

Canonical and noncanonical argument structures - How much underspecified is the lexical entry of a verb?

Dieter Wunderlich (Berlin/Düsseldorf), DGfS-Jahrestagung Köln 25.02.06

The most robust classification of verbs relates to the number of arguments: intransitive verbs have one argument, transitive verbs have two, and ditransitive verbs have three. Another, likewise important property of verbs is argument asymmetry: the arguments of a verb are always strictly ordered. This ordering is often semantically determined, but in some verbs, notably the symmetric ones, it is not. Nevertheless, the arguments of a verb such as ‘meet’ are put into an arbitrary order. Argument ranking is mapped onto linear syntactic order in a so-called neutral context, with the highest argument first. However, whether the more specific morpho-syntactic means of realizing arguments are symmetric or asymmetric, depends on the language type. They are asymmetric in languages of the generalized case type because case is an asymmetric notion, regardless of whether it is spelled out by pronominal affixes, agreement morphemes, morphological case or syntactic position. It is one of the main advantages of this type that lexical asymmetry is preserved in overt case. I only consider this type here.

It is in the spirit of minimalism that lexical entries are maximally underspecified. For a canonical intransitive verb we only need to know that it has one argument. In some languages, a subclass of intransitives (the agentive ones) has to be characterized by the feature +control, alternatively, a different subclass by +affected. Possible are also features such as +telic or +animate. In German, +control regulates the possibility of passive, and +telic the choice of perfect auxiliary (*sein* vs. *haben*). Sporadically, intransitive verbs are marked by a case feature: ERG in some agentive intransitives of Hindi, and ACC in some experiencer intransitives of German.

(1) canonical intransitives: VERB(x), with possible additional features for subclasses

More interesting are transitive verbs. Canonical transitive verbs need no more information than the order of the two arguments, which can mostly be predicted from semantic information (‘Agents outrank all other semantic roles’). To encode the ranking, one feature on just one of the two arguments is sufficient. One can designate either the higher or the lower argument. The higher argument intrinsically bears the feature +lr (‘there is a lower argument’, which is short for ‘there is another argument which is lower than this one’), while the lower argument intrinsically bears the feature +hr (‘there is a higher argument’). For convenience, +hr is read as underlying ACC, and +lr as underlying ERG, thus demonstrating that these two hierarchy features constitute abstract case.

(2) canonical transitives: VERB{x, Z_{ACC}} or VERB{x_{ERG}, Z}

Turning to ditransitive verbs, the ordering of arguments is fully determined if either a further feature is added (such as underlying DAT on the medial argument), or two further instances of the features used already are distributed. This latter option has a number of advantages: it reduces the number of features by decomposing DAT into ACC,ERG; and it also allows to account for cross-linguistic variation: either the medial or the lowest argument is most-marked intrinsically.

(3) canonical ditransitives:

direct vs. indirect object: VERB{x, y_{ACC,ERG}, z_{ACC}} or VERB{x_{ERG}, y_{ACC,ERG}, z}

primary vs. secondary object: VERB{x, y_{ACC}, z_{ACC,ACC}}

[ACC,ACC = ‘more deeply embedded’]

In terms of overt case, dative is +hr,+lr, accusative is +hr, ergative is +lr, and nominative is []_{CASE} (i.e. fully underspecified). There is no overt case +hr,+hr, i.e. secondary objects cannot be marked by a specific case: they can be marked by accusative, or they can remain unmarked.

Intransitive verbs often alternate with transitive ones, which add either a higher argument (4a) or a lower one (4b). Hybrids of the former type are found in many languages.

- (4) a. Die Vase zerbrach. Er zerbrach die Vase.
 ‘The vase broke.’ ‘He broke the vase.’
 BREAK(y) BREAK { x_{ERG} , y} entry: BREAK {(x_{ERG}), y}
- b. Er fuhr langsam. Er fuhr einen Audi/das Gepäck/einen Kilometer.
 ‘He drove slowly.’ ‘He drove an Audi/the luggage/one kilometer.’
 DRIVE(x) DRIVE {x, y_{ACC} } entry: DRIVE {x, (y_{ACC})}

(5) Intransitive-transitive hybrids in Basque (Joppen & Wunderlich 1995:143)

- a. Jon igo-tzen da.
 John move.up-IMPF AUX.3N
 ‘John moves up.’
- b. Jon-ek maleta-k igo-tzen ditu.
 John-ERG suitcase-pl take.up-IMPF AUX.3plN.3sE
 ‘John takes the suitcases up.’ entry: MOVE_UP {(x_{ERG}), y}

The verbal predicate is formed from an open proposition by means of λ -abstraction. The λ -abstractors constitute the theta-roles, and they can be associated with further information, which is relevant for argument realization.

- (6) λ -abstraction: $\lambda z \lambda x \text{ VERB } \{x, z_{\text{ACC}}\}$, $\lambda z \lambda y \lambda x \text{ VERB } \{x, y_{\text{ACC,ERG}}, z_{\text{ACC}}\}$, etc.

So far, canonical verbs have been considered, where the abstract case features serve to indicate argument hierarchy. In principle, there are two options for introducing minor classes purely structurally: (i) A canonical verb of higher order is mimicked, (ii) A non-canonical argument pattern is derived.

(i) Mimicking a canonical verb of higher order. Zero-place predicates can be turned to intransitive verbs (‘it rains’), and intransitive predicates can be turned to transitive verbs. In German, either an empty higher argument is added, realized by the expletive *es*, or an empty lower argument is added, realized by a reflexive. The different ranking can be made visible by the feature +hr (=ACC) assigned to the empty or the non-empty argument. The structural feature in (7) invites for the reading of being affected, which is typical for experiencers, and since the reflexive in (8) is bound to the subject, again an experiencer reading can arise (the same argument is both affected and non-affected).

- (7) Impersonal transitives in German (8) Inherent reflexives in German
- a. Es ekelte ihn. a. Er schämte sich.
 it disgusted he.ACC he.NOM was.ashamed himself
 ‘He was disgusted.’ ‘He was ashamed.’
- b. $\lambda y \lambda x \text{ BE_DISGUSTED}(y)$ b. $\lambda y \lambda x \text{ BE_ASHAMED}(x)$
 ACC ACC
 ($\rightarrow es$) ($\rightarrow +\text{refl}$)

Some verbs allow both patterns without any shift in meaning, which shows that the patterns do not contribute to compositional meaning. The lexical entry then simply is: $\lambda y \lambda x \text{ BE_DISGUSTED}(y)$, with both orderings of the theta-roles possible.

- (9) Alternation between impersonals and inherent reflexives
- es ekelte ihn er ekelte sich ‘it made him sick’
 es freute ihn er freute sich ‘he was pleased’
 es wunderte ihn er wunderte sich ‘he was surprised’

Even more remarkable is the canonical ditransitive pattern shown in (10b), in which both objects are expletive (compared with the similar recipient construction in (10a)).

(10) Canonical ditransitive reflexives in German

- a. Ich trank mir einen (Rausch) an.
I.NOM drunk I.DAT.REFL INDEF.ACC (drunkenness) at
'I drunk too much.'

$\lambda z \lambda y \lambda x$ [DRINK(x,u) & BECOME POSS(y,z)]

- b. Ich arbeitete mir einen (*N) ab.
I.NOM worked I.DAT.REFL INDEF.ACC off
'I worked too much.'

$\lambda z \lambda y \lambda x$ [WORK(x) & ??] = $\lambda z \lambda y \lambda x_{\text{ERG}}$ WORK-TOO-MUCH(x)

(ii) Derivation of a non-canonical argument pattern. Intransitive verbs can exceptionally be marked by ergative (e.g. in Hindi), or by accusative (German *mich dürstet* 'I am thirsty'). In transitive verbs, either the lower or the higher argument can be marked in addition to the inherent hierarchy (see (11) vs. (12)). Combining the inherent hierarchy feature with the lexical feature, the optimal result in each case is dative.

(11) NOM-DAT verbs in German

- a. Sie folgte ihm.
She.NOM followed he.DAT
'She followed him.'

- b. $\lambda y \lambda x$ FOLLOW {x, y_{ACC}}
ERG
(→ DAT)

(12) DAT-NOM verbs in German

- a. Ihm gefiel der Roman.
he.DAT liked the.NOM novel
'He liked the novel.'

- b. $\lambda y \lambda x$ LIKE {x_{ERG}, y}
ACC
(→ DAT)

The presence of a lexical feature can 'invite' certain readings by comparison with the inherent features of canonical transitive verbs. The lexical feature +lr (= ERG) invites the reading of being active, whereas the lexical feature +hr (=ACC) invites the reading of being affected. Lexical features thus motivate certain semantic groupings of verbs. However, lexical marking often constitutes a historical relic, and one cannot predict the semantic subclass of a verb from the presence of such a feature, and neither the other way around.

In the lexically marked 2-place verbs above ((11) & (12)), one needs one feature for the ranking and another feature for the lexical stipulation, both features taken from the set {+lr (= ERG), +hr (=ACC)}. The situation is similar with ditransitive verbs that are lexical marked; but it can get more complex. In Icelandic, one also finds ACC-ACC and DAT-ACC verbs, as well as NOM-DAT-DAT verbs. To account for these verbs, one can use the same feature set as above, and only in some rare instances the respective negative feature value is needed.

(13) Lexically marked ditransitive verbs:

- a. weil der Arzt den Patienten einem Test unterzog
because the.NOM doctor the.ACC patient a.DAT test exposed
'because the doctor exposed the patient to a test'

- b. Ég skilaði henni peningunum.
I.NOM returned her.DAT the money.DAT
'I returned her back the money.'

For a verb that reverses accusative and dative in the ditransitive pattern, the lexical representation is shown in (14b), while a verb with double dative object needs one feature more (14c). One never finds the highest argument of ditransitives to be lexically marked, probably because it is always an agent or causer.

(14) Lexical entries of canonical and non-canonical ditransitives:

- | | | | | |
|----|----------------------|-------------------|-------------|--|
| | λz | λy | λx | VERB {X, Y _{ACC,ERG} , Z _{ACC} } |
| a. | (→ ACC | DAT | NOM) | |
| b. | ERG
(→ DAT | ACC | NOM) | |
| c. | ERG
(→ DAT | ERG
DAT | NOM) | |

It is of course possible to enrich the lexical entries by assigning additional semantic roles or eventive roles. By and large, these roles are irrelevant for the question of argument realization. However, it seems impossible to go beyond the minimal system of lexical entries sketched above. To handle these lexical entries, a system of universal constraints is necessary, specifically ordered and possibly also parameterized for a particular language. Accordingly, lexical entries are adapted to such a constraint system; it is not possible to determine minimal entries without having some system in mind that calculates how a lexical feature affects overt case. (Wunderlich 2003)

(15) MAX(lexF): Lexical features are realized overtly.

Can be violated, e.g. in the Icelandic example (16), which contains a dative object verb.

(16) Control in German vs. Icelandic

- | | | | | | |
|----|-------|---------------|-----------|------------|-----------------------|
| a. | *Ich | hoffe | geholfen | zu werden. | |
| | I.NOM | hope | helped to | AUX | |
| b. | Ég | vonast til a□ | ver□a | hjálp□. | ‘I hope to be helped’ |
| | I.NOM | hope | for to | AUX | helped |

(17) DEFAULT: Each clausal domain contains the unmarked ‘linker’ (nominative).

Explains why the switch from DAT-NOM to NOM-ACC in (18) is due to the loss of only *one* lexical feature, although two arguments are realized differently.

Is violated in ACC or DAT-ACC verbs.

(18) Alternating DAT-NOM and NOM-ACC in German (with slightly different readings)

- | | | | | |
|----|-------|-----------|---------------------|-----------------|
| a. | Mir | schmeckte | der Braten/ | ?der Thymian. |
| | I.DAT | enjoyed | the.NOM roast meat/ | ?the thyme |
| b. | Ich | schmeckte | den Thymian/ | ?den Braten. |
| | I.NOM | tasted | the.ACC thyme/ | ?the roast meat |

(19) UNIQUENESS: Each linker applies only once in a domain.

Can be parameterized according to each single case. In Japanese, for instance, the ban on DOUBLE-ACC holds stronger than the ban on DOUBLE-DAT.

Is violated in German DOUBLE-ACC verbs such as *lehren* ‘teach’, *fragen* ‘ask’.

As a rule, verbs that are formed by means of a productive process (such as denominal verbs, locative alternation, causativization, cognate or internal object, applicative) are canonical, and the way in which the alternation functions already defines the argument hierarchy. There is nothing in these processes that gives rise to a lexical feature.

In both German and Icelandic one also finds genitive competing with accusative as a lexical marker on verbs. Moreover, one often finds arguments that have to be realized by means of a specific preposition (such as *hoffen auf* ‘hope’, *glauben an* ‘believe’). These cases require language-particular extensions of the set of lexical features in obvious ways, they do not pose any additional theoretical problem.

A further remark: Lexical marking yields *more* specific case patterns than the canonical ones predicted by argument hierarchy. It is also possible that *less* specific case patterns arise, given certain sortal, referential or aspectual conditions. E.g., NOM-NOM occurs with Japanese stative verbs, and also occurs with unspecific objects or with 1st/2nd person subjects. All so-called differential object/ differential subject marking is due to certain markedness constraints (such as *ACC/stative, *ACC/-spec, *ERG/1,2), and thus does not belong to the idiosyncratic information of lexical entries.

Choctaw, a Muskogean language of Oklahoma and Mississippi, shows a rich system of lexical marking. For transitive verbs, one finds the alternating patterns shown in (20).

(20) Alternating patterns of pronominal affixes in Choctaw

a. NOM-ACC and NOM-DAT (Davies 1986: 110,112)

- | | |
|------------------------|--------------------------|
| i. Chi-alikchi-li-tok. | ii. Chim-alikchi-li-tok. |
| 2A-doctor-1N-past | 2D-doctor-1N-past |
| ‘I doctored you.’ | ‘I doctored you.’ |

anoksita ‘admire’, shilli ‘comb’, akammi ‘corral’, mokoffi ‘let loose’, ...

b. NOM-DAT and ACC-DAT (Davies 1986: 128, 121)

- | | |
|-----------------------|------------------------|
| i. Chi-nokkilli-li-h. | ii. Chi-sa-nokkilli-h. |
| 2D-hate-1N-pred | 2D-1A-hate-pred |
| ‘I hate you.’ | ‘I hate you.’ |

komota ‘fear’, nokpalli ‘desire’, noktalha ‘mistrust’, nokowa ‘angry at’, ...

c. NOM-DAT and DAT-ACC (Davies 1986: 112, 86)

- | | |
|------------------------|------------------------|
| i. Chim-ihaksi-li-tok. | ii. Chi-am-ihaksi-tok. |
| 2D-forget-1N-past | 2A-1D-forget-past |
| ‘I forgot you.’ | ‘I forgot you.’ |

lhakoffi ‘miss’, ahchiba ‘tired of’, kania ‘lose’, achokma ‘like’, ...

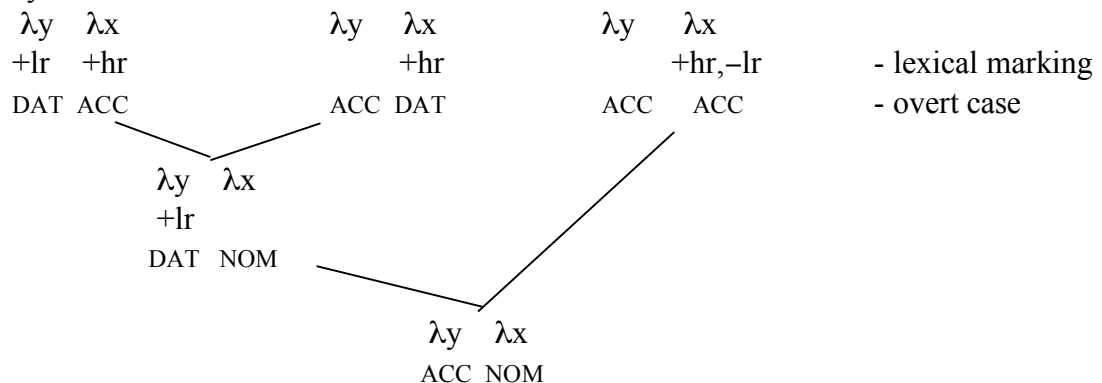
d. NOM-ACC and ACC-ACC (Davies 1986: 65)

- | | |
|--------------------|---------------------|
| i. Chi-banna-li-h. | ii. Chi-sa-banna-h. |
| 2A-want-1N-pred | 2A-1A-want-pred |
| ‘I want you.’ | ‘I want you.’ |

only three verbs: banna ‘want’, yimmi ‘believe’, and anoktoklo ‘doubt’

The tree in (21) illustrates how these alternations function. In the left branches, only one lexical feature is dropped (20a,b), whereas in the right branches more feature changes occur (20c,d).

(21) Lexically marked transitive verbs of Choctaw



Interestingly, free-standing nominals of Choctaw ignore these lexical features; they always show the pattern NOM-(OBL). We account for it by parameterizing $MAX_M(\text{lexF})$ for morphology.

- (22) Hattak-at holisso-(ya) im-ihaksi-tok.
 man-NOM book-(OBL) 3D-forget-PAST
 ‘The man forgot the book.’

Summary:

Every account of lexical underspecification highly depends on the assumed theoretical background.

Our account of argument structure in generalized case-based languages relies on the following assumptions:

- a minimal feature set (two features),
- strict argument hierarchy,
- possible empty abstraction over argument variables,
- a minimal set of constraints regulating the mapping between lexical information and morpho-syntactic patterns,
- minimal assignment of features:
 - to determine the argument hierarchy of lexical items (canonical case patterns),
 - to specify the actual function of an argument role (noncanonical case patterns),
 - to specify overt case (in all of its various sorts of appearance).

References

- Davies, William D. (1986) Choctaw verb agreement and universal grammar. Dordrecht: Reidel.
 Wunderlich, Dieter (2003) Optimal case patterns. German and Icelandic compared. In Ellen Brandner & Heike Zinsmeister (eds.) *New perspectives on case theory*. Stanford: CSLI.