RELS attribute of the two valence alternates differ (as in the two alternates of *spray* in *spray/load* alternation we discussed earlier). In some cases, the alternation might be different construals of the same event for some verbs, but not others, as Rappaport Hovav & Levin (2008) claim for the English ditransitive alternations, which adds the meaning of transfer for verbs like *send*, but not for verbs like *promise*; a KEY change would be involved (with the addition of a *cause-possess-rel*) for the first verb only. But KEY shifts and diathesis alternations do not always involve a change in meaning. The same elementary predications can be present in as the CONTENT values of two alternants, with each alternant designating a different elementary predication as the KEY.

Koenig & Davis propose this not only for cases in which there is no obvious meaning difference between two alternants of a verb, but also for different verbs that appear to be truth-conditionally equivalent, one famous example being the verbs of commercial exchange *buy* and *sell* (but see Van Valin (1999): 387-388), Levin & Rappaport Hovav (2005: 20), and Wechsler (2005a) for arguments that *buy* and *sell* are not equivalent). Koenig and Davis argue that a commercial event involves two reciprocal actions, an exchange of goods (which involves giving goods and obtaining goods) and an exchange of money (which involves giving money and obtaining money). Individual verbs might select one or the other these four relations, thus accounting for the differences in subject and object selection. As shown in (36) and (37), each of these verbs contains four elementary predications: one *exch-give-rel* and one *obtain-rel* for the transfer of goods, and one of each for the counter-transfer of money. *Buy* designates the *obtain-rel* representing the transfer of goods as the KEY, while *sell* designates the *exch-give-rel* representing the transfer of goods as the KEY. Other verbs, such as *pay* or *charge*, choose elementary predications representing the counter-transfer as the KEY. In all cases, the same linking constraints apply between the KEY and the ARG-ST list, yielding the different argument realizations of these verbs while preserving their underlying semantic commonality. The relevant portions of the entries for *buy* and *sell* in (36) and (37) below illustrate: critically, the KEY relation for *buy* is not the same as that for *sell*.

(36)   A representation of the relevant parts of the lexical entry for *buy*:

$$
\left[
\begin{array}{l}
\text{CONTENT}
\left[
\begin{array}{ll}
\text{KEY} & \boxed{7} \\[2pt]
\text{RELATIONS} & \left\langle
\boxed{7}
\left[
\begin{array}{l}
\text{ACT} \ \boxed{1}\ (buyer) \\
\text{UND} \ \boxed{2} \\
\text{SOA} \ \boxed{5}
\left[
\begin{array}{l}
possess\text{-}rel \\
\text{ACT} \ \boxed{1} \\
\text{UND} \ \boxed{2}\ (goods)
\end{array}
\right]
\end{array}
\right],
\left[
\begin{array}{l}
\text{ACT} \ \boxed{1}\ (buyer) \\
\text{UND} \ \boxed{4} \\
\text{SOA} \ \boxed{6}
\left[
\begin{array}{l}
possess\text{-}rel \\
\text{ACT} \ \boxed{3}\ (seller) \\
\text{UND} \ \boxed{4}\ (money)
\end{array}
\right]
\end{array}
\right],
\right. \\[20pt]
\qquad\quad
\left[
\begin{array}{l}
obtain\text{-}rel \\
\text{ACT} \ \boxed{3}\ (seller) \\
\text{UND} \ \boxed{4} \\
\text{SOA} \ \boxed{6}
\left[
\begin{array}{l}
possess\text{-}rel \\
\text{ACT} \ \boxed{3}\ (seller) \\
\text{UND} \ \boxed{4}\ (money)
\end{array}
\right]
\end{array}
\right],
\boxed{8}
\left[
\begin{array}{l}
exch\text{-}give\text{-}rel \\
\text{ACT} \ \boxed{3}\ (seller) \\
\text{UND} \ \boxed{2} \\
\text{SOA} \ \boxed{5}
\left[
\begin{array}{l}
possess\text{-}rel \\
\text{ACT} \ \boxed{1} \\
\text{UND} \ \boxed{2}\ (goods)
\end{array}
\right]
\end{array}
\right]
\left.
\right\rangle
\end{array}
\right] \\[40pt]
\text{ARG-ST} \quad \langle \text{NP:}\boxed{1}, \text{NP:}\boxed{2}, \text{PP(FROM):}\boxed{8} \rangle
\end{array}
\right]
$$

(37) A representation of the relevant parts of the lexical entry for *sell*:

$$
\left[
\begin{array}{l}
\text{CONTENT}
\left[
\begin{array}{ll}
\text{KEY} & \boxed{7} \\[2pt]
\text{RELATIONS} & \left\langle
\boxed{7}
\left[
\begin{array}{l}
\text{ACT} \ \boxed{1}\ (seller) \\
\text{UND} \ \boxed{2} \\
\text{SOA} \ \boxed{6}
\left[
\begin{array}{l}
\text{POSSESS-REL} \\
\text{ACT} \ \boxed{3}\ (buyer) \\
\text{UND} \ \boxed{2}\ (goods)
\end{array}
\right]
\end{array}
\right],
\left[
\begin{array}{l}
\text{ACT} \ \boxed{1}\ (seller) \\
\text{UND} \ \boxed{4} \\
\text{SOA} \ \boxed{6}
\left[
\begin{array}{l}
\text{POSSESS-REL} \\
\text{ACT} \ \boxed{1} \\
\text{UND} \ \boxed{4}\ (money)
\end{array}
\right]
\end{array}
\right]
\right\rangle
\end{array}
\right] \\[30pt]
\text{ARG-ST} \quad \langle \text{NP:}\boxed{1}, \text{NP:}\boxed{2}, \text{PP(TO):}\boxed{7} \rangle
\end{array}
\right]
$$

As a final example of semantic alternations in the fine-grained meaning possibilities of verbs, we consider here the source-final product alternation exemplified in (38) where the direct object can be either the final product or the material source of the final product. Davis proposes that the (38a) sentences involve an alternation between the two meanings of entries represented in (39). We adapt Davis (2001) to make it consistent with Koenig & Davis (2006) and also treat the alternation as an alternation of meaning of *entries*. Note that in the meaning alternation described in (39), we use, informally, a double-headed arrow. One of the potential drawbacks of a lexical rule approach to valence alternations is that it requires selecting one or the other alternant as basic and the other as derived (e.g., is the inchoative or the causative basic?). This is not always an easy decision, as Goldberg (1995) or Levin & Rappaport Hovav (1994) have pointed out. Sometimes, morphology provides a clue, although in different languages the clues may point in different directions. French, and other Romance languages, use a "reflexive" clitic as a detransitivizing affix. In English, though, there is no obvious "basic" form or directionality. It is to avoid committing ourselves to a directionality in the meaning relation described in (39) that we eschews treating

it as a lexical rule.

(38)   a. Kim made/carved/sculpted/crafted a toy (out of the wood).

    b. Kim made/carved/sculpted/crafted the wood into a toy.

(39)
$$\begin{bmatrix} \text{CONTENT} & \begin{bmatrix} \text{KEY} & \boxed{3}\begin{bmatrix} \textit{affect-incr-th-rel} \\ \text{ACT} & \boxed{1} \\ \text{SOA} & \boxed{2}\textit{(final product)} \end{bmatrix} \\ \text{RELS} & \langle \boxed{3}, \boxed{4} \rangle \end{bmatrix} \end{bmatrix}$$

$$\longleftrightarrow$$

$$\begin{bmatrix} \text{CONTENT} & \begin{bmatrix} \text{KEY} & \boxed{4}\begin{bmatrix} \textit{affect-incr-th-rel} \\ \text{ACT} & \boxed{1} \\ \text{UND} & \textit{(source material)} \\ \text{SOA} & \begin{bmatrix} \textit{affect-incr-th-rel} \\ \text{ACT} & \boxed{1} \\ \text{SOA} & \boxed{2}\textit{(final product)} \end{bmatrix} \end{bmatrix} \\ \text{RELS} & \langle \boxed{3}, \boxed{4} \rangle \end{bmatrix} \end{bmatrix}$$

Although most diathesis alternations can be modeled as alternations in meaning and as KEY shifts, some arguably cannot. We discuss the active/passive alternation here, but impersonals, as well as raising structures exemplified in (28) are good candidates too. The semantic relations of actives and long passives, as in (40), are practically identical and the difference between the two alternates is pragmatic in nature. Arguably, then, passives are a degenerate case of the subset relationship between RELS attributes, where the RELS values of the two entries are identical and so are the two entries' KEY. But this raises the question of whether linking in passives violates the constraints in (18)–(20), especially (18, which links the value of ACT to the first element of ARG-ST.

(40)   a. Fido dug a couple of holes.

    b. A couple of holes were dug by Fido.

One typical HPSG method for modeling valence alternations like passives is through lexical rules (see Chapter ??) with one alternant serving as input and the other as output; the main effect of the lexical rule in such an approach is to alter the ARG-ST of the input, going from the ARG-ST of *give* to that of *given* in (8). Critically, we must assume that the output cannot be subject to linking constraint (18), since the actor argument is not linked to the first member of the ARG-ST list. A simplified representation of what such a rule would look like is provided in (41) where we assume that the input to the rule must be transitive.

(41) $\begin{bmatrix} \textit{trans-vb} \\ \text{HEAD} \quad \textit{verb} \\ \text{ARG-ST} \ \left\langle \text{NP}_{\boxed{1}}, \text{NP}_{\boxed{2}} \right\rangle \oplus \boxed{3}\textit{list} \end{bmatrix} \mapsto \begin{bmatrix} \text{HEAD} \quad \begin{bmatrix} \text{VFORM} \ \textit{pass} \end{bmatrix} \\ \text{ARG-ST} \ \left\langle \text{NP}_{\boxed{2}} \right\rangle \oplus \boxed{3} \oplus \left\langle \text{PP}[_{by}]_{\boxed{1}} \right\rangle \end{bmatrix}$

To sum up, in contrast to most meaning-driven alternations, valence alternations like the active/passive are modeled through the use of lexical rules that alter the ARG-ST of "base" entries. Which alternations pattern with active/passive and require positing lexical rules that alter a "base" entry's ARG-ST list is, as of yet, not settled. We turn now to roles that are putatively present semantically, but not realized syntactically at all.

6 Extended ARG-ST

Most of this chapter focuses on cases where semantic roles linked to the ARG-ST list are arguments of the verb's core meaning. But in quite a few cases, complements (or even subjects) of a verb are not part of this basic meaning; consequently, the ARG-ST list must be extended to include elements beyond the basic meaning. We consider three cases here, illustrated in (42)–(44).

Resultatives, illustrated in (42), express an effect, which is caused by an action of the type denoted by the basic meaning of the verb. The verb *fischen* 'to fish' is a simple intransitive verb (42a) that does not entail that any fish were caught, or any other specific effect of the fishing.

(42) a. dass er fischt
         that he fishes
         'that he is fishing'

      b. dass er ihn leer   fischt
         that he it   empty fishes
         'that he is fishing it empty'

      c. wegen   der Leerfischung der   Nordsee
         because.of the empty.fishing of.the North.See.GEN
         'because of the North Sea being fished empty'

In (42b) we see a resultative construction, with an object NP and a secondary predicate AP. The meaning is that he is fishing, causing it (the body of water) to become empty of fish. Müller (2002) posits a German lexical rule applying to the verb that augments the ARG-ST list with an NP and AP, and adds the causal semantics to the CONTENT (see Wechsler (2005b) for a similar analysis of English

resultatives). The existence of deverbal nouns like *Leerfischung* 'fishing empty', which takes the body of water as an argument in genitive case (see 42c) confirms that the addition of the object is a lexical process, as noted by Müller (2002).

Romance clause-union structures as in (43) have long been analyzed as cases where the complements of the complements of a clause-union verb (*faire* in (43)) are complements of the clause-union verb itself (Aissen 1979).

(43) Johanna a   fait   manger les enfants.
      Johanna have made eat   the children
      'Johanna had the children eat.'

Within HPSG, the "union" of the two verbs' dependents is modeled via the composition of ARG-ST lists of the clause union verb, following Hinrichs & Nakazawa (1994) (this is a slight simplification, see Chapter ?? for details).

Sentence (44) illustrates a slightly different point, namely that some semantic modifiers, such as *souvent* in (44), can be realized as complements, and thus should be added as members of ARG-ST (or members of the DEPS list if one countenances such an additional list).

(44) Mes amis   m' ont   souvent aidé.
      My friends me have often   helped
      'My friends often helped me.'

Abeillé & Godard (1997) have argued that many adverbs including negative adverbs and negation in French are complements of the verb and Kim & Sag (2002) extended that view to some uses of negation in English. In contrast to resultatives, which affect the meaning of the verb, or to clause union, where one verb co-opts the argument structure of another verb, what is added to the ARG-ST list in these cases is typically considered a semantic adjunct and a modifier in HPSG (thus it selects the verb or VP via the MOD attribute).

Another case of an adjunct that behaves like a complement is found in (45). The clitic *en* expressing the cause of death is not normally an argument of the verb *mourir* 'die', but rather an adjunct (Koenig & Davis 2006):

(45) Il en   est mort
      He of.it is   dead.PERF.PAST
      'He died of it' (Koenig & Davis 2006, ex. 12a)

On the widespread assumption (at least within HPSG) that pronominal clitics are verbal affixes (see Miller & Sag 1997), the adjunct cause of the verb *mourir* must be

represented within the entry for *mourir*, so as to trigger affixation by *en*. Bouma, Malouf & Sag (2001) discuss such cases and other cases where "adverbials" as they call them, can be part of a verb's lexical entry. To avoid mixing those adverbials with the argument structure list (and have to address their relative obliqueness with syntactic arguments of verbs), they introduce yet an additional list, the dependents list (abbreviated as DEPS) which includes the ARG-ST list but also a list of adverbials. Each adverbial selects for the verb on whose DEPS list it appears as argument, as shown in (46). But, of course, not all verb modifiers can be part of the DEPS list and Bouma, Malouf & Sag discuss at length some of the differences between the two kinds of "adverbials."

$$(46) \quad verb \Rightarrow \begin{bmatrix} \text{CONT|KEY} & \boxed{2} \\ \text{HEAD} & \boxed{3} \\ \text{DEPS} & \boxed{1} \oplus list\left(\begin{bmatrix} \text{MOD} & \begin{bmatrix} \text{HEAD} & \boxed{3} \\ \text{key} & \boxed{2} \end{bmatrix} \end{bmatrix} \right) \\ \text{ARG-ST} & \boxed{1} \end{bmatrix}$$

Although the three cases we have outlined result in an extended ARG-ST, the ways in which this extension arises differ. In the case of resultatives, the extension results partly or wholly from changing the meaning in a way similar to Rappaport Hovav & Levin (1998): by adding a causal relation, the effect argument of this causal relation is added to the membership in the base ARG-ST list (see Section 5 for a definition of the attributes KEY and RELS; suffice it to say for now that a *cause-rel* is added to the list of relations that are the input of the rule). The entries of the clause union verbs are simply stipulated to include on their ARG-ST lists the syntatic arguments of their (lexical) verbal arguments. Finally, (negative) adverbs that select for a verb (VP) are added to the ARG-ST of the verb they select. A simplified representation of all three processes is provided in (47)-(49).

$$(47) \quad \begin{bmatrix} \text{KEY} & \boxed{2} \\ \text{RELS} & \boxed{1}\langle \dots \boxed{2} \dots \rangle \end{bmatrix} \mapsto \begin{bmatrix} \text{KEY} & \boxed{3} \textit{cause-rel} \\ \text{RELS} & \boxed{1} \oplus \boxed{3} \end{bmatrix}$$

$$(48) \quad \begin{bmatrix} \text{ARG-ST} & \left\langle \dots, \begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{ARG-ST} & \boxed{1} \end{bmatrix} \right\rangle \bigcirc \boxed{1} \end{bmatrix}$$

$$(49) \quad \begin{bmatrix} \text{ARG-ST} & \boxed{1} \end{bmatrix} \mapsto \begin{bmatrix} \text{ARG-ST} & \boxed{1} \bigcirc \langle \text{ADV}_{neg} \rangle \end{bmatrix}$$

7 Is ARG-ST universal?

In this section, we briefly consider the question of whether something akin to Figure 1 offers a satisfying account of the grammatical encoding of semantic ar-

guments across all languages. Because of its role in accounting for the syntax of basic clauses, the presence of an ARG-ST list on lexical entries comes with expectations about the syntactic realization of semantic roles. In recent work, Koenig & Michelson (2015) argue these expectations are not universally borne out, based on data from Oneida (Norther Iroquoian). The only grammatical reflex of semantic arguments in Oneida, they argue, is inflectional: the referencing of semantic arguments by so-called pronominal prefixes, which are better thought of as agreement markers à la Evans (2002). Koenig and Michelson distinguish between grammatical and syntactic arguments. Grammatical arguments include not only syntactic arguments (that is, those on ARG-ST) but also inflectional referencing of semantic roles. Some ordering analogous to linking in other languages is present in Oneida, because the semantic roles are not arbitrarily associated with agreement morphemes, but this can be captured in an ordered list of semantic indices , called INFL-STR in this proposal. INFL-STR is part of the morphological information relevant to word-internal inflectional processes, what Anderson (1992) calls 'Morphosyntactic Representation' in his treatment of Georgian agreement markers. The ordering of semantic indices on INFL-STR insures that the predicator is properly inflected. For example, the prefix *lak-* occurs if a third singular masculine proto-agent argument is acting on a first singular proto-patient argument as in *lak-hlo·lí-heʔ* 'he tells me' (habitual aspect), whereas the prefix *li-* occurs if a first singular proto-agent argument is acting on a third masculine singular argument, as in *li-hlo·lí-heʔ* 'I tell him' (habitual aspect).

In (50) and (51) we show the distinction between grammatical arguments realized as syntactic arguments (as in most languages) and those that are not (as in Oneida). Here, Koenig and Michelson follow the encoding of linking constraints as implicational constraints, as in Koenig & Davis (2003), although nothing critical hinges on that choice.

$$(50) \quad \begin{bmatrix} \text{CONTENT} & \begin{bmatrix} \textit{cause-rel} \\ \text{CAUSER} & \boxed{1} \end{bmatrix} \\ \text{ARG-ST} & \langle \text{NP}, ... \rangle \end{bmatrix} \implies \begin{bmatrix} \text{ARG-ST} & \langle \text{NP:}\boxed{1}, ... \rangle \end{bmatrix}$$

$$(51) \quad \begin{bmatrix} \text{CONTENT} & \begin{bmatrix} \textit{cause-rel} \\ \text{CAUSER} & \boxed{1}\textit{anim} \end{bmatrix} \end{bmatrix} \implies \begin{bmatrix} \text{INFL-STR} & \langle \boxed{1}, ... \rangle \end{bmatrix}$$

(50) constrains the association between a cause and a *synsem* member of the ARG-ST list; (51) constrains the semantic *index* of the cause to be the first member of INFL-STR.

If a language like Oneida (Northern Iroquoian) only includes an ordering of semantic indices for inflectional purposes and constraints such as (51) and no

ARG-ST list, a number of predictions follow, which Koenig and Michelson claim are borne out. Briefly summarizing their evidence, the relation between semantic arguments and external phrases, when they occur, is not necessarily one of co-indexing, no binding constraints exist between external phrases (e.g., condition C violations can be found), there are no valence alternations, and no syntactic constraints on extraction. In other words, Oneida contains none of the evidence supporting the presence of an ARG-ST list and an ordering of syntactic arguments along an obliqueness hierarchy we have discussed in this chapter. The ARG-ST list may thus not a universal attribute of words, though present in the overwhelming majority of languages. Linking, understood as constraints between semantic roles and members of the ARG-ST list, is then but one possibility; constraints that relate semantic roles to an INFL-STR list of semantic indices is also an option. In languages that exclusively exploit that latter possibility, syntax is indeed simpler.

8 The lexical approach to argument structure

We end this chapter with a necessarily brief comparison between the approach to argument structure we describe in this chapter with other approaches to argument structure that have developed since the 1990's. This chapter describes a *lexical approach to argument structure*, which is typical of research in HPSG. The basic tenet of such approaches is that lexical items include argument structures, which represent essential information about potential argument selection and expression, but abstract away from the actual local phrasal structure. In contrast, *phrasal approaches*, which are common both in Construction Grammar and in transformational approaches such as Distributed Morphology, reject such lexical argument structures. Let us briefly review the reasons for preferring a lexical approach. (This section is drawn from Müller & Wechsler (2014b), which may be consulted for more detailed and extensive argumentation).

In phrasal approaches to argument structure, components of a verb's apparent meaning are actually 'constructional meaning' contributed directly by the phrasal structure. The linking constraints of the sort discussed above are then said to arise from the interaction of the verb meaning with the constructional meaning. For example, agentive arguments tend to be realized as subjects, not objects, of transitive verbs. On the theory presented above, that generalization is captured by the linking constraint (18), which states that the ACTOR argument of an *act-und-rel* ('actor-undergoer relation') is mapped to the initial item in the ARG-ST list. In a phrasal approach, the agentive semantics is directly associated with the subject position in the phrase structure. In transformational theories, a

silent 'light verb' (usually called 'little *v*') heads a projection in the phrase structure and assigns the agent role to its specifier (the subject). In constructional theories, the phrase structure itself assigns the agent role. In either type of phrasal approach, the agentive component of the verb meaning is actually expressed by the phrasal structure into which the verb is inserted.

The lexicalist's predicate argument structure provides essential information for a verb's potential combination with argument phrases. If a given lexical entry could only combine with the particular set of phrases specified in a single vALfeature, then the lexical and phrasal approaches would be difficult to distinguish: whatever information the lexicalist specifies for each vALlist item could, on the phrasal view, be specified instead for the phrases realizing those list items. But crucially, the verb need not immediately combine with its specified arguments. Alternatively it can meet other fates: it can serve as the input to a lexical rule; it can combine first with a modifier in an adjunction structure; it can be coordinated with another word with the same predicate argument 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). Here we consider two of these, lexical rules and coordination.

The predicate argument structure is 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.

(52)   a. The rabbits were nibbling the carrots.

   b. The carrots were being nibbled (by the rabbits).

   c. a large, partly nibbled, orange carrot

   d. the quiet, nibbling, old rabbits

   e. the rabbit's nibbling of the carrots

   f. The rabbit gave the carrot a nibble.

   g. The rabbit wants a nibble (on the carrot).

   h. The rabbit nibbled the carrot smooth.

Verbs undergo morpholexical operations like passive (52d), as well as antipassive, causative, and applicative in other languages. They have cognates in other parts of speech such as adjectives (52c,d) and nouns (52e,f,g). Verbs have been argued to form complex predicates with resultative secondary predicates (52h), and with serial verbs in other languages.

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 syntax 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 (52) 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 and the referential index of Arg1 is optionally assigned to a *by*-phrase. The same rules of syntax dictate the position of the subject, whether the verb is active or passive. 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 referent of the noun that it modifies, see Bresnan (1982) and Bresnan et al. (2015: Chapter 3). The phrasal approach, in which the agent role is assigned to the subject position, is too rigid.

Nor could this be solved by associating each syntactic environment with a different meaningful phrasal construction: an active construction with agent role in the subject position; a passive construction with agent in the *by*-phrase position; etc. The problem for that view is that that one lexical rule can feed another. In the example above, the output of the verbal passive rule (see (52d)) feeds the adjective formation rule (see (52e)).

A verb can also be coordinated with another verb with the same valence requirements. The two verbs then share their dependents. This causes problems for the phrasal view, especially when a given dependent receives different semantic roles from the two verbs. For example, in an influential phrasal analysis, Hale & Keyser (1993) derived denominal verbs like *to saddle* through noun incorporation out of a structure akin to [PUT a saddle ON x]. Verbs with this putative derivation routinely coordinate and share dependents with verbs of other types:

(53)   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.[4]

Under the phrasal analysis the two verbs place contradictory demands on a single phrase structure. But on the lexical analysis, this is simple V^0 coordination.

To summarize, a lexical argument structure is an abstraction or generalization over various occurrences of the verb in syntactic contexts. To be sure, one key use of that argument 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

---

[4]http://www.jouetthouse.org/index.php?option=com_content&view=article&id=56&Itemid=63, 21.07.2012

Draft of 9th September 2019, 12:08

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 phrasal approach prematurely commits to a single phrasal position for the realization of a semantic argument. In contrast, a lexical argument structure gives a word the appropriate flexibility to account for the full range of expressions found in natural language.

References

Abeillé, Anne & Danièle Godard. 1997. The syntax of French negative adverbs. In Paul Hirschbühler & F. Marineau (eds.), *Negation and polarity: syntax and semantics*, 1–17. Amsterdam: John Benjamins.

Ackerman, Farrell, Robert Malouf & John Moore. 2017. Symmetrical objects in Moro: challenges and solutions. *Journal of Linguistics* 53. 3–50.

Ackerman, Farrell, Robert Malouf & John Moore. 2018. Symmetric objects in Moro. http://csli-publications.stanford.edu/HPSG/2013/, accessed 2018-7-5. Paper presentend at the 20th International Conference on Head-Driven Phrase Structure Grammar, Freie Universität Berlin.

Ackerman, Farrell & John Moore. 2001. *Proto-properties and grammatical encoding: a correspondence theory of argument selection*. Stanford: CSLI Publications.

Aissen, Judith. 1979. *The syntax of causative constructions*. New York: Garland.

Anderson, Stephen. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Artawa, Ketut. 1994. *Ergativity and Balinese syntax*. Bundoora, Australia: La Trobe University dissertation.

Baker, Mark. 1988. *Incorporation*. Chicago: Chicago University Press.

Baker, Mark. 1997. Thematic roles and syntactic structures. In Liliane Haegeman (ed.), *Elements of grammar*, 73–137. Kluwer.

Beavers, John. 2010. The structure of lexical meaning: why semantics really matters. *Language* 86. 821–864.

Bouma, Gosse, Robert Malouf & Ivan Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory* 19(1). 1–65.

Bresnan, Joan. 1982. The passive in lexical theory. In Joan Bresnan (ed.), *The mental representation of grammatical relations* (MIT Press Series on Cognitive Theory and Mental Representation), 3–86. Cambridge, MA/London: MIT Press.

Bresnan, Joan, Ash Asudeh, Ida Toivonen & Stephen Mark Wechsler. 2015. *Lexical-functional syntax*. 2nd edn. Oxford: Blackwell Publishers Ltd. DOI:10.1002/9781119105664

Carlson, Greg. 1998. Thematic roles and the individuation of events. In Susan Rothstein (ed.), *Events and grammar*, 35–51. Dordrecht: Kluwer.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris Publications.

Copestake, Ann, Dan Flickinger, Carl Pollard & Ivan Sag. 2005. Minimal Recursion Semantics: An Introduction. *Research on Language and Computation* 3. 281–332. DOI:10.1007/s11168-006-6327-9

Copestake, Ann, Alex Lascarides & Dan Flickinger. 2001. An algebra for semantic construction in constraint-based grammars. In *Proceedings of the 39th ACL meeting*. Toulouse, France.

Davis, Anthony. 1996. *Lexical semantics and linking in the hierarchical lexicon*. Stanford University dissertation.

Davis, Anthony. 2001. *Linking by types in the hierarchical lexicon*. Stanford: CSLI Publications.

Davis, Anthony & Jean-Pierre Koenig. 2000. Linking as constraints on word classes in a hierarchical lexicon. *Language* 76(1). 56–91.

Dowty, David. 1989. On the semantic content of the notion of 'thematic role'. In Gennaro Chierchia, Barbara Partee & Raymond Turner (eds.), *Properties, types, and meaning*, vol. 2, 69–129. Dordrecht: Kluwer.

Dowty, David R. 1979. *Word meaning and montague grammar: the semantics of verbs and times in generative semantics and in Montague's PTQ*. Vol. 7 (Studies in Linguistics and Philosophy). Dordrecht, Netherlands: Reidel. DOI:10.1007/978-94-009-9473-7

Dowty, David R. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619.

Evans, Nicholas. 2002. The true status of grammatical object affixes: evidence from Bininj Gun-Wok. In Nicholas Evans & Hans-Jürgen Sasse (eds.), *Problems of polysynthesis*, 15–50. Berlin: Akademie Verlag.

Fillmore, Charles. 1968. The case for case. In Emmon Bach & Robert Harms (eds.), *Universals in linguistic theory*, 1–87. New York: Holt, Rinehart & Winston.

Fillmore, Charles. 1977. The case for case reopened. In Peter Cole & Jerold Sadock (eds.), *Grammatical relations*, 59–81. New York: Academic Press.

Foley, William & Robert Van Valin. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.

Gawron, Jean Mark. 1986. Situations and prepositions. *Linguistics and Philosophy* 9. 327–382.

Draft of 9th September 2019, 12:08

Gerdts, Donna. 1992. Morphologically-mediated relational profiles. In Laura Buszard-Welcher et al. (ed.), *Proceedings of the 18th annual meeting of the berkeley linguistics society*, 322–337. Berkeley, CA: Berkeley Linguistics Society.

Goldberg, Adele. 1995. *Constructions: a construction grammar approach to argument structure*. Chicago: Chicago University Press.

Greenberg, Joseph H. 1966. *Language universals*. The Netherlands: Mouton.

Guilfoyle, Eithne, Henrietta Hung & Lisa Travis. 1992. Spec of IP and Spec of VP: two subjects in Austronesian languages. *Natural Language and Linguistic Theory* 10(3). 375–414.

Hale, Kenneth & Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20*, 53–109. Cambridge, MA: MIT Press.

Hinrichs, Erhard & Tsuneko Nakazawa. 1994. Linearizing AUXs in German verbal complexes. In John Nerbonne, Klaus Netter & Carl Pollard (eds.), *German in Head-driven Phrase-Structure Grammar*, 11–37. Stanford: CSLI Publications.

Jackendoff, Ray. 1987. The status of thematic relations in linguistic theory. *Linguistic Inquiry* 18. 369–411.

Keenan, Edward & Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8. 63–99.

Keyser, Samuel & T. Roeper. 1984. On the middle and ergative constructions in English. *Linguistic Inquiry*. 2nd ser. 15. 381–416.

Kim, Jong-Bok & Ivan Sag. 2002. Negation without head movement. *Natural Language and Linguistic Theory* 20(2). 339–412.

Koenig, Jean-Pierre & Anthony Davis. 2001. Sublexical modality and the structure of lexical semantic representations. *Linguistics and Philosophy* 24. 71–124.

Koenig, Jean-Pierre & Anthony Davis. 2003. Semantically transparent linking in HPSG. In Stefan Müller (ed.), *Proceedings of the HPSG 2003 conference*, 222–235. Stanford: CSLI Publications.

Koenig, Jean-Pierre & Anthony Davis. 2006. The key to lexical semantic representations. *Journal of Linguistics* 42. 71–108.

Koenig, Jean-Pierre & Karin Michelson. 2015. Invariance in argument realization: the case of Iroquoian. *Language* 91. 1–47.

Ladusaw, William & David Dowty. 1988. Toward a nongrammatical account of thematic roles. In Wendy Wilkins (ed.), *Syntax and semantics 21: thematic relations*, 62–74. New York: Academic Press.

Landman, Fred. 2000. *Events and plurality: The Jerusalem lectures*. Dordrecht: Kluwer.

Levin, Beth. 1993. *English verb classes and alternations.* Chicago, IL: Chicago University Press.

Levin, Beth & Malka Rappaport Hovav. 1994. A preliminary analysis of causative verbs in English. In Lila Gleitman & Barbara Landau (eds.), *The acquisition of the lexicon*, 35–77. Cambridge, MA: MIT Press.

Levin, Beth & Malka Rappaport Hovav. 2005. *Argument realization.* Cambridge, UK: Cambridge University Press.

Manning, Christopher. 1996. *Ergativity: argument structure and grammatical relations.* Stanford, CA: CSLI publications.

Manning, Christopher & Ivan A Sag. 1999. Dissociations between argument structure and grammatical relations. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional aspects of linguistic explanation*, 63–78. Stanford: CSLI Publications.

Mauner, Gail & Jean-Pierre Koenig. 2000. Linguistic vs. conceptual sources of implicit agents in sentence comprehension. *Journal of Memory and Language* 43. 110–134.

Miller, Philip & Ivan Sag. 1997. French clitic movement without clitics or movement. *Natural Language and Linguistic Theory* 15(3). 573–639.

Müller, Stefan. 2002. *Complex predicates: Verbal complexes, resultative constructions, and particle verbs in German* (Studies in Constraint-Based Lexicalism 13). Stanford, CA: CSLI Publications.

Müller, Stefan & Stephen Mark Wechsler. 2014a. Lexical approaches to argument structure. *Theoretical Linguistics* 40(1–2). 1–76.   DOI:10.1515/tl-2014-0001

Müller, Stefan & Stephen Mark Wechsler. 2014b. Lexical approaches to argument structure. *Theoretical Linguistics* 40(1–2). 1–76.   DOI:10.1515/tl-2014-0001

Pinker, Steven. 1989. *Learnability and cognition: the acquisition of argument structure.* Cambridge, MA: MIT Press.

Pollard, Carl J. & Ivan A. Sag. 1987. *Information-based syntax and semantics* (CSLI Lecture Notes 13). Stanford, CA: CSLI Publications.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.

Rappaport Hovav, Malka & Beth Levin. 1998. Building verb meanings. In Miriam Butt & Wilhelm Geuder (eds.), *The projection of arguments*, 97–134. Stanford: CSLI Publications.

Rappaport Hovav, Malka & Beth Levin. 2008. The English dative alternation: the case for verb sensitivity. *Journal of Linguistics* 44. 129–167.

Reis, Marga. 1982. Zum Subjektbegriff im Deutschen. In Werner Abraham (ed.), *Satzglieder im Deutschen – Vorschläge zur syntaktischen, semantischen und prag-*

matischen Fundierung (Studien zur deutschen Grammatik 15), 171–211. Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag.

Sag, Ivan A., Thomas Wasow & Emily M. Bender. 2003. *Syntactic theory: A formal introduction.* 2nd edn. (CSLI Lecture Notes 152). Stanford, CA: CSLI Publications.

Schein, Barry. 2002. Events and the semantic content of thematic relations. In Gerhard Preyer & Georg Peter (eds.), *Logical form and language*, 263–344. Oxford: Oxford University Press.

Van Valin, Robert. 1999. Generalized semantic roles and the syntax-semantics interface. In *Empirical issues in formal syntax and semantics 2*, 373–388. The Hague, The Netherlands: Thesus.

Wechsler, Stephen. 1995a. Preposition selection outside the lexicon. In Raul Aranovich, William Byrne, Susanne Preuss & Matha Senturia (eds.), *Proceedings of WCCFL XIII*, 416–431. Stanford: CSLI Publications.

Wechsler, Stephen. 1995b. *The semantic basis of argument structure.* Stanford, CA: CSLI Publications.

Wechsler, Stephen. 1999. HPSG, GB, and the Balinese bind. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and constructional aspects of linguistic explanation*, 179–195. Stanford, CA: CSLI publications.

Wechsler, Stephen. 2005a. What is right and wrong about little v. In Mila Vulchanova & Tor A. Åfarli (eds.), *Grammar and beyond – essays in honour of Lars Hellan.* Novus Press.

Wechsler, Stephen Mark. 2005b. Resultatives under the event-argument homomorphism model of telicity. In Nomi Erteschik-Shir & Tova Rapoport (eds.), *The syntax of aspect—Deriving thematic and aspectual interpretation*, 255–273. Oxford University Press.

Wechsler, Stephen & I Wayan Arka. 1998. Syntactic ergativity in Balinese: an argument structure based theory. *Natural Language and Linguistic Theory* 16. 387–442.