For LING1111 class Spring 2009 – April 28 **Lars Hellan**

A basic introduction to using attribute-value matrices (AVMs) in linguistic representation.

The following note takes as its point of departure
i) an elementary view of classical transformational analyses of a configurational language like Norwegian;
ii) an elementary view of situation structure as constituted by ‘actants’ (participants’);
iii) an elementary conception of ‘grammatical functions’ such as ‘subject of’, ‘object of’, etc.

The goal of the note is to give a typologically broad basis for an appreciation of the conciseness offered by AVMs as a representational format.
Section 1 explains the basicfeature inventories and rules of composition of AVMs, according to one among current definitions of such structures (used in HPSG, and in the computational system LKB), with examples centering around verb constructions.
Section 2 applies the AVM formalism to representing verbal derivation involving morphological Causatives and Passive.
Section 3 states AVM representations of Equi and Raising constructions.
Section 4 gives AVM profiles for a wider range of construction types such as Applicatives and Ergative constructions.
Section 5 gives a brief introduction to how parsing grammars can be constructed relative to the construction representations reviewed, thereby indicating some basic algorithms of HPSG grammars and LFG grammars.
Appendix 1 gives an introduction to a system of type labels for a typologically rich inventory of construction types, related to their AVM representations.

1. Basic composition of AVMs

As an example, the sign structure for a ditransitive construction can be viewed first as follows: here are the GFs (grammatical functions) subject, direct object and indirect object, and the referents of these GFs exposed inside ACTNTS, which is an overview of the participants in the situation expressed by the construction:

(1)
$$\begin{bmatrix} \text{GF} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix}\text{INDX } \boxed{1}\end{bmatrix} \\ \text{OBJ} & \begin{bmatrix}\text{INDX } \boxed{2}\end{bmatrix} \\ \text{IOBJ} & \begin{bmatrix}\text{INDX } \boxed{3}\end{bmatrix} \end{bmatrix} \\ \text{ACTNTS} & \begin{bmatrix} \text{ACT1 } \boxed{1} \\ \text{ACT2 } \boxed{2} \\ \text{ACT3 } \boxed{3} \end{bmatrix} \end{bmatrix}$$

The paths ‘$\text{SUBJ} \begin{bmatrix}\text{INDX } \boxed{1}\end{bmatrix}$’ and ‘$\text{ACT1 } \boxed{1}$’ both lead to the same entity, or index – $\boxed{1}$; such paired use of boxed numbers is called ‘reentrancy’ or ‘identity’, and the boxed numbers themselves we may call ‘identifiers’). (Instead of boxes, we can use ‘#’, so that the above pair comes out as ‘$\text{SUBJ} \begin{bmatrix}\text{INDX \#1}\end{bmatrix}$’ and ‘ACT1 #1’. For the representation of paths, equivalent to ‘$\text{SUBJ}\left[\text{INDX}\ \#1\right]$’ we can use ‘SUBJ|INDX #1’ or ‘SUBJ.INDX #1’)

Then there are more things to expose: syntactically, the POS category of the head of the construction (which is ‘verb’), and semantically, the situation-index of the situation, and its Aktionsart. Most straightforward will be to expose them as follows:

(2)
$$\begin{bmatrix} \text{HEAD verb} \\ \text{GF} \begin{bmatrix} \text{SUBJ} \begin{bmatrix}\text{INDX } \boxed{1}\end{bmatrix} \\ \text{OBJ} \begin{bmatrix}\text{INDX } \boxed{2}\end{bmatrix} \\ \text{IOBJ} \begin{bmatrix}\text{INDX } \boxed{3}\end{bmatrix} \end{bmatrix} \\ \text{INDX index} \\ \text{ACTNTS} \begin{bmatrix} \text{ACT1 } \boxed{1} \\ \text{ACT2 } \boxed{2} \\ \text{ACT3 } \boxed{3} \end{bmatrix} \\ \text{AKTART aktionsart} \end{bmatrix}$$

Notice that the attribute INDX here occurs in two environments: as an ‘outermost’ attribute, and on the path inside of SUBJ. Can attributes occur in as many environments as we please? Every formalism has rules regulating their use. In our case, where we follow rules laid out in Copestake (2002), the rule can be seen by the following:

The attribute SUBJ introduces a *sign*: that is reasonable, since subjects are noun phrases, full-fledged signs. What (1) and (2) expose are also *sign* structures, since *constructions* are just complex signs. The *sign* status can be indicated by *types*, which stand for the kinds of objects that are characterizable by the attributes in question. (Attributes are used to sort things, and what are appropriate sorting attributes depends on what kind of thing one is sorting: fish markets require certain attributes, universities other, and linguistic signs still other.) If we enter type next to the bracket comprising the attributes, (2) comes out as (3) (type is consistently entered in *italics*):

(3)
$$\mathit{sign}\begin{bmatrix} \text{HEAD } \mathit{verb} \\ \text{GF} \begin{bmatrix} \text{SUBJ } \mathit{sign}\begin{bmatrix}\text{INDX } \boxed{1}\ \mathit{index}\end{bmatrix} \\ \text{OBJ } \mathit{sign}\begin{bmatrix}\text{INDX } \boxed{2}\ \mathit{index}\end{bmatrix} \\ \text{IOBJ } \mathit{sign}\begin{bmatrix}\text{INDX } \boxed{3}\ \mathit{index}\end{bmatrix} \end{bmatrix} \\ \text{INDX } \mathit{index} \\ \text{ACTNTS} \begin{bmatrix} \text{ACT1 } \boxed{1}\ \mathit{index} \\ \text{ACT2 } \boxed{2}\ \mathit{index} \\ \text{ACT3 } \boxed{3}\ \mathit{index} \end{bmatrix} \\ \text{AKTART } \mathit{aktionsart} \end{bmatrix}$$

Here we see that the attribute INDX now consistently occurs inside the bracket prefixed by *sign*. So that is the rule regulating the use of the attribute INDX: it is used

for sorting the content of a *sign*. Likewise, HEAD, GF, ACTNTS and AKTART are attributes used for sorting the content of *sign*. One says that the type *sign* **declares** these attributes – that is to say, for an object of the kind ‘sign’, these are sorting attributes allowed to occur inside the bracket representing a *sign*.
Values of attributes always belong to some *type*, and in (3) we have entered the name of the type in question: *aktionsart*, *verb* and *index* are all types. Types generally have their supertypes or subtypes – in (4) we enumerate some:

(4)
verb has *pos* (part of speech: ‘ordklasse’) as supertype;
aktionsart has *state, event, activity, process, accomplishment, achievement*, ... as subtypes.

So when in (3) we see the somewhat non-imaginative-looking ‘AKTART *aktionsart*’, this is to be understood with *aktionsart* holding the place of all its possible subtypes mentioned in (4).
A given type can be the value of more than one attribute: for instance, in (3), we see that *index* is a value of both INDX and ACT1, ACT2, ACT3. Intuitively that is reasonable – many principles of sorting can lead to the same type of thing.
Conceivably, an attribute could be declared by more than one type as well – a given sorting attribute might be relevant in many connections. Here, though, formal systems differ, and in the one used here (see Copestake op. cit.), *an attribute can be declared only by one type*. If that type has *sub*types, then the attribute appears in all the subtypes as well (by what is called ‘inheritance’); but you never find an attribute inside two types which do not have a common *super*type.
Under such regulations, it is in practical expositions defensible to leave out many of the types mentioned in (3): since INDX, HEAD, GF etc are all declared by *sign*, the mention in (3) of *sign* can be left out: it is predictable. Moreover, if we have nothing to say about an attribute on a given occasion, there is no need to mention it – by the general attribute declaration of the type concerned, one knows which attributes are in principle there. In a case like representing the ditransitive verb construction in general, the part worth keeping is thus actually just (1) plus the HEAD specification – the rest of (3) is given by general regulations and declarations of the system. However, to illustrate the general feature architecture, we expose the richness exemplified in (3).

In analyzing a causative construction where the ‘base’ construction is transitive, with a ‘causer’ added making the construction ditransitive, a reasonable display could be:

(5)
$$
\textit{sign}\left[\begin{array}{l}
\text{HEAD } \textit{verb} \\
\text{GF } \left[\begin{array}{l}
\text{SUBJ } \textit{sign}\left[\text{INDX } \boxed{1}\ \textit{index}\right] \\
\text{OBJ } \textit{sign}\left[\text{INDX } \boxed{2}\ \textit{index}\right] \\
\text{IOBJ } \textit{sign}\left[\text{INDX } \boxed{3}\ \textit{index}\right]
\end{array}\right] \\
\text{INDX } \textit{index} \\
\text{ACTNTS } \left[\begin{array}{l}
\text{PRED cause-rel} \\
\text{ACT1 } \boxed{1}\ \textit{index} \\
\text{ACT2 } \left[\begin{array}{l}
\text{ACT1 } \boxed{2}\ \textit{index} \\
\text{ACT2 } \boxed{3}\ \textit{index}
\end{array}\right]
\end{array}\right] \\
\text{AKTART } \textit{aktionsart}
\end{array}\right]
$$

This exposes the subject as a causer, and the caused situation as one whose ACT1 (‘underlying subject’) is realized as (‘surface’) object and its ACT2 (‘underlying object’) as (‘surface’) indirect object. This is a picture we want to maintain.
However, keeping in mind the ‘traffic rules’ just laid out, one may ask whether the type *index*, which we have said is the value of ACT2, can be the type that declares the ‘inner’ attributes ACT1 and ACT2 in (5). That would not seem too plausible – ACTNTS is an attribute which unfolds a whole situation with its participants, and not just an identifying index. This is one problem with (5). Another is that in talking about the caused situation, we may want to also specify its *aktionsart*; and we may want to say that the caused situation is exactly the same situation as the one expressed by the ‘underlying’ verb, i.e., the root of the verb. Such expression of ‘sameness’ would typically involve identifiers, where the identifier would have all of the attributes AKTART, INDX and ACTNTS in its scope, and introduced by the outer ACT2 in (5) (see (8) below for an illustration, where the identifier is the boxed ‘$\boxed{6}$’). But if the value of this ACT2 is *index*, we would have the type *index* declaring the attribute INDX, which seems strange by itself, and here it would also be illegal, since INDX is aleady declared by *sign*. A similar collision would apply for AKTART.
Below is one possible reconciliation of interests, see (6): First, one lets the type *index* in general declare an attribute KIND, whose value may be either *sit* or *indiv*. The type *sit* in turn declares the attributes PRED, ACT0, ACT1,.., AKTRT, and is the value also of ACTNTS. In general, AKTART, declared by *sign*, and AKTRT, declared by *sit,* are re-entered for value, and likewise INDX, declared by *sign*, and ACT0, declared by *sit*. In this way, we avoid having nominally the same attribute declared by different types:

(6)
$$
\mathit{sign}\begin{bmatrix}
\text{HEAD } \mathit{verb} \\
\text{GF} \begin{bmatrix}
\text{SUBJ } \mathit{sign}\begin{bmatrix}\text{INDX } \boxed{1}\,\mathit{index}\end{bmatrix} \\
\text{OBJ } \mathit{sign}\begin{bmatrix}\text{INDX } \boxed{2}\,\mathit{index}\end{bmatrix} \\
\text{IOBJ } \mathit{sign}\begin{bmatrix}\text{INDX } \boxed{3}\,\mathit{index}\end{bmatrix}
\end{bmatrix} \\
\text{INDX } \boxed{4} \\
\text{ACTNTS } \mathit{sit-rel} \begin{bmatrix}
\text{PRED cause-rel} \\
\text{ACT0 } \boxed{4} \\
\text{ACT1 } \boxed{1}\ \mathit{index} \\
\text{ACT2 } \mathit{index} \begin{bmatrix}\text{KIND } \mathit{sit-rel} \begin{bmatrix}\text{ACT1 } \boxed{2}\ \mathit{index} \\ \text{ACT2 } \boxed{3}\ \mathit{index}\end{bmatrix}\end{bmatrix} \\
\text{AKTRT } \boxed{5}
\end{bmatrix} \\
\text{AKTART } \boxed{5}
\end{bmatrix}
$$

What we thereby secure is that a verb construction will always have its full semantic content assembled in a single structure, the one introduced by ACTNTS, for use, e.g., when we want to show how this full semantic structure behaves in a larger causative structure (see discussion below). At the same time, when discussion requires exposure only of the verb in isolation, there is no need to mention the ‘duplicate’ attributes ACT0 and AKTRT, since their value is anyway the same as those of INDX and AKTART.

Thus, a representation of a ditransitive non-causative construction will, for practical purposes, be as simple as originally conceived, i.e., as in (7):

(7)
$$\begin{bmatrix} \text{HEAD verb} & \\ \text{GF} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{INDX } \boxed{1} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{INDX } \boxed{2} \end{bmatrix} \\ \text{IOBJ} & \begin{bmatrix} \text{INDX } \boxed{3} \end{bmatrix} \end{bmatrix} \\ \text{ACTNTS} & \begin{bmatrix} \text{ACT1 } \boxed{1} \\ \text{ACT2 } \boxed{2} \\ \text{ACT3 } \boxed{3} \end{bmatrix} \end{bmatrix}$$

Which is not unimportant: simple matters should have a simple representation.

2. Derivation

Let us now expose more clearly the sense in which a morphologically causative construction is *derived* from its non-causative counterpart. The notion ‘derivation’ suggests a definable input and a definable output. The structure in (6) represents the output now in question; the input we represent by an attribute ‘DTR’ (for ‘daughter’), regarding the derived form as in some sense having the input form as a ‘daughter constituent’. As is shown in (8) below, the type of this ‘daughter’ is *sign*, and relative to this ‘input’ sign, the ACTNTS value is set identical to the ‘caused’ situation (the ACT2 of ‘cause-rel’), and the referents of the input SUBJ and OBJ are those of the output OBJ and IOBJ, respectively:

(8) Construction headed by verb derived by morphological causative:

$$sign\begin{bmatrix} \text{HEAD } \textit{verb} & \\ \text{GF} & \begin{bmatrix} \text{SUBJ } \textit{sign}\begin{bmatrix} \text{INDX } \boxed{1}\textit{index} \end{bmatrix} \\ \text{OBJ } \textit{sign}\begin{bmatrix} \text{INDX } \boxed{2}\textit{index} \end{bmatrix} \\ \text{IOBJ } \textit{sign}\begin{bmatrix} \text{INDX } \boxed{3}\textit{index} \end{bmatrix} \end{bmatrix} \\ \text{INDX } \boxed{4} & \\ \text{ACTNTS } \textit{sit-rel} & \begin{bmatrix} \text{PRED cause-rel} \\ \\ \text{ACT0 } \boxed{4} \\ \text{ACT1 } \boxed{1}\textit{index} \\ \text{ACT2 } \textit{index}\begin{bmatrix} \text{KIND } \boxed{6}\textit{sit-rel}\begin{bmatrix} \text{ACT1 } \boxed{2}\textit{index} \\ \text{ACT2 } \boxed{3}\textit{index} \end{bmatrix} \end{bmatrix} \\ \text{AKTRT } \boxed{5} \end{bmatrix} \\ \text{AKTART } \boxed{5} & \\ \text{DTR } \textit{sign} & \begin{bmatrix} \text{GF} \begin{bmatrix} \text{SUBJ } \textit{sign}\begin{bmatrix} \text{INDX } \boxed{2}\textit{index} \end{bmatrix} \\ \text{OBJ } \textit{sign}\begin{bmatrix} \text{INDX } \boxed{3}\textit{index} \end{bmatrix} \end{bmatrix} \\ \text{ACTNTS } \boxed{6} \end{bmatrix} \end{bmatrix}$$

In an AVM like this, as said above, one normally enters only identities and specific values not given by the general type and declarations system. (8) can thereby be shrunken down to (9) (also abbreviating ‘KIND’ to ‘K’):

(9) A leaner version of (8)

$$\left[\begin{array}{ll}
\text{HEAD } \textit{verb} & \\
\text{GF} & \left[\begin{array}{ll} \text{SUBJ} & [\text{INDX } \boxed{1}] \\ \text{OBJ} & [\text{INDX } \boxed{2}] \\ \text{IOBJ} & [\text{INDX } \boxed{3}] \end{array}\right] \\
\text{ACTNTS} & \left[\begin{array}{ll} \text{PRED cause-rel} & \\ \text{ACT1 } \boxed{1} & \\ \text{ACT2} & \left[\text{K } \boxed{6} \left[\begin{array}{l} \text{ACT1 } \boxed{2} \\ \text{ACT2 } \boxed{3} \end{array}\right]\right] \end{array}\right] \\
\text{DTR} & \left[\begin{array}{ll} \text{GF} & \left[\begin{array}{ll} \text{SUBJ} & [\text{INDX } \boxed{2}] \\ \text{OBJ} & [\text{INDX } \boxed{3}] \end{array}\right] \\ \text{ACTNTS } \boxed{6} & \end{array}\right]
\end{array}\right]$$

Suppose now that this causative construction undergoes *passivization*, through the agent becoming implicit, and the object being promoted to subject. This is another derivational process, and can be represented in exactly the same way as above. The whole structure (9) is now a ‘DTR’ relative to the new structure. A substantive difference between the causative and the passive derivation is that the situational content of the passive construction is presumably the same as that of its ‘active’ input, hence the values of ACTNTS and DTR|ACTNTS can be set identical in the new structure, as seen in (10).

(10) Verb construction derived by morphological causative and passive:

$$\left[\begin{array}{ll}
\text{HEAD } \textit{verb} & \\
\text{GF} & \left[\begin{array}{ll} \text{SUBJ} & [\text{INDX } \boxed{2}] \\ \text{OBJ} & [\text{INDX } \boxed{3}] \end{array}\right] \\
\text{ACTNTS } \boxed{7} & \\
\text{DTR} & \left[\begin{array}{ll}
\text{HEAD } \textit{verb} & \\
\text{GF} & \left[\begin{array}{ll} \text{SUBJ} & [\text{INDX } \boxed{1}] \\ \text{OBJ} & [\text{INDX } \boxed{2}] \\ \text{IOBJ} & [\text{INDX } \boxed{3}] \end{array}\right] \\
\text{ACTNTS } \boxed{7} & \left[\begin{array}{ll} \text{PRED cause-rel} & \\ \text{ACT1 } \boxed{1} & \\ \text{ACT2} & \left[\text{K } \boxed{6} \left[\begin{array}{l} \text{ACT1 } \boxed{2} \\ \text{ACT2 } \boxed{3} \end{array}\right]\right] \end{array}\right] \\
\text{DTR} & \left[\begin{array}{ll} \text{GF} & \left[\begin{array}{ll} \text{SUBJ} & [\text{INDX } \boxed{2}] \\ \text{OBJ} & [\text{INDX } \boxed{3}] \end{array}\right] \\ \text{ACTNTS } \boxed{6} & \end{array}\right]
\end{array}\right]
\end{array}\right]$$

The implicitness of the causative subject in this passsive construction is represented by the circumstance that in the situational content indicated by ‘ ACTNTS $\boxed{7}$ ’ of the passive output, there is an ACT1 whose index is not among the indices of the realized GFs. As for the explicit arguments, tracing, e.g., the referent indicated by ‘ $\boxed{2}$ ’ in the path GF.SUBJ.INDX $\boxed{2}$ down through the AVM, one sees that this referent comes up in the paths DTR.GF.OBJ.INDX $\boxed{2}$, and DTR.DTR.GF.SUBJ.INDX $\boxed{2}$, reflecting its GF realizations throughout the derivation; its semantic status is represented inside ACTNTS $\boxed{7}$.

The following are examples of (9) and (10):

(11) a. (example of (9), from Citumbuka (Jean Chavula, p.c.))
Mary wa-ka-mu-phik-**isk**-a Tumbikani nchunga
Mary 1SM-pst-1OM-cook-Caus-fV Tumbikani beans
'Mary made Tumbikani cook beans'

b. (example of (10), from Kiswahili, based on Vitale, 1981:165, quoted in Kroeger, 2004, p.196, ex. (11a))
mke wake a-li-pi-**ish**-**w**-a uji na Sudi
wife his S.agr-PAST-cook-CAUS-Pass-Ind gruel by Sudi
‘His wife was made to cook gruel by Sui’

It may be noted that in these derivations, *the semantics is monotonic*: from step to step, the semantics either stays *the same* (as in passive), or something is *added* to it (as in causativization), but nothing is ever *deleted* in the ‘input’ semantic structure. (*Syntactically* one might say that the linkage ‘causer -subject’ is deleted at the step of passivization in (10); however, in the semantic representation of the last step, the referent $\boxed{1}$ remains, reflecting the circumstance that in the resulting passive construction, the causative meaning is still maintained.) Such monotonicity has been one of the key characteristics of *transformations*, and in the format of rules producing representations like (9) and (10), often called *lexical rules* (for reasons to be seen), monotonicity is a defining property.

Let us look at the structure of another construction type traditionally counted as derived by a transformation, namely Raising-from-subject-to-object, as in (12a), which semantically is a near-equivalent of (12b):

(12) a. Jeg så henne komme
b. Jeg så at hun kom

Both express a relationship between a perceiver, ’I’, and a situation perceived, that ‘she came’. The syntactic structure of (12b) may be said to be isomorphic to this semantic structure, having two syntactic arguments of the verb; however, the syntactic structure of (12a), which has three syntactic argument constituents, is not isomorphic to this semantic structure. The AVM (13) for (12a) exposes this ‘skewed’ relation syntax-semantics. In (13), ACT2 is what is ‘seen’, constituted by the combination of the meanings of OBJ and SECPR in that the INDX of SECPR is equated with the event-index of ‘komme’, namely ACT0 of *komme-rel*, and the single participant of this event – its ACT1 – is the referent of the object. Thus, the second main semantic constituent – ACT2 – is related to both the OBJ, as its INDX, and to the SECPR, as its ACT1.

(13)
$$\left[\begin{array}{ll} \text{HEAD } \textit{verb} & \\ \text{GF} & \left[\begin{array}{ll} \text{SUBJ} & \left[\text{INDX } \boxed{1}\right] \\ \text{OBJ} & \left[\text{INDX } \boxed{2}\right] \\ \text{SECPR} & \left[\text{INDX } \boxed{3}\right] \end{array}\right] \\ \text{ACTNTS} & \left[\begin{array}{ll} \text{PRED se-rel} & \\ \text{ACT1 } \boxed{1} & \\ \text{ACT2} & \left[\text{K} \left[\begin{array}{l} \text{PRED komme-rel} \\ \text{ACT0 } \boxed{3} \\ \text{ACT1 } \boxed{2} \end{array}\right]\right] \end{array}\right] \end{array}\right]$$

Having thus represented the syntactic and semantic structure of (12a), where is the 'Raising' factor? In the times of the transformational model, there was no semantic representation offered like here, and to provide a structure with the essential properties of ACTNTS in (13), this had to be posited as a 'deep structure', from which the GF constellation of (13) was derived. Is there by itself any reason to build a syntactic derivation here, which would be essentially as in (14)?

(14)
$$\left[\begin{array}{ll} \text{HEAD } \textit{verb} & \\ \text{GF} & \left[\begin{array}{ll} \text{SUBJ} & \left[\text{INDX } \boxed{1}\right] \\ \text{OBJ} & \left[\text{INDX } \boxed{2}\right] \\ \text{SECPR} & \left[\begin{array}{l} \text{HEAD verb}\left[\text{FORM infinitive}\right] \\ \text{INDX } \boxed{3} \end{array}\right] \end{array}\right] \\ \text{ACTNTS } \boxed{4} & \\ \text{DTR} & \left[\begin{array}{ll} \text{SUBJ} & \left[\text{INDX } \boxed{1}\right] \\ \text{COMP} & \left[\begin{array}{l} \text{GF}\left[\text{SUBJ } \boxed{2}\right] \\ \text{INDX } \boxed{3} \end{array}\right] \\ \text{ACTNTS } \boxed{4} & \left[\begin{array}{ll} \text{PRED se-rel} & \\ \text{ACT1 } \boxed{1} & \\ \text{ACT2} & \left[\text{K} \left[\begin{array}{l} \text{PRED komme-rel} \\ \text{ACT0 } \boxed{3} \\ \text{ACT1 } \boxed{2} \end{array}\right]\right] \end{array}\right] \end{array}\right] \end{array}\right]$$

If the DTR in (14) is to be seen as an analogue to a deep structure of (12a), it should presumably also occur as DTR in a parallel derivation giving the *at*-clause construction, looking as in (15); the effect of this derivation would be essentially specifying the verb form as being *finite,* as opposed to the underspecified form of the DTR of both (14) and (15):

(15)

$$\begin{bmatrix}
\text{HEAD} & \textit{verb} \\
\text{GF} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix}\text{INDX } \boxed{1}\end{bmatrix} \\ \text{COMP} & \begin{bmatrix} \text{HEAD verb}\begin{bmatrix}\text{FORM finite}\end{bmatrix} \\ \text{GF}\begin{bmatrix}\text{SUBJ } \boxed{2}\end{bmatrix} \\ \text{INDX } \boxed{3} \end{bmatrix} \end{bmatrix} \\
\text{ACTNTS} & \boxed{4} \\
\text{DTR} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix}\text{INDX } \boxed{1}\end{bmatrix} \\ \text{COMP} & \begin{bmatrix} \text{GF}\begin{bmatrix}\text{SUBJ } \boxed{2}\end{bmatrix} \\ \text{INDX } \boxed{3}\end{bmatrix} \\ \text{ACTNTS } \boxed{4} & \begin{bmatrix} \text{PRED se-rel} & \\ \text{ACT1} & \boxed{1} \\ \text{ACT2} & \begin{bmatrix} \text{K} & \begin{bmatrix}\text{PRED komme-rel} \\ \text{ACT0 } \boxed{3} \\ \text{ACT1 } \boxed{2}\end{bmatrix}\end{bmatrix}\end{bmatrix}\end{bmatrix}
\end{bmatrix}$$

Since the rule in (15) doesn't accomplish very much, most grammars using the present formalism would state the verb entry corresponding to (12b) directly, and thereby also the entry corresponding to (12a), which is thus (13) (minus the embedded PRED-specification). That is to say, there would be no lexical rule corresponding to Raising-from-subject-to-object, but just one entry with a sign specification of the form of (13) for the raised version in (12), and another entry for the non-raised version. (But the issue of what is the preferable approach is perhaps not totally settled.)

For Raising-from-subject-to-subject, similar considerations apply. Analogously to (13) as a representation of (12a), (16b) will represent (16a):

(16)

a. Han synes syk.

b.

$$\begin{bmatrix}
\text{HEAD} & \textit{verb} \\
\text{GF} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix}\text{INDX } \boxed{1}\end{bmatrix} \\ \text{SECPR} & \begin{bmatrix}\text{INDX } \boxed{2}\end{bmatrix}\end{bmatrix} \\
\text{ACTNTS} & \begin{bmatrix} \text{PRED synes-rel} & \\ \text{ACT1} & \begin{bmatrix} \text{K} & \begin{bmatrix}\text{PRED syk-rel} \\ \text{ACT0 } \boxed{2} \\ \text{ACT1 } \boxed{1}\end{bmatrix}\end{bmatrix}\end{bmatrix}
\end{bmatrix}$$

Exercise: Make an AVM for *Hun ble sett komme*.

We have now illustrated some essential conventions of the AVM formalism, we have introduced a very specific version of the formalism for representing the syntax and semantics of *signs*, and we have shown how morphologically encoded derivation like

Causativization and Passive can be represented in the formalism through something called *lexical rules*. In the case of the Passive construction, the function of the lexical rule corresponds to what in a TG-grammar could be effected through a *transformation.* We have however also seen a case where what in TG-grammar is represented through a transformation can perhaps be adequately represented in a *non-derivational* sign, given the semantic part of the AVM as a locus for specifying what a Deep structure could be invoked for representing in a TG-grammar.

3. Raising and Equi constructions

We now proceed with generalizing the above account of Raising, and in comparing the AVM representation of *Raising* with that of *Equi*. First the generalization:

In both (13) and (16), the embedded verb is intransitive, so that its actant can be identified as ‘ACT1’. However, suppose that the embedded verb is transitive, and has undergone Passive, as in

(17) Jeg så ham bli sparket.

In this case, the general structure of an AVM of Raising-from-subject-to-object as given in (13), repeated,

(13)
$$\begin{bmatrix} \text{HEAD } \textit{verb} & \\ \text{GF} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix}\text{INDX } \boxed{1}\end{bmatrix} \\ \text{OBJ} & \begin{bmatrix}\text{INDX } \boxed{2}\end{bmatrix} \\ \text{SECPR} & \begin{bmatrix}\text{INDX } \boxed{3}\end{bmatrix} \end{bmatrix} \\ \text{ACTNTS} & \begin{bmatrix} \text{PRED se-rel} & \\ \text{ACT1 } \boxed{1} & \\ \text{ACT2} & \begin{bmatrix} \text{K} & \begin{bmatrix} \text{PRED komme-rel} \\ \text{ACT0 } \boxed{3} \\ \text{ACT1 } \boxed{2} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

is inadequate, since the ‘raised’ NP in this case does not represent the ACT1 of the embedded verb, but rather its ACT2. What this ACT2 of ‘bli sparket’ has in common with the ACT1 of ‘komme’ is that they both have a claim to the underlying subject position of the verb, be it as a ‘passive’ subject or an ‘active’ subject. This could mean that if we are to provide a general schema of this Raising construction, it should be as in (18), where we specify that the SECPR has a subject identical to the upper object, whereas we leave open how the ACTants inside the embedded clause are realized:

(18)
$$\begin{bmatrix} \text{HEAD } \textit{verb} \\ \text{GF} \begin{bmatrix} \text{SUBJ} \begin{bmatrix}\text{INDX } \boxed{1}\end{bmatrix} \\ \text{OBJ} \begin{bmatrix}\text{INDX } \boxed{2}\end{bmatrix} \\ \text{SECPR} \begin{bmatrix} \text{GF} \begin{bmatrix}\text{SUBJ}\begin{bmatrix}\text{INDX } \boxed{2}\end{bmatrix}\end{bmatrix} \\ \text{INDX } \boxed{3}\end{bmatrix}\end{bmatrix} \\ \text{ACTNTS} \begin{bmatrix} \text{PRED se-rel} \\ \text{ACT1 } \boxed{1} \\ \text{ACT2} \begin{bmatrix}\text{K}\begin{bmatrix}\text{ACT0 } \boxed{3}\end{bmatrix}\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

A similar amendment will apply for Raising-of-subject-to-subject, for cases like *Han synes å bli sparket.*
The diagram (18) is a bit similar to what one will have for an Equi-construction such as *Han ba meg komme*: the difference is that in the general schema for the latter type of construction – Equi-from-object-to-comp – the upper object will correspond to an independent actant in the upper clause, as in (19), which it does not in (18):

(19)
$$\begin{bmatrix} \text{HEAD } \textit{verb} \\ \text{GF} \begin{bmatrix} \text{SUBJ} \begin{bmatrix}\text{INDX } \boxed{1}\end{bmatrix} \\ \text{OBJ} \begin{bmatrix}\text{INDX } \boxed{2}\end{bmatrix} \\ \text{COMP} \begin{bmatrix} \text{GF} \begin{bmatrix}\text{SUBJ}\begin{bmatrix}\text{INDX } \boxed{2}\end{bmatrix}\end{bmatrix} \\ \text{INDX } \boxed{3}\end{bmatrix}\end{bmatrix} \\ \text{ACTNTS} \begin{bmatrix} \text{PRED be-rel} \\ \text{ACT1 } \boxed{1} \\ \text{ACT2 } \boxed{2} \\ \text{ACT3} \begin{bmatrix}\text{K}\begin{bmatrix}\text{ACT0 } \boxed{3}\end{bmatrix}\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

If 'Raising' is supposed to include also constructions with embedded adjectives rather than verbs, as in (16), or PPs, as in *Jeg så ham under bordet*, then a question arising from using the schema (18) as a general schema for Raising will be if adjectives and prepositions can generally be assumed to have subjects. This question we leave for the time being.

Why are the rules for Passive and Causativization as we have now conceived them often called *Lexical Rules*? In our discussion so far, all they have been used for is to derive the specification of one construction type from the specification of another construction type (– although in one case, the contrast between the types is reflected in a verbal affix). The label 'lexical' stems from an analytic strategy of some frameworks working with AVMs, namely that of attributing as many properties as possible of a construction to its *head* – this is reflected in 'Lexical' in 'LFG', and 'Head' in 'HPSG'. By consequence, relations between construction types are then construed as relations between the heads of these construction types; and when the relations are construed as 'rules', they thus become referrable to as 'lexical rules'.

The analytic strategy in question is motivated by concerns of *building parsers*. What we are doing so far is not exactly that, only *representing properties* of constructions. How to make parsers that can actually produce displays of these properties for any actual sentence belonging to a given construction type, is a further step, which logically presupposes the inventory of representations of construction types now being created. The step is not a very large step; but we will reserve it for later. Still, as for what to call the rules in question, we may as well call them ‘lexical rules’ from the outset.

4. Further construction types

We will now construct general AVM schemata for some further construction types, some of them involving ‘lexical rules’, representing phenomena discussed in (Kroeger 2004). We will also introduce type names for the various construction types, brief, but with some transparency as to what the constructions contain.

4.1. Applicatives

In many cases, Applicative formation can be viewed as a derivation from a construction with an oblique argument to one where the governee of the oblique is promoted to object, and where the semantics remains the same. At least two cases can be distringuished – where the ‘input’ construction has no object, as in (20), and one where it already has an object, as in (21), where in turn two options are distinguished, A and B:

(20) Applicative formation: intrObl -> tr

$$
\begin{bmatrix}
\text{HEAD } \textit{verb} & \\
\text{GF} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{INDX } \boxed{1} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{INDX } \boxed{2} \end{bmatrix} \end{bmatrix} \\
\text{ACTNTS } \boxed{6} & \begin{bmatrix} \text{ACT1 } \boxed{1} \\ \text{ACTobl } \boxed{2} \end{bmatrix} \\
\text{DTR} & \begin{bmatrix} \text{GF} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{INDX } \boxed{1} \end{bmatrix} \\ \text{OBL} & \begin{bmatrix} \text{GOV} \begin{bmatrix} \text{INDX } \boxed{2} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{ACTNTS } \boxed{6} & \end{bmatrix}
\end{bmatrix}
$$

(21) Applicative formation: trObl -> ditr

A. obl -> ob and original ob -> ob2

$$\begin{bmatrix} \text{HEAD } \textit{verb} & \\ \text{GF} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix}\text{INDX } \boxed{1}\end{bmatrix} \\ \text{OBJ} & \begin{bmatrix}\text{INDX } \boxed{3}\end{bmatrix} \\ \text{OBJ2} & \begin{bmatrix}\text{INDX } \boxed{2}\end{bmatrix} \end{bmatrix} \\ \text{ACTNTS } \boxed{6} & \begin{bmatrix} \text{ACT1 } \boxed{1} \\ \text{ACT2 } \boxed{2} \\ \text{ACTobl } \boxed{3} \end{bmatrix} \\ \text{DTR} & \begin{bmatrix} \text{GF} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix}\text{INDX } \boxed{1}\end{bmatrix} \\ \text{OBJ} & \begin{bmatrix}\text{INDX } \boxed{2}\end{bmatrix} \\ \text{OBL} & \begin{bmatrix}\text{GOV} \begin{bmatrix}\text{INDX } \boxed{3}\end{bmatrix}\end{bmatrix} \end{bmatrix} \\ \text{ACTNTS } \boxed{6} & \end{bmatrix} \end{bmatrix}$$

B. obl -> iob and original ob -> ob

$$\begin{bmatrix} \text{HEAD } \textit{verb} & \\ \text{GF} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix}\text{INDX } \boxed{1}\end{bmatrix} \\ \text{IOBJ} & \begin{bmatrix}\text{INDX } \boxed{3}\end{bmatrix} \\ \text{OBJ} & \begin{bmatrix}\text{INDX } \boxed{2}\end{bmatrix} \end{bmatrix} \\ \text{ACTNTS } \boxed{6} & \begin{bmatrix} \text{ACT1 } \boxed{1} \\ \text{ACT2 } \boxed{2} \\ \text{ACTobl } \boxed{3} \end{bmatrix} \\ \text{DTR} & \begin{bmatrix} \text{GF} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix}\text{INDX } \boxed{1}\end{bmatrix} \\ \text{OBJ} & \begin{bmatrix}\text{INDX } \boxed{2}\end{bmatrix} \\ \text{OBL} & \begin{bmatrix}\text{GOV} \begin{bmatrix}\text{INDX } \boxed{3}\end{bmatrix}\end{bmatrix} \end{bmatrix} \\ \text{ACTNTS } \boxed{6} & \end{bmatrix} \end{bmatrix}$$

(A is a pattern exemplified in example (8) in Appendix 1 below. B is a pattern corresponding to *jeg brakte kaffe til ham* -> # *jeg tilbrakte ham kaffe*.)

4.2. Ergative constructions, and role, case and agreement

In displaying ergativity, we need to expose the features relevant in *morphological ergativity*, and those relevant in *syntactic ergativity*. In the latter case, the essential factor is whether *roles* like agent and patient get realized as subject or object. Morphological ergativity can reside in *agreement* between transitive verb and object corresponding to that between intransitive verb and subject, or in *case* of the object of a transitive verb being equal to that of a subject of an intransitive verb. None of these factors have been encoded in our AVMs so far, so let us first consider possible encodings in general.

Roles will naturally be associated with the ACTants of a situation – in the form ‘ACT1.ROLE ag’, ‘ACT2.ROLE pat’, etc. Since the value of ACT1 etc. is *index*, this means that *index* has to declare ROLE as an attribute alongside KIND.

Case of an NP may be associated with the head noun, as the simplest case. To associate an attribute directly with the type *noun*, we may let *noun* declare that attribute. CASE will then be one attribute declared by *noun*; others will soon be mentioned.

When an NP *agrees* with a verb or adjective, certain properties of that NP are thereby signalled, such as noun class (gender), number, person. Some of these properties may be encoded by attributes declared by *noun*, such as noun class and person; number may rather be attributed to the index of the NP. Anyhow, one has to distinguish between these properties as *properties of an NP signalled* in an agreement relation, and the circumstance that an NP is at all *targeted* by an agreement relation. Across languages, subjects are often targeted by agreement, objects as well often, although less often; even indirect objects and second objects may be the targets of agreement. The properties of an NP that are signalled in an agreement relation in a given language may be the same regardless of the GF of the NP (e.g., in Kiswahili, a subject marker and an object marker represent the same range of properties of the targeted

NPs), but whether the NP is agreement-targeted depends on its GF (relative to the language in question – so, in Kiswahili, an object NP can be agreement targeted, in German, not). We may encode the targeting relationship by an attribute AGR-TARGET +/-, declared by the type *noun*, so that a noun sign like *den Hund* in *Die Katze hat den Hund zerbissen* will include (22) in its specification (including number in the head specification):

(22)
$$\left[\begin{array}{l} \text{HEAD } \textit{noun} \left[\begin{array}{l}\text{CASE acc}\\ \text{NCL-GND masc}\\ \text{PERS 3}\\ \text{NUM sg}\\ \text{AGR-TARGET -}\end{array}\right] \\ \\ \text{INDX } \boxed{1}\left[\text{ROLE pat}\right] \\ \text{ACTNTS} \left[\begin{array}{l}\text{PRED hund-rel}\\ \text{ACT0 } \boxed{1}\end{array}\right]\end{array}\right]$$

In the specification of a verb, when a subject or object NP has to agree with it, this is encoded in the valence specification of the verb. Thus, in a German transitive verb (which agrees only with the subject), if the agreement marker carried by the verb indicates 3.p sg, the following partial AVM is induced for the verb construction:

(23)
$$\left[\begin{array}{l} \text{HEAD } \textit{verb} \\ \text{GF} \left[\text{SUBJ} \left[\text{HEAD noun} \left[\begin{array}{l}\text{PERS 3}\\ \text{NUM sg}\\ \text{AGR-TARGET +}\end{array}\right]\right]\right]\end{array}\right]$$

We can now state the profile of *syntactic ergativity*, as in (24)

(24)
$$\left[\begin{array}{l} \text{HEAD } \textit{verb} \\ \text{GF} \left[\begin{array}{l}\text{SUBJ } \left[\text{INDX } \left[\text{ROLE pat}\right]\right]\\ \text{OBJ } \left[\text{INDX } \left[\text{ROLE ag}\right]\right]\end{array}\right]\end{array}\right]$$

The profile of *morphological* ergativity by *case* is as follows:

(25)
$$\left[\begin{array}{l} \text{HEAD } \textit{verb} \\ \text{GF} \left[\begin{array}{l}\text{SUBJ } \left[\text{HEAD}\left[\text{CASE erg}\right]\right]\\ \text{OBJ } \left[\text{HEAD}\left[\text{CASE abs}\right]\right]\end{array}\right]\end{array}\right]$$

The profile of *morphological* ergativity by *agreement* is as follows:

(26)
$$\left[\begin{array}{l} \text{HEAD } \textit{verb} \\ \text{GF} \left[\begin{array}{l}\text{SUBJ } \left[\text{HEAD}\left[\text{AGR-TARGETED -}\right]\right]\\ \text{OBJ } \left[\text{HEAD}\left[\text{AGR-TARGETED +}\right]\right]\end{array}\right]\end{array}\right]$$

Merging these profiles together, the ‘fully ergative’ profile would become:

(27)

$$\left[\begin{array}{l} \text{HEAD } \textit{verb} \\ \text{GF} \left[\begin{array}{l} \text{SUBJ} \left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \text{CASE erg} \\ \text{AGR-TARGETED -} \end{array}\right] \\ \text{INDX} \left[\text{ROLE pat}\right] \end{array}\right] \\ \text{OBJ} \left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \text{CASE abs} \\ \text{AGR-TARGETED +} \end{array}\right] \\ \text{INDX} \left[\text{ROLE ag}\right] \end{array}\right] \end{array}\right] \end{array}\right]$$

5. Composing a construction representation in parsing

We now address algorithmic strategies for composing constructional representations of sentences given as a string. For instance, an AVM of the full construction *Felix hat Fido zerbissen* can be viewed as in (28), and we will consider how such a representation can be constructed by a parser, i.e., an assembling mechanism processing the string word by word.

(28)

$$\left[\begin{array}{l} \text{HEAD } \textit{verb} \\ \text{GF} \left[\begin{array}{l} \text{SUBJ} \left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \text{CASE nom} \\ \text{PERS 3} \\ \text{NUM sg} \\ \text{AGR-TARGETED +} \end{array}\right] \\ \text{INDX } \boxed{1} \left[\text{ROLE ag}\right] \\ \text{ACTNTS} \left[\text{PRED Felix-rel}\right] \end{array}\right] \\ \text{OBJ} \left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \text{CASE acc} \\ \text{PERS 3} \\ \text{NUM sg} \\ \text{AGR-TARGETED -} \end{array}\right] \\ \text{INDX } \boxed{2} \left[\text{ROLE pat}\right] \\ \text{ACTNTS} \left[\text{PRED Fido-rel}\right] \end{array}\right] \end{array}\right] \\ \text{ACTNTS} \left[\begin{array}{l} \text{PRED zerbeissen-rel} \\ \text{ACT1 } \boxed{1} \\ \text{ACT2 } \boxed{2} \end{array}\right] \\ \text{AKTART accomplishment} \end{array}\right]$$

In HPSG, the standard strategy for this purpose is to specify the *head* of a verbal construction as a skeleton of the whole construction, with underspecified values for those attributes whose content depends on what words actually occur with the verb. In the first place, for the verb stem *zerbeissen*, this will be an AVM such as (29) (ignoring the feature ‘AGR-TARGETED’ in the following):

(29)
$$\left[\begin{array}{l}
\text{HEAD } \textit{verb} \\
\text{GF} \left[\begin{array}{l}
\text{SUBJ} \left[\begin{array}{l} \text{HEAD}\left[\text{CASE nom}\right] \\ \text{INDX } \boxed{1}\left[\text{ROLE ag}\right] \end{array}\right] \\
\text{OBJ} \left[\begin{array}{l} \text{HEAD}\left[\text{CASE acc}\right] \\ \text{INDX } \boxed{2}\left[\text{ROLE pat}\right] \end{array}\right]
\end{array}\right] \\
\text{ACTNTS} \left[\begin{array}{l} \text{PRED zerbeissen-rel} \\ \text{ACT1 } \boxed{1} \\ \text{ACT2 } \boxed{2} \end{array}\right] \\
\text{AKTART accomplishment}
\end{array}\right]$$

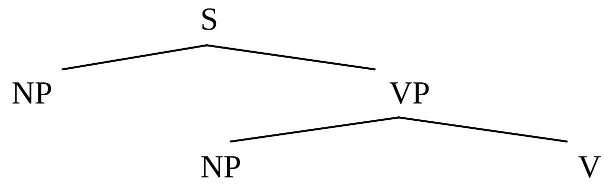
Supplementing this specification is a mechanism for distributing the representations of words combining with *zerbeissen* into the right slots of (29), to yield (28). If one assumes that the verb makes its combinations corresponding to the structure of a standard syntactic tree, and that this tree will have a VP containing the complements, with the subject sitting right underneath the S node, one can posits two ‘programs’ for combination, one for inside the VP, and one for outside the VP, standardly named ‘COMPS’ and ‘SPR’, respectively. These ‘programs’ are lists of constituents to be combined with, which is to say that the attributes COMPS and SPR take *lists* as values, the content of the lists being determined by the valence of the verb, i.e., the valence-bound arguments present in the construction. In this list specification for *zerbeissen*, COMPS and SPR will have one member each, the member on the COMPS list being identical to the value of GF.OBJ, and the member on the SPR list being identical to the value of GF.SUBJ. This is displayed in (30):

(30)
$$\left[\begin{array}{l}
\text{HEAD } \textit{verb} \\
\text{GF} \left[\begin{array}{l}
\text{SUBJ } \boxed{3} \left[\begin{array}{l} \text{HEAD}\left[\text{CASE nom}\right] \\ \text{INDX } \boxed{1}\left[\text{ROLE ag}\right] \end{array}\right] \\
\text{OBJ } \boxed{4} \left[\begin{array}{l} \text{HEAD}\left[\text{CASE acc}\right] \\ \text{INDX } \boxed{2}\left[\text{ROLE pat}\right] \end{array}\right]
\end{array}\right] \\
\text{SPR } \left\langle \boxed{3} \right\rangle \\
\text{COMPS } \left\langle \boxed{4} \right\rangle \\
\text{ACTNTS} \left[\begin{array}{l} \text{PRED zerbeissen-rel} \\ \text{ACT1 } \boxed{1} \\ \text{ACT2 } \boxed{2} \end{array}\right] \\
\text{AKTART accomplishment}
\end{array}\right]$$

The verb and the adjacent object NP are registered by the ‘assembling’ mechanism as constituting a constellation where that NP *instantiates* the item posited on the COMPS list. The item $\boxed{4}$ in (30) is thus identified with the specification of the sign represented by the object NP. This is the plan for how the specification of ‘Fido’ reaches its place in (28). Correspondingly for ‘Felix’, and the item $\boxed{3}$. When an item on a list is registered in this manner, it is deleted from the list, signifying that the valence of the verb is ‘satisfied’ with regard to this category. Thus, the item on the COMPS list is deleted, but its value is ‘stored’ as the value of OBJ, given the reentrancy between these items, and likewise for the item on the SPR list. A string is

accepted by the assembling mechanism only if, at the end, both of these lists are empty.
To illustrate, given a syntactic structure of the form (31) (ignoring *hat*),

(31)



when the mechanism reads the VP local tree, it is instructed to interpret the NP as instantiating the item in the COMPS list in (30); it thus identifies the symbol $\boxed{4}$ with the sign structure already built up for *Fido,* and deletes the item from the COMPS list. Correspondingly when it reads the S local tree, it is instructed to interpret the NP as instantiating the item in the SPR list in (30); it thus identifies the symbol $\boxed{3}$ with the sign structure already built up for *Felix*, and deletes the item from the SPR list. The resulting AVM is thereby (32), where '$\langle\ \rangle$' means 'empty list'; (32) is the same as (28), except for the valence lists:

(32)
$$\left[\begin{array}{l} \text{HEAD } \textit{verb} \\ \text{GF} \left[\begin{array}{l} \text{SUBJ} \left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l}\text{CASE nom}\\ \text{PERS 3}\\ \text{NUM sg}\end{array}\right] \\ \text{INDX } \boxed{1}\left[\text{ROLE ag}\right] \\ \text{ACTNTS}\left[\text{PRED Felix-rel}\right]\end{array}\right] \\ \text{OBJ} \left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l}\text{CASE acc}\\ \text{PERS 3}\\ \text{NUM sg}\end{array}\right] \\ \text{INDX } \boxed{2}\left[\text{ROLE pat}\right] \\ \text{ACTNTS}\left[\text{PRED Fido-rel}\right]\end{array}\right] \end{array}\right] \\ \text{SPR } \langle\ \rangle \\ \text{COMPS } \langle\ \rangle \\ \text{ACTNTS} \left[\begin{array}{l}\text{PRED zerbeissen-rel}\\ \text{ACT1 } \boxed{1}\\ \text{ACT2 } \boxed{2}\end{array}\right] \\ \text{AKTART accomplishment} \end{array}\right]$$

At this point. the specifications for PERS and NUM on both SUBJ and OBJ reflect properties of the actually occurring NPs. However, for the subject, also these properties are required, not by the verb lexeme, but by the *inflected verb*, signalled by the form *hat*. As it is the fully inflected form of the verb which is processed by the assembling mechanism, also the 3.sg. nature of the subject is specified in the input to the mechanism, thus, (33) rather than (30) is that input (but the output is (32) still):

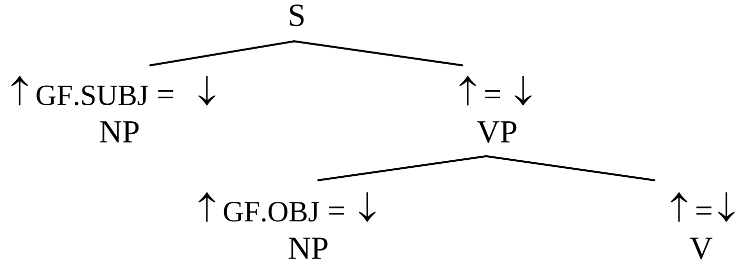
(33)

$$\begin{bmatrix} \text{HEAD } \textit{verb} \\ \text{GF} \begin{bmatrix} \text{SUBJ } \boxed{3} \begin{bmatrix} \text{HEAD} \begin{bmatrix} \text{CASE nom} \\ \text{PERS 3} \\ \text{NUM sg} \end{bmatrix} \\ \text{INDX } \boxed{1} \begin{bmatrix} \text{ROLE ag} \end{bmatrix} \end{bmatrix} \\ \text{OBJ } \boxed{4} \begin{bmatrix} \text{HEAD} \begin{bmatrix} \text{CASE acc} \end{bmatrix} \\ \text{INDX } \boxed{2} \begin{bmatrix} \text{ROLE pat} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{SPR } \left\langle \boxed{3} \right\rangle \\ \text{COMPS } \left\langle \boxed{4} \right\rangle \\ \text{ACTNTS} \begin{bmatrix} \text{PRED zerbeissen-rel} \\ \text{ACT1 } \boxed{1} \\ \text{ACT2 } \boxed{2} \end{bmatrix} \\ \text{AKTART accomplishment} \end{bmatrix}$$

In this ‘cancellation’ strategy, the situation where a string has less items than what the verb requires, is accommodated through the requirement that all items on the valence lists must have been processed. If a string has more items than the valence lists accommodate, it will mean that there are items that the assmebling mechanism has not been able to accommodate; in this case too, the string is then not accepted. Obviously, the specifications associated with the individual verb lexemes are crucial in this approach.

A similar basic view underlies the algorithms used in LFG. It may be illustrated as follows, with the proviso that the actual AVMs used here do not in all respects coincide with those used in LFG.
The syntactic analysis would be as above, and the verb specification as in (33) (minus the valence lists); however, the syntactic tree would be annotated in the manner shown in (34):

(34)



These annotations indicate that in building the AVM corresponding to the tree, the specification of the subject NP (signalled by ‘↓’ over the subject node, having been composed at that point) should be inserted at the feature path GF.SUBJ of the dominating node, and correspondingly for the object NP. ‘↑=↓’ over each V-projection node means that the basic AVM structure of the dominating node is equal to that of the node itself, except for whatever supplementary specification may be provided by the NPs. In this way, the inclusion of the contents of the NPs will plenish the underspecified SUBJ and OBJ slots of (33) to become (32) (minus the valence lists) in analogy to how this is done with the valence-list strategy for HPSG.

Also LFG has mechanisms for preventing the situation where a string has less items than the verb requires, as well as that where a string has more items than the valence of the verb allows. The names of these mechanisms are ‘completeness’ and ‘coherence’, respectively; they are implemented in computational applications of LFG grammars, but less exposed in the general view of the analysis than we saw for HPSG.
In the following we relate primarily to the HPSG formalism. However, it may be noted that in standard HPSG grammars, a field of attributes like those we have under ‘GF’ is normally not assumed – instead, it is assumed that valence lists sufficiently indicate GFs. Notably, in such a design, the view of the full construction will have no indication of what are the grammatical functions realized and by what, since the valence lists are all empty by the time the top S node has been processed. We regard that as a drawback of that design.
In return, it may be noted that in standard LFG analyses, where GFs are most explicitly exposed for the full construction, no explicit semantic representation is exposed in the way we do it in (33)/(32) – in that respect, thus, the present illustration is not representative of standard LFG.

Let us now first return to the notion ‘lexical rule’, from the perspective of HPSG or LFG as theories of *parsing*. Situated in frameworks where the main specification of a construction is encoded in the specification of a lexeme – as in these theories as theories of parsing -, statements of the effect of constructional operations like Passive, Applicative, Causativization, etc. must be related to specifications of lexemes as well, viz. as statements *deriving one lexeme from another*. And that is what a ‘lexical rule’ is. In such a setting, a statement like (20) for an Applicative construction will need to be enriched with information about valence lists, but otherwise stay the same, thus (20) will emerge as in (20’):

(20’) Applicative formation: intrObl -> tr

$$\begin{bmatrix} \text{HEAD } \textit{verb} \\ \text{GF} \begin{bmatrix} \text{SUBJ } \boxed{3} \begin{bmatrix} \text{INDX } \boxed{1} \end{bmatrix} \\ \text{OBJ } \boxed{4} \begin{bmatrix} \text{INDX } \boxed{2} \end{bmatrix} \end{bmatrix} \\ \text{SPR } \langle \boxed{3} \rangle \\ \text{COMPS } \langle \boxed{4} \rangle \\ \text{ACTNTS } \boxed{6} \begin{bmatrix} \text{ACT1 } \boxed{1} \\ \text{ACTobl } \boxed{2} \end{bmatrix} \\ \text{DTR} \begin{bmatrix} \text{GF} \begin{bmatrix} \text{SUBJ } \boxed{3} \begin{bmatrix} \text{INDX } \boxed{1} \end{bmatrix} \\ \text{OBL } \boxed{5} \begin{bmatrix} \text{GOV } \boxed{4} \begin{bmatrix} \text{INDX } \boxed{2} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{SPR } \langle \boxed{3} \rangle \\ \text{COMPS } \langle \boxed{5} \rangle \\ \text{ACTNTS } \boxed{6} \end{bmatrix} \end{bmatrix}$$

Then we illustrate the above outline with a simple syntactic construction. Let us consider *Katt fanger mus* (as it might occur in a newspaper headline, avoiding questions pertaining to determiners):

(35)
$$\begin{bmatrix} \text{HEAD verb} & \\ \text{GF} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{INDX } \boxed{1} \\ \text{ACTNTS} \begin{bmatrix} \text{PRED katt-rel} \end{bmatrix} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{INDX } \boxed{2} \\ \text{ACTNTS} \begin{bmatrix} \text{PRED mus-rel} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{ACTNTS} & \begin{bmatrix} \text{PRED fange-rel} \\ \text{ACT1 } \boxed{1} \\ \text{ACT2 } \boxed{2} \end{bmatrix} \end{bmatrix}$$

Here we first have to digress into how to represent NPs. As seen in (35), the subject and object have a semantic specification apart from INDX, namely a PRED-value. The attribute PRED is one we have so far used only for situational specifications, with *sit-rel* as the type taken as value by ACTNTS and introducing (declaring) the attributes PRED, ACT0, ACT1, etc. (cf. (6) above). Obviously both situations and individuals can be named, but it may seem a bit awkward if individual-names are to be introduced under the bells of situational specification. Let us stipulate the following types and inheritance relations:[1]

relation
[PRED name
ACT0 index]

sit-rel
[ACT0 sit-index]

indiv-rel
[ACT0 indiv-index]

index

sit-index
[KIND sit-rel]

indiv-index

Sit-rel is thereby in practice a 'recursive' category, as we established in section 1, for use when participants inside of a situation are themselves situations. Thus, this is a recursion to be made use of for ACT1, ACT2, ACT3 and ACTobl, but not inside of ACT0 itself when specified as 'ACT0 sit-index'. How to ensure the latter is a technical point we need not go into here; observe, however, that for an *indiv*-level participant, there is no further recursion inside its specification. In this respect, individuals are the 'base' entities in the system.

A more explicit version of (35) is thereby (36):

---

[1] Earlier on, and in class, we also considered the features 'ROLE' and 'HAECC', the latter to be introduced later in the text.

(36)

$$\begin{bmatrix} \text{HEAD verb} & \\ \text{GF} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{INDX } \boxed{1} \text{ indiv-index} \\ \text{ACTNTS indiv-rel} \begin{bmatrix} \text{PRED katt-rel} \\ \text{ACT0 } \boxed{1} \end{bmatrix} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{INDX } \boxed{2} \text{ indiv-index} \\ \text{ACTNTS indiv-rel} \begin{bmatrix} \text{PRED mus-rel} \\ \text{ACT0 } \boxed{2} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{INDX } \boxed{0} \text{ sit-index} & \\ \text{ACTNTS sit-rel} & \begin{bmatrix} \text{PRED fange-rel} \\ \text{ACT0 } \boxed{0} \\ \text{ACT1 } \boxed{1} \\ \text{ACT2 } \boxed{2} \end{bmatrix} \end{bmatrix}$$

We then turn to the illustration of syntactic combination. (35)/(36) is obviously the result of a composition, produced by a parsing algorithm taking the string *katt fanger mus* as input, and having at its disposal lexical specifications for *katt, fange,* and *mus*, and rules for combining these. Conceiving these in the HPSG way described above, the lexical entry for *fange* will be (37) – like (35)/(36), but with the valence lists in addition:

(37)

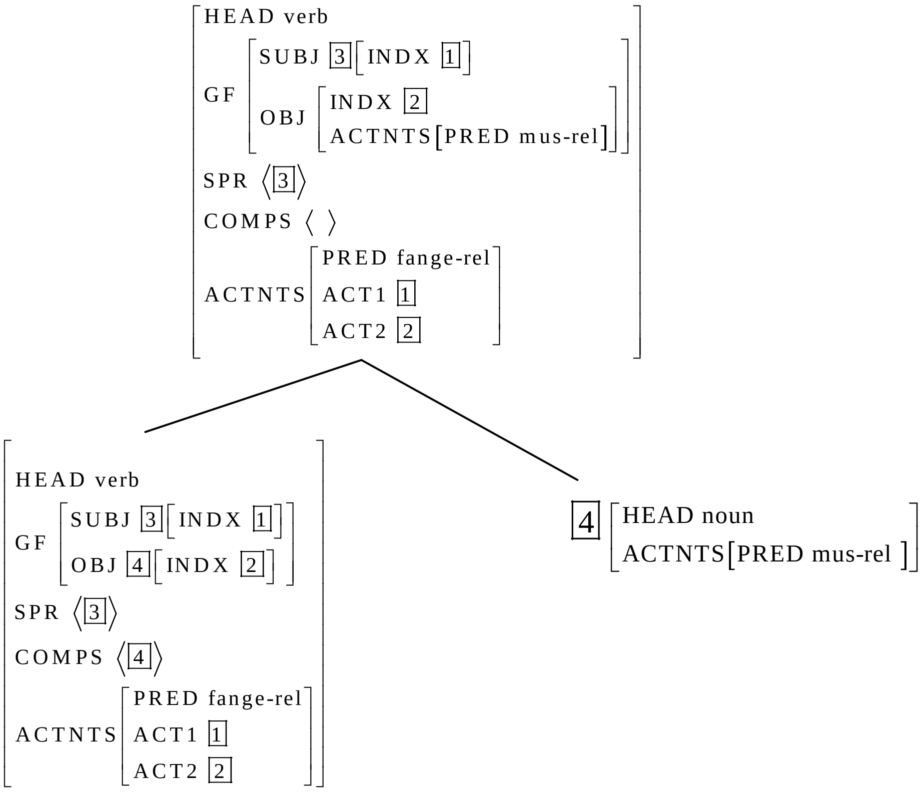
$$\begin{bmatrix} \text{HEAD verb} & \\ \text{GF} & \begin{bmatrix} \text{SUBJ } \boxed{3} \begin{bmatrix} \text{INDX } \boxed{1} \end{bmatrix} \\ \text{OBJ } \boxed{4} \begin{bmatrix} \text{INDX } \boxed{2} \end{bmatrix} \end{bmatrix} \\ \text{SPR} & \left\langle \boxed{3} \right\rangle \\ \text{COMPS} & \left\langle \boxed{4} \right\rangle \\ \text{ACTNTS} & \begin{bmatrix} \text{PRED fange-rel} \\ \text{ACT1 } \boxed{1} \\ \text{ACT2 } \boxed{2} \end{bmatrix} \end{bmatrix}$$

The lexical entries for *katt* and *mus* will be (38a,b) (here redundantly showing the INDX – ACT0 reentrancies, not to be repeated later):

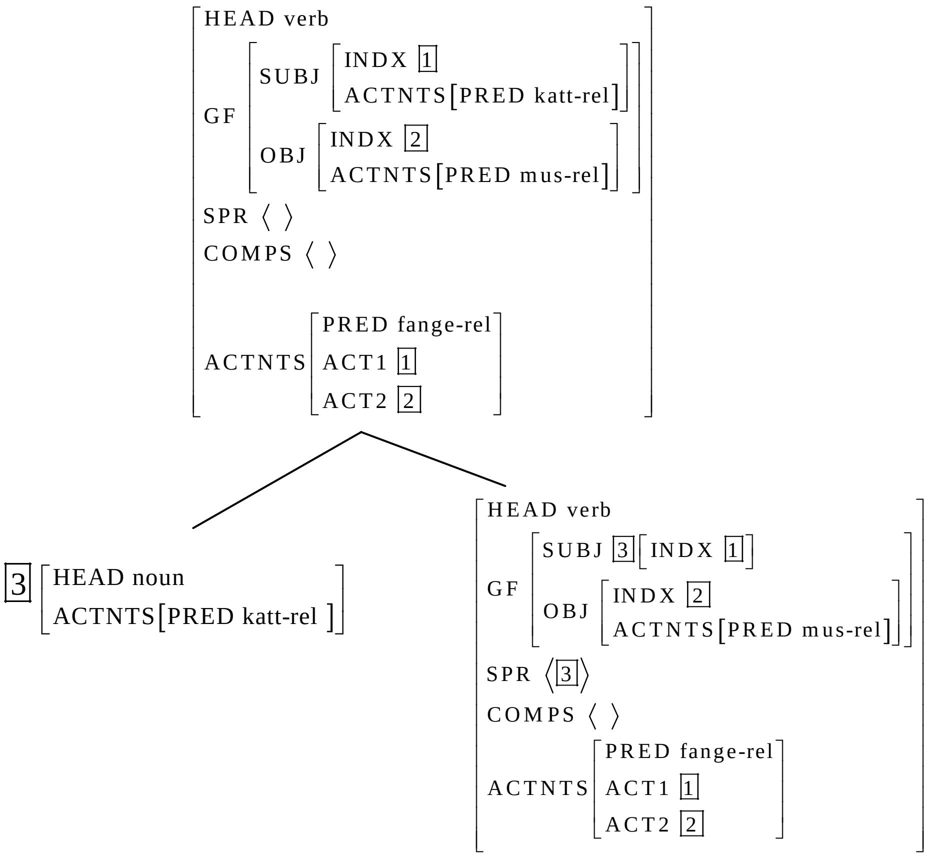
(38) a. $\begin{bmatrix} \text{HEAD noun} & \\ \text{INDX } \boxed{1} & \\ \text{ACTNTS} & \begin{bmatrix} \text{PRED katt-rel} \\ \text{ACT0 } \boxed{1} \end{bmatrix} \end{bmatrix}$ b. $\begin{bmatrix} \text{HEAD noun} & \\ \text{INDX } \boxed{1} & \\ \text{ACTNTS} & \begin{bmatrix} \text{PRED mus-rel} \\ \text{ACT0 } \boxed{1} \end{bmatrix} \end{bmatrix}$

The process which combines the verb with its object can be represented as in (39a), indicating that the NP which combines with the verb inside of the VP is registered as its OBJ; and similarly, the NP combining with that VP is registered as the SUBJ of the verb, as indicated in (39b).

(39) a. **Combining V and OBJ ('VP' as dominating node):**



b. **Combining SUBJ and VP ('S' as dominating node):**



These illustrations are formally a bit mixed, in combining AVM notation with tree notation. We can use a 'clean' AVM notation using an attribute like 'DTR' employed in connection with lexical rules; since there are two daughters in the structures in question, the attributes are 'HEAD-DTR' and 'NONHEAD-DTR'. The rules (in maximally general form) underlying the processes in (39a,b) can then be stated as AVMs declaring the complex sign structures in question as being valid structures:

(40) a. Combining V and OBJ ('VP' as dominating node – cf. (39a)):

$$
\begin{bmatrix}
\text{HEAD} & \boxed{5}\ \text{verb} \\
\text{GF} & \boxed{1}\begin{bmatrix} \text{SUBJ} & \boxed{3} \\ \text{OBJ} & \boxed{4} \end{bmatrix} \\
\text{SPR} & \left\langle \boxed{3} \right\rangle \\
\text{COMPS} & \langle\ \rangle \\
\text{INDX} & \boxed{0} \\
\text{ACTNTS} & \boxed{2} \\
HEAD-DTR & \begin{bmatrix} HEAD & \boxed{5} \\ GF & \boxed{1} \\ SPR & \left\langle \boxed{3} \right\rangle \\ COMPS & \left\langle \boxed{4} \right\rangle \\ INDX & \boxed{0} \\ ACTNTS & \boxed{2} \end{bmatrix} \\
NONHEAD-DTR & \boxed{4}\begin{bmatrix} HEAD\ noun \end{bmatrix}
\end{bmatrix}
$$

b. Combining SUBJ and VP ('S' as dominating node – cf. (39b)):

$$
\begin{bmatrix}
\text{HEAD} & \boxed{5}\ \text{verb} \\
\text{GF} & \boxed{1}\begin{bmatrix} \text{SUBJ} & \boxed{3} \end{bmatrix} \\
\text{SPR} & \langle\ \rangle \\
\text{COMPS} & \langle\ \rangle \\
\text{INDX} & \boxed{0} \\
\text{ACTNTS} & \boxed{2} \\
HEAD-DTR & \begin{bmatrix} HEAD & \boxed{5} \\ GF & \boxed{1} \\ SPR & \left\langle \boxed{3} \right\rangle \\ COMPS & \langle\ \rangle \\ INDX & \boxed{0} \\ ACTNTS & \boxed{2} \end{bmatrix} \\
NONHEAD-DTR & \boxed{3}\begin{bmatrix} HEAD\ noun \end{bmatrix}
\end{bmatrix}
$$

In the statements in (40), nothing is said about the linking between GFs and semantic participants – it is assumed that this will be stated in the general description of the verb types involved. All that is ensured in that respect in these statements is that the GF and ACTNTS specifications of mother and daughter are identical. This is an assumption one can make in those cases where one sees the argument structure of the entire construction as mirrored in the valence of the verb, and the combination in question as being one where a valence-bound item is 'filled in'. Notably, the combination then changes the valence list in question – for COMPS in (40a) and SPR

in (40b) -, but the over-all inventory of GFs, and the general semantic structure, reflected in the attributes GF and ACTNTS, respectively, remain constant. Not all combinations in grammar are like that – those presently considered are among the main cases.
The information given in (40) might also well be presented in tree form, analogously to (39). In that respect, (39) is just more explicit, and of course renders a specific construction, whereas (40) are general statements. The only factor that the conventional phrase structure trees of Generative Grammar (as opposed to Dependency Grammar) do represent and the AVMs in (40) not, is the *linear ordering* between the daughters. (Technically, this can be added also in an AVM; but (40) doesn't show it.)

How, more generally speaking, does a phrase structure tree in Generative Grammar relate to a statement like either of those in (40)? Both can be used either for the representation of a concrete string, or as a general licensor of strings. The main difference is this: a Generative Grammar tree represents just a partition of a constituent into subconstituents, giving them names and stating their linear order. A statement like in (40) also reflects a partition of a constituent, *C*, into subconstituents, *D1* and *D2*, but it in addition shows *how the grammatical and semantic content of C is distributed among D1 and D2***.** From the viewpoint of grammar as a compositional mechanism, using smaller content blocks to form larger content blocks, statements like those in (40) thus try to get at the essence of combination between such blocks, and thus at the essence of grammatical combination, whereas phrase structure trees provide a mere inventory of which blocks there are and where they occur. Thus, the present AVM-based approach is the more ambitious one, but also the more complex.

(Note that approaches aiming to show how the grammatical and semantic content of a constituent C is distributed among its subconstituents D1 and D2 do not necessarily have to be based on the AVM formalism – *Montague Grammar* is an example**.**)

Appendix 1

In a typological linguistic approach, it is desirable to have a mechanism or notation by which one can efficiently keep stock of what construction types a given language has, and how this construction inventory relates to inventories of other languages. During the past year colleagues at the Universities of Ghana and Leiden and at NTNU have been developing a system for labeling structures in such a way as to facilitate and enhance comparison of structures within and between languages. In principle the system is universal, but it is also very specific, giving a fairly detailed encoding of language-specific syntactic and semantic features. The project has been developed so far using mainly material from Norwegian and from Ghanaian languages especially Kwa, and from some Bantu languages.

One aspect of the system which makes it relevant in an exposition of AVM notation, is that the labels used in the notation can be systematically related to AVM structures, in such a way that combinations of labels in a labeling string can be reflected in the merger of the individual AVMs into larger AVMs, representing the argument structure of full constructions. This is illustrated below.

The scheme is intended to be theory neutral, although it is informed by such formal grammatical theories as HPSG and LFG, and the computation systems that implement grammars based on them. By keeping the notation aloof of any particular grammatical formalism, it is hoped that it can be found useful across frameworks and for research foci including language documentation, typology, and formal grammar. A wiki page facilitating example annotation and information sharing is currently in development at NTNU, where constructions and annotated example sentences can be viewed and discussed (www.typecraft.org). Thus, an inventory of Norwegian types are found at this site under `research/projects/Verbconstructions`.

The notation for designating construction types consists of strings of letters and hyphens called **templates**, composed by **labels**. The basic structural parts of a template are referred to as **slots**. A slot is filled by one or more labels. In the slot specification, the following conventions are observed:

* Slots are interconnected by '-' (hyphen).
* Distinct items inside a slot are interconnected by '_' (underline).
* An item label containing neither '-' nor '_' is an uninterrupted string of letters. If it acts as a complex label, the internal composition is indicated by alternation between small and capital letters (however, no labels are distinguished solely in terms of CAP vs. not).

Constructions with a Verb as head have a template structure with maximum *seven* slots, with the following types of content:

Slot 1: A label for Part of Speech of the head, and – connected by underline – the category of possible formatives marked on the head. The formatives may be realized as affixes, tones, stem formation (as in Semitic), vowel change, reduplication, and more – the realization mode as such is not displayed, only the category expressed.
Slot 2: A label for valency specification - like `intr`, `tr`, `ditr`, and varieties thereof.
Slot 3: Zero or more labels for specification of syntactic constituents.
Slot 4: Zero or more labels for specification of participant roles.
Slot 5: A label for aspect and Aktionsart, written in CAPS.
Slot 6: A label for the situation type expressed by the construction, written in CAPS.

Slot 7: A linking between the slot 6 type and specifications in slots 2-4, of relevance especially for contents whose expression varies crosslinguistically.

Of these, slots 1, 2 and 3 represent well understood areas of specification, and can build on much consensus across frameworks. Slots 4 and 5 are less robust, but have a core of consensus to build on. Slot 6 is still at a highly preliminary state of development.
Slots 1 and 2 are obligatorily filled, the others not. A slot not filled is not displayed: the labels defined for the various slots are distinct, hence no specification can be misread with regard to which slot it concerns. Likewise, no labels are distinguished in terms of CAP vs. not.

The following template exemplifies the notation:

(1) `v-tr-suAg_obAffincrem-COMPLETED_MONODEVMNT`
(E.g. *the boy eats the cake*)

The template reads as follows:
Slot 1: the head is verb;
Slot 2: the syntactic frame is *transitive*;
Slot 4: the thematic roles expressed are 'agent', by the subject, and 'incrementally affected', by the object;
Slot 5: the Aktionsart is characterized as 'completed monotonic development'.
Nothing here occupies slots 3 or 6. Since none of 'suAg', 'obAffincrem' or 'COMPLETED_MONODEVMNT' is a defined slot 3- or 6-specification, there is no ambiguity as to which slot is empty.

The following is an example of a Serial Verb Construction in Akan, where we illustrate the use of slots 6 and 7:

(2)
`sv_aspID-v1tr-v1obIDv2su-v1suAg_v1obEjct-v2tr-v2suTh_v2obEndpt-`
`CONTACTEJECTION-LAUNCHERv1su_MOVERv1ob_TARGETv2ob`

Kofi	**to-o**	**ne**	**nan**	**wɔ-ɔ**	**Kwame**
Kofi	throw-PST	3Poss	leg	pierce-PST	Kwame
N	V	Pron	N	V	N

'Kofi kicked Kwame'

This Template reads as:
- The construction is an SVC;
- aspect (PST) is identical across the verbs;
- the first v-construction has a transitive verb, whose object is (referentially) identical to the subject of the second verb, and its participant roles are 'agent and ejected'; the second v-construction also has a transitive verb, with the participant roles 'theme' and 'endpoint';
- the whole construction expresses the situation-type 'CONTACTEJECTION', that is, 'ejection (by a LAUNCHER) with the ejected (MOVER) obtaining contact with an expressed TARGET'. The second line of the template details how these conceptual

roles are distrubuted on the participants of the actual coonstruction, in a non-isomorphic pattern.

Each template in this system can be seen as a label of a *type* of construction. Thus, the template (1) can serve as a type label for the construction represented in AVM format by (3):

(3)
$$\begin{bmatrix} \text{HEAD verb} \\ \text{GF} \begin{bmatrix} \text{SUBJ} \begin{bmatrix} \text{INDX } \boxed{1} \begin{bmatrix} \text{ROLE agent} \end{bmatrix} \end{bmatrix} \\ \text{OBJ} \begin{bmatrix} \text{INDX } \boxed{2} \begin{bmatrix} \text{ROLE aff-increm} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{INDX ref-index} \\ \text{ASPECT completed} \\ \text{ACTNTS} \begin{bmatrix} \text{ACT1 } \boxed{1} \\ \text{ACT2 } \boxed{2} \end{bmatrix} \\ \text{AKTIONSART monotonic\_development} \end{bmatrix}$$

Just as (1)/(2) is composed of parts each having their independent interpretations, so (3) - or any *sign* AVM – is a complex of various dimensions of specification, and these complex structures can be interrelated. To illustrate, (3) can be seen as a merger of smaller AVMs each associated with one of the constituent labels in (1). The associations are:

(4) v - $\begin{bmatrix} \text{HEAD verb} \end{bmatrix}$

tr - $\begin{bmatrix} \text{GF} \begin{bmatrix} \text{SUBJ} \begin{bmatrix} \text{INDX } \boxed{1} \end{bmatrix} \\ \text{OBJ} \begin{bmatrix} \text{INDX } \boxed{2} \end{bmatrix} \end{bmatrix} \\ \text{ACTNTS} \begin{bmatrix} \text{ACT1 } \boxed{1} \\ \text{ACT2 } \boxed{2} \end{bmatrix} \end{bmatrix}$

suAg - $\begin{bmatrix} \text{GF} \begin{bmatrix} \text{SUBJ} \begin{bmatrix} \text{INDX} \begin{bmatrix} \text{ROLE agent} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$

obAffincrem - $\begin{bmatrix} \text{GF} \begin{bmatrix} \text{OBJ} \begin{bmatrix} \text{INDX} \begin{bmatrix} \text{ROLE aff-increm} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$

COMPLETED_MONODEVMNT - $\begin{bmatrix} \text{ASPECT completed} \\ \text{AKTIONSART monotonic\_development} \end{bmatrix}$

In Hellan and Dakubu 2009, a large set of labels are associated with such AVMs, along with definitions in words spelling out the intended content. For any combination of labels constituting a template, such a merged AVM can be constructed. Below is a list of attributes serving inside of the AVMs used in this inventory; it represents an enlargement relative to the inventory we have worked with so far, but is of the same format. In this list, features in boldface are 'outermost' in a *sign* path, and features in italics are next in the path.

The following general definitions are adopted:

A ***direct syntactic argument*** of a verb is any nominal constituent syntactically directly related to the verb (as subject-of, direct object-of, or indirect object-of), and any clausal constituent with either of these functions. (This *in*cludes expletive subjects and objects, and *ex*cludes clausal constituents in extraposed position; it also excludes any NP or clause governed by a preposition. It also excludes NPs carrying locative case as in Finno-Ugric or Caucasian languages.)

With this notion 'direct syntactic argument', we define three basic valency notions:

intr = intransitive, i.e., with only SUBJECT as direct syntactic argument.

tr = transitive, i.e., with SUBJECT and one OBJECT as direct syntactic arguments.

ditr = ditransitive, i.e., with SUBJECT and two OBJECTs as direct syntactic arguments.

A direct syntactic argument is ***standardly linked*** when it has referential content and serves a semantic argument function relative to the verb. (This *ex*cludes expletive subjects and objects and 'raised' full NPs.)

HEAD	part of speech and other properties associated with the head of a construction
FORMATIVES	list of affixes, tones, stem formation (as in Semitic), reduplication, and other formatives marked on the head constituent
CASE	case (mainly for nouns, pronouns and determiners)
DEF	definiteness (mainly for nouns, pronouns and determiners)
REAL	realization status: dropped, cliticized, normal (mainly for pronouns)
AGR-TARGET	the constituent is targeted by agreement marking on the head of the construction (mainly for nominals)
TAM	Tense/aspect/mood (mainly for verbs)
GF	grammatical function
SUBJ	subject sign
OBJ	object sign; in combination with IOBJ, OBJ is 'direct object', and in combination with OBJ2, OBJ is 'first object'
IOBJ	indirect object, to be used in combination with OBJ
OBJ2	second object, to be used in combination with OBJ
COMP	sentential complement (when not being classified as object)
OBL	oblique, i.e., a PP where the governed NP has a role defined relative to the head V, and it thus is the semantics of that NP, and not the semantics of the PP as a whole, which is of interest
PRESENTED	'presented' NP in a presentational construction
SECPR	secondary predicate
IDNT	complement of an identifying predicate
ADVBL	'adverbial complement', i.e., a PP, Adv or AdvP serving as complement, where it is the semantics of the whole constituent which is of interest
PRTCL	'particle', with aspectual or less tangible impact
VID	'verbid', a VP serving a bit like an OBL
GOV	governee, used in connection with a preposition for its inherent GF (roughly, an abbr. for 'GF | OBJ')
INDX	referential index
ROLE	participant role ('theta-role')
KIND	‘kind’, i.e., situation or individual
CLASS	class, i.e., inherent properties

XACT	'exposed actant': in 'raising' and 'equi' constructions, XACT coincides with the subject of the infinitive, and in non-verbal secondary predicates it coincides with the ACT1 of the predicate.
ACTNTS	'actants', i.e., participants of the situation type expressed by the head of the construction
ACT0	index of the situation type expressed by the construction
ACT1	actant 1
ACT2	actant 2
ACT3	actant 3
ACTobl	actant expressed by the NP complement of an oblique
LOC	locative argument
DIR	directional argument
PRED	predicate (used only with grammatically expressed meanings)
ASPECT	aspect
AKTIONSART	Aktionsart

Values

+/-	
copula	value of HEAD: a subtype of *verb*
drop	value of HEAD | REAL: dropped, in the sense 'pro-drop'
clit	value of HEAD | REAL: cliticized
nomin	value of HEAD | CASE
decl-compl	value of HEAD
yes-no-compl	value of HEAD
wh-compl	value of HEAD
infin-compl	value of HEAD
gerund	value of HEAD | TAM
infinitive	value of HEAD | TAM
irrealis	value of HEAD | TAM
cause-rel	value of ACTNTS| PRED
binary-rel	value of ACTNTS| PRED
part-of	value of ACTNTS| PRED
spatial-coord-of	value of ACTNTS| PRED
concur	value of ACTNTS| PRED
explet	value of INDX: expletive, i.e., referentially void
spatial	value of INDX | CLASS
bodypart	value of INDX | CLASS
sign	value of GF: sign
oriented-obj	value of ACT1 and ACT2: oriented object, a super-type of paths, direction indicators and locomotors (movers)

Examples

Below are a few templates reflecting the operation of ‘lexical rules’, with AVMs displaying the ‘output’ constellation, followed by examples from Citumbuka (spoken in Malawi; examples by courtesy of Ms Jean Chavula, University of Malawi). The affixes reflecting the operations are given in boldface:

(5) `intrRf` = intransitive formed (from transitive) by a reflexive.

$$\begin{bmatrix} \text{GF} & \begin{bmatrix}\text{SUBJ} & \begin{bmatrix}\text{INDX} & \boxed{1}\end{bmatrix}\end{bmatrix} \\ \text{ACTNTS} & \begin{bmatrix}\text{ACT1} & \boxed{1} \\ \text{ACT2} & \boxed{1}\end{bmatrix}\end{bmatrix}$$

(Ex. Citumbuka:

Mwana	*wa-**ji**-timb-a*	*na*	*ndodo*.
1-child	1SM-Refl-beat-FV	prep	7 stick

‘The child beat herself with a stick.’)

(6) `intrRp` = intransitive formed (from transitive) by a reciprocal.

$$\begin{bmatrix} \text{GF} & \begin{bmatrix}\text{SUBJ} & \begin{bmatrix}\text{INDX} & \boxed{1}\end{bmatrix}\end{bmatrix} \\ \text{ACTNTS} & \begin{bmatrix}\text{ACT1} & \boxed{1} \\ \text{ACT2} & \boxed{1}\end{bmatrix}\end{bmatrix}$$

(Ex. Citumbuka:

Ti-ka-pulik-izg-**an**-a

SM-Past-hear-Cs-Recip-FV ‘We listened to each other’)

(7) `trAp-obAobl` = transitive formed (from intransitive oblique) through Applicative, where `obAobl` indicates that the first object is 'promoted' from oblique.

$$\begin{bmatrix} \text{GF} & \begin{bmatrix}\text{SUBJ sign} \begin{bmatrix}\text{INDX} & \boxed{1}\end{bmatrix} \\ \text{OBJ sign} \begin{bmatrix}\text{INDX} & \boxed{2}\end{bmatrix}\end{bmatrix} \\ \text{ACTNTS} & \begin{bmatrix}\text{ACT1} & \boxed{1} \\ \text{ACTobl} & \boxed{2}\end{bmatrix}\end{bmatrix}$$

(Ex. Citumbuka:

Temwani	wa-gon-**er**-a	mphasa	
1Temwani	1SM-sleep-Ap-Fv	9mat	'Temwani has slept on a mat')

(8) `ditrAp-obAobl` = ditransitive construction derived through Applicative, where `obAobl` indicates that the first object is 'promoted' from oblique through Applicative.

$$\begin{bmatrix} \text{GF} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix}\text{INDX } \boxed{1}\end{bmatrix} \\ \text{OBJ} & \begin{bmatrix}\text{INDX } \boxed{2}\end{bmatrix} \\ \text{OBJ2} & \begin{bmatrix}\text{INDX } \boxed{3}\end{bmatrix} \end{bmatrix} \\ \text{ACTNTS} & \begin{bmatrix} \text{ACT1 } \boxed{1} \\ \text{ACT2 } \boxed{3} \\ \text{ACTobl } \boxed{2} \end{bmatrix} \end{bmatrix}$$

(Ex.Citumbuka:
Mary wa-ka-mu-phik-**ir-**a Temwa nchunga
M. 1SM-pst-1OM-cook-Appl-fV Temwa beans
'Mary cooked beans for Temwa')

(9) `ditrCs-obCsu_ob2Cob` = ditransitive construction derived through Causativization, where `obCsu_ob2Cob` indicates that the first object is derived from underlying subject, and the second object is derived from underlying object. Semantically speaking, the construction expresses a person-causer, a two-actant caused event, and the ACT1 of the caused event is expressed as first object and ACT2 of the caused event as second object. In the AVM, the representation of the derivational history is short-cut to show only the resulting GF constellation linked to the resulting semantics.

$$\begin{bmatrix} \text{GF} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix}\text{INDX } \boxed{1}\end{bmatrix} \\ \text{OBJ} & \begin{bmatrix}\text{INDX } \boxed{2}\end{bmatrix} \\ \text{OBJ2} & \begin{bmatrix}\text{INDX } \boxed{3}\end{bmatrix} \end{bmatrix} \\ \text{ACTNTS} & \begin{bmatrix} \text{PRED cause} & \\ & \\ \text{ACT1 } \boxed{1} & \\ \text{ACT2} & \begin{bmatrix} \text{ACT1 } \boxed{2} \\ \text{ACT2 } \boxed{3} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

(Ex.Citumbuka:
Mary wa-ka-mu-phik-**isk**-a Tumbikani nchunga
Mary1SM-pst-1OM-cook-Caus-fV Tumbikani beans
'Mary made Tumbikani cook beans')

(10) `ditrOblApCs-oblCsu_obAobl_ob2Cob` = ditransitive-plus-oblique construction derived through Causativization followed by Applicative, where `oblCsu` indicates that the oblique is derived by Causativization from underlying subject, `obAobl` that the first object is 'promoted' by Applicative from underlying oblique, and `ob2Cob` that the second object is derived by Causativization from underlying object. Semantically speaking, the construction expresses a causation with a person-causer, a three-actant caused event, and the ACT1 of the caused event (the causee) expressed as oblique and the ACT2 of the caused event as second object, whereas ACT3 of the caused event, the ‘applied object’, takes the position of first object.

$$
\begin{bmatrix}
\text{GF} & \begin{bmatrix} \text{SUBJ} & \begin{bmatrix}\text{INDX } \boxed{1}\end{bmatrix} \\ \text{OBJ} & \begin{bmatrix}\text{INDX } \boxed{3}\end{bmatrix} \\ \text{OBJ2} & \begin{bmatrix}\text{INDX } \boxed{2}\end{bmatrix} \\ \text{OBL} & \begin{bmatrix}\text{GOV} \begin{bmatrix}\text{INDX } \boxed{4}\end{bmatrix}\end{bmatrix} \end{bmatrix} \\
\text{ACTNTS} & \begin{bmatrix} \text{PRED cause} & \\ \text{ACT1} & \boxed{1} \\ \text{ACT2} & \begin{bmatrix} \text{ACT1} & \boxed{4} \\ \text{ACT2} & \boxed{2} \\ \text{ACTobl} & \boxed{3} \end{bmatrix} \end{bmatrix}
\end{bmatrix}
$$

(Ex.Citumbuka:

Tumbikani wa-ka-mu-phik-**isk-ir**-a	Temwa	nchunga	kwa Mary
Tumbikani 1SM-pst-1OM-cook-Caus-Appl-fV	Temwa	beans	'to' Mary

'Tumbikani made Mary cook beans for Temwa' or
'Tumbikani had Mary have the beans cooked for Temwa')

References

Bresnan, Joan. 2001. *Lexical Functional Syntax*. Oxford: Blackwell Publishers.

Copestake, Ann. 2002. *Implementing Typed Feature Structure Grammars*. Stanford: CSLI Publications.

Hellan, Lars and Mary Esther Kropp Dakubu. 2009. 'Enumerating Verb Constructions Cross-linguistically'. Ms., NTNU and University of Ghana.

Kroeger, Paul.2004. *Analyzing Syntax*. Cambridge: Cambridge University Press.

Pollard, Carl, and Ivan Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: Chicago University Press.

Sag, Ivan, Thomas Wasow and Emily Bender. 2003. Syntactic Theory. Stanford: CSLI Publications.

Vitale, Anthony J. 1981. *Swahili Syntax*. Dordrecht and Cinnaminson, NJ: Foris Publications.

Solution to AVM for *Hun ble sett komme*
:

$$
\left[\begin{array}{ll}
\text{HEAD } \textit{verb} & \\
\text{GF} & \left[\begin{array}{ll}
\text{SUBJ} & \left[\text{INDX } \boxed{2}\right] \\
\text{SECPR} & \left[\text{INDX } \boxed{3}\right]
\end{array}\right] \\
\text{ACTNTS } \boxed{4} & \\
\text{DTR} & \left[\begin{array}{ll}
\text{HEAD } \textit{verb} & \\
\text{GF} & \left[\begin{array}{ll}
\text{SUBJ} & \left[\text{INDX } \boxed{1}\right] \\
\text{OBJ} & \left[\text{INDX } \boxed{2}\right] \\
\text{SECPR} & \left[\text{INDX } \boxed{3}\right]
\end{array}\right] \\
\text{ACTNTS } \boxed{4} & \left[\begin{array}{ll}
\text{PRED se-rel} & \\
\text{ACT1 } \boxed{1} & \\
\text{ACT2} & \left[\begin{array}{ll}
\text{K} & \left[\begin{array}{l}
\text{PRED komme-rel} \\
\text{ACT0 } \boxed{3} \\
\text{ACT1 } \boxed{2}
\end{array}\right]
\end{array}\right]
\end{array}\right]
\end{array}\right]
\end{array}\right]
$$