

Why is there morphology?

Dieter Wunderlich (Düsseldorf), 23th Annual Meeting of the DGfS, AG 12

1. Introduction

In many (but not all) languages, the computational system of grammar is organized in two domains: words are formed in the morphology, while phrases are formed in the syntax.

What belongs to the morphology and what to the syntax may change during times. There is evidence from many languages that historical shifts took place along the following lines, which may be thought of as potential stages of a cycle.

(1) The morphology-syntax cycle (MS cycle):

- Morphology is reduced; syntactic elements (auxiliaries, particles) take over functions of morphology.
- Syntactic constituents become clitics and finally affixes; morphology is created.
- Affixes fuse, and morphology becomes less transparent.

The assumption of a MS cycle suggests that morphology and syntax are interchangeable computational options, differing only in whether the minimal elements are bound or free.

Many linguists claim that there is only *one* computational system (called syntax), and that morphology *is* syntax, although different types of movement are assumed (head movement vs. XP movement).

An open question is whether all known types of morphological and syntactic constructions participate in the MS cycle. Moreover, why do some languages seem to stop at some stage in the cycle? Why is there so little evidence for more than one cycle? (Is this only because our observation span for human languages is too short?)

2. In all basic aspects, morphology and syntax are alike.

Under the perspective of abstract morpho-syntax, morphology and syntax are indeed very similar:

- They share the same set of categories,
- the same types of interpretations (functor-argument relations; modification),
- the same types of structures (binary branching, head principle, recursivity), and
- they often correspond to each other according to the *mirror principle*.

3. However, morphology differs from syntax structurally.

Morphology is more restricted than syntax.

- No movement of affixes: neither topicalization, nor focusing, nor raising within a word. Often no specific options for the first or the last position of a word.
- No agreement and no binding (anaphoric relations) between constituents of a word.
- Constituents are not overtly marked for their role (as by case in the syntax).
- Strict locality: non-adjacent affixes are blind for each other, except that they all may relate to the same root.
- No scope ambiguities with respect to arguments: all structural arguments are in the scope of a morphological operator such as negation (*un-*) or repetition (*re-*).
- Weak referential interpretation: non-head nouns mostly have no referential capacity (except in the polysynthetic languages).
- Uniqueness: Morphological objects are often blocked if there exist alternative lexical options (**Stehler* vs. *Dieb*).

Moreover, morphology differs from syntax in terms of regularity.

Morphology is more irregular than syntax, and not only because the smaller morphological objects are stored more easily than the bigger syntactic ones.

- The morphological surface often depends on phonological conditions (assimilation, syllable structure, stress assignment, existence of allomorphs). There may exist syllables for the only purpose of stress regularization (Italian *-esk*, Quechua *-ni*).
- The principle of contrast is more relevant for the smaller than the bigger units: Morphological objects are organized in paradigms, which, however, often contain syncretisms, gaps and substitutions.
- There exist morphological taboos that restrict the expression of first and second person in canonical encounters (many indigenous languages of the Americas, of Australia and New Guinea).
- There exist morphology-semantics mismatches, where the morphological ordering does not reflect semantic computation adequately.

4. Consequently, morphology is less effective than syntax.

Hence, avoid morphology.

Considering both the differences in structure and the differences in regularity, morphology is less effective than syntax and might be better avoided.

- There exist languages with nearly no morphology (Vietnamese, Chinese).
- Most Creole languages start out with syntax only.
- Small populations mostly have languages with rich morphology. For to be diffused to a larger area, a language has to be effective in using transparent and simple means.

5. Why, then, is there morphology at all?

- If the assumption of an MS cycle is right, morphological stages of a language are unavoidable. But why should a language go through stages that are less effective? There may be parts of morphology that are less affected by the above-mentioned structural restrictions.
- Some linguistic functions are either exclusively or predominantly realized by morphological means: agreement, case, voice, tense-aspect. It is these parts of morphology that are even more effective than corresponding syntactic means.
- Agreement and case are morphological means in order to establish relations between syntactic constituents. These means thus presuppose a division of labor between morphology and syntax; paradigms serve as the interface between these two domains.

6. Morphology is more differentiated than syntax.

- Typologically, languages with 'rich morphology' show more variation than languages with 'poor morphology'. It is difficult to predict any correlation between rich morphology and the existence of certain functions of morphology. There are languages with rich morphology that totally lack agreement (Malayalam), and likewise languages with rich morphology that totally lack case (Yimas). There are also languages with rich morphology that neither have agreement nor case (Algonquian).

- Some types of argument linking systems are only known from the morphology of small population languages: the inverse type, the active type, the salience type, the case stacking type. A common feature of some of these types is the reference to cognitive scales of salience (such as animacy, specificity, discourse prominence).

7. Hypothesis: Morphology represents a historically elder state in the evolution of languages.

- Catastrophic evolution theories assume that the computational system of language appeared *all at once*. Since there are so many types of argument linking systems incompatible with each other, these theories have serious problems.
- More realistic evolution theories assume that the computational system of language appeared *step by step*. What are the possible steps? One assumption is that the bifurcated MS system is a later product. First, only one computational domain for the concatenation of items was invented. Because of the more restricted structural capacity of morphology, we may identify this system with what we now call morphology.
- How do we get evidence of former steps in the evolution of language? (i) Consider small language populations that have had few contacts for thousands of years. (ii) These languages may show relics of former states.
- Particularly interesting in this respect are polysynthetic languages showing extensive use of morphological means, while syntax is reduced to a marginal (adjunct-like) status.
- Since agreement and case are linguistic functions that presuppose the MS distinction, the relics have to be found especially among those languages in which argument linking exploits cognitive scales rather than case.

8. The inverse type of argument linking is a potential medial step towards modern language types.

The inverse type is known from the Algonquian languages.

- The direct marker encodes that the higher argument is more salient, while the inverse marker encodes that the lower argument is more salient. (person hierarchy: 2 > 1 > 3)
- With two third animate persons, only one person can be unmarked (proximate), the other person must be marked by obviative (as less salient). Obviation also serves as a reference-tracking device in more complex clauses. (proximate > obviative)
- The Algonquian verb has plural suffixes, which for the 3rd person are differentiated of whether it is the more or less salient one. Furthermore, there are proclitics for 1st and 2nd person, but only the more salient person can be marked.
- Syntactic arguments have adjunct status; they may have possessors, obviative or postpositional marking.

Examples from Plains Cree (Wolfart 1981)

- (2) Transitive animate verbs with 1st and 3rd person
- a. Ni -seekih -aa -naan atim. ‘We scare the dog’
 1 -scare -DIRECT -1pl dog
- b. Ni -seekih -iko -naan atim ‘The dog scares us’
 1 -scare -INVERSE -1pl dog
- (3) Transitive animate verbs with two 3rd persons
- a. waapam-eew naapeew siisiip-a ‘The man sees the duck (obv)’
 see-DIRECT man duck-OBV
- b. waapam-ik naapeew siisiip-a ‘The duck (obv) sees the man’
 see-INVERSE man duck-OBV
- c. waapam-eew naapeew-a siisiip ‘The duck sees the man (obv)’
 see-DIRECT man-OBV duck
- d. waapam-ik naapeew-a siisiip ‘The man (obv) sees the duck’
 see-INVERSE man-OBV duck
- (4) Coreference of the possessor with a core argument
 Note the difference between *o-stees-a* ‘his older brother_{obv}’ and *o-stees-iyiw* ‘his_{obv} older brother’. prox(imate) indicates the non-obviative person.
- a. caan waapam-eew o-stees-a w-iik-ihk
 John see-DIRECT [3-brother]-OBV 3-house-at
 ‘John_{prox} saw his_{prox} older brother_{obv} at his_{prox} house’
- a. caan waapam-eew o-stees-a w-iik-iy-ihk
 John see-DIRECT [3-brother]-OBV 3-house-OBV-at
 ‘John saw his older brother_{obv} at his_{obv} house’
- (5) Coreference between dependent and main clause
- a. naapeew atimw-a waapam-eew ee-sipweehte-t
 man dog-OBV see-DIRECT CONJUNCT-leave-3
 ‘The man_{prox} saw the dog as he_{prox} left’
- b. naapeew atimw-a waapam-eew ee-sipweehte-yit
 man dog-OBV see-DIRECT CONJUNCT-leave-3OBV
 ‘The man saw the dog_{obv} as he_{obv} left’

The inverse type of argument linking crucially differs from the case type.

- Case marks the lower argument role (+hr: accusative) or the higher argument role (+lr: ergative), possibly in the context of a more or less salient person. Accusative is preferred for 1st person, while ergative is preferred for 3rd person. (Dixon 1994, Stiebels 2000)

- (6) a. *direct setting* b. *inverse setting*
- | | | | | | |
|-------------|-------------|-----------|-------------|-------------|-----------|
| λy | λx | VERB(x,y) | λy | λx | VERB(x,y) |
| +hr | +lr | | +hr | +lr | |
| 3 | 1 | | 1 | 3 | |
| | (erg) | | acc | erg | |

+hr ‘there is a higher role’

+lr ‘there is a lower role’

- Theme marks the association of the less salient person with the lower argument (+hs/+hr: direct) or with the higher argument (+hs/+lr: inverse); hence, it marks salience in the context of argument role. There are further markers for the more or the less salient person.

+hs ‘there is a more salient argument’

+ls ‘there is a less salient argument’

(7)	a. <i>direct setting</i>	b. <i>inverse setting</i>
	$\lambda y \quad \lambda x \quad \text{VERB}(x,y)$	$\lambda y \quad \lambda x \quad \text{VERB}(x,y)$
	+hr +lr	+hr +lr
	3 1	1 3
	+hs	+hs
	DIRECT	INVERSE
	3obv 3	3 3obv
	+hs	+hs
	DIRECT	INVERSE

The inverse type is only possible in a morphology-based system.

- An inverse type system is computationally more complex than a case system. The inverse system requires to mark relative salience in the context of abstract case, whereas the case system only requires to mark abstract case (possibly in the context of absolute salience features).
- Further disadvantages of the inverse type: it requires salience differences even when they are not obvious; it blocks certain generalizations over similar settings; it can distinguish syntactic arguments only indirectly.
- There is a lot of evidence that inverse systems may have developed into case systems. There is no indication that a case system ever has developed into an inverse system.

Today, many languages are found that show relics of the inverse type. Inverse systems allow for reanalysis in several ways.

- The more salient argument is marked of whether it is the higher argument (direct) or the lower argument (inverse). (Arizona Tewa (Tanoan), Klaiman 1993)

(8)	a. Ne’i k ^w iyó ná:-tay	b. Ne’i k ^w iyó-di wó:-tay
	this woman 2sg.DIR-know	this woman-OBL 2sg.INV-know
	‘You know this woman.’	‘This woman knows you.’

several sets of person prefixes: DIR: +ls/+lr; INV: +ls/+hr

INV might be reanalyzed as accusative, and DIR as ergative.

- The less salient argument of the inverse setting is marked oblique (\rightarrow passive). (All Tanoan languages, Kroskrity 1985) See example (8b) above.
- The inverse type is restricted to 3rd person (Navajo (Athabaskan), Jelinek 1990)

(9)	a. ‘ashkii yi-z-taʃ	b. ‘ashkii bi-z-taʃ
	boy _i 3.DIR-3 _i -kick	boy _i 3.INV-3 _i -kick
	‘He kicked the boy.’	‘The boy kicked him.’

3.DIR: The more salient 3rd person (he) is the higher argument.

3.INV: The more salient 3rd person (he) is the lower argument.

- Certain constructions show gaps for inverse settings. (Straits Salish, Jelinek & Demers 1994)

- (10) a. kwəniŋ-t =lə =sən b. *kwəniŋ-t -oŋəs -əs =lə
 help-TRANS =PAST=1.NOM help-TRANS -loc.ACC -3.ERG=PAST
 ‘I helped him’ ‘He helped me’
- c. kwəniŋ-t -ŋ =lə =sən
 help-TRANS -PASS =PAST =1.NOM
 ‘I was helped’

- Specialized operations have been developed for the realization of inverse settings. (Agent Focus in Tzotzil (Mayan) only applies in inverse 3→3 settings; Aissen 1999). (cp=completive aspect, fc=final clitic)

- (11) a. Vo'on i-j-kolta. b. Li antz i-kolta-on li petul-e.
 1sg CP-1.ERG-help the woman CP-help-AF **the Pedro-FC**
 ‘It was me who helped him.’ ‘It was the woman who helped Pedro.’

- An inverse morpheme can exist redundantly, even if the language has morphological case. (Jyarong (Tibeto-Burman), DeLancey 1981)

- (12) a. nga-ma nang hetho-e b. nang-ma nga hetho-h-ang
 1-ERG 2 teach-1pl 2-ERG 1 teach-INV-1sg
 ‘We will teach you.’ ‘You will teach me.’
- c. nga-ma ate hetho-ang d. ate-ma nga-nang hetho-h-ang
 1-ERG 3 teach-1sg 3-ERG 1-ACC teach-INV-1sg
 ‘I will teach him.’ ‘He will teach me.’

- In the family of Tanoan languages, some members still have an inverse morpheme and polysynthetic properties such as noun incorporation (Tiwa, Towa), while Tewa lacks such a morpheme, but instead has developed a further class of salience person markers.

- Switch-references between main and dependent clause. (Many languages)

Possible reinterpretations:

- a. inverse voice → inactive marking
 → inverse salience → accusative case
 → passive voice
- b. direct voice → active marking
 → direct salience → ergative case
- c. obviative → switch-reference
 → agreement

9. Other important phenomena in the MS distinction

- Typical phenomena that indicate a **rigid morphology-syntax division** are morphological islands, where no non-head part of a morphological object can interact with other syntactic constituents (by anaphoric relation, or extraction).
- Typical phenomena that indicate **erosion of the morphology-syntax division** are raisings of morphological elements into other syntactic constituents (such as preverb and object agreement raising in Hungarian, object clitic raising in Romance, possessor clitic raising in Bulgarian).

10. Why is there morphology? Three answers have been put forward:

- ◆ Morphology evolves from the process of fusion (MS cycle).
- ◆ There are effective linguistic means (such as agreement and case) that establish some balance between morphology and syntax.
 - ‘Early’ syntax lacks these means: it builds up complex NPs, as well as complex verbs and syntactic domains of verbs (independent, dependent). There must be a further step involving argument linking: How are NPs related to complex verbs as their arguments?
- ◆ Rich morphology was a necessary step in the evolution of syntax. Morphology allows for various types of argument linking, so that the most effective type (case linking) could have been invented. Many relics of the inverse type suggest how the evolution may have taken place.
 - Inverse type: specifies the relative salience of the higher/the lower argument
 - Salience type: specifies the more salient argument of whether it is the higher/the lower argument
 - Split type: specifies the higher argument (ergative) only if it is less salient than the lower argument; specifies the lower argument (accusative) only if it is more salient than the higher one.
 - Generalized case type: specifies the higher/the lower argument independent of salience.

References

- Aissen, Judith. 1999a. Markedness and subject choice in Optimality Theory. *Natural Language and Linguistic Theory* 17: 673-711.
- Aissen, Judith. 1999b. Agent focus and inverse in Tzotzil. *Language* 75: 451-485.
- Dahlstrom, Amy. 1991. *Plains Cree Morphosyntax*. New York: Garland.
- DeLancey, Scott. 1981. An interpretation of split ergativity and related patterns. *Language* 57: 626-657.
- Dixon, R.M.W. 1994. *Ergativity*. Cambridge University Press.
- Fabri, Ray. 1996. The inverse morphology of Plains Cree (Algonquian). In Geert Booij & Jaap van Marle (eds.) *Yearbook of Morphology 1995*, 17-41. Dordrecht: Kluwer.
- Jelinek, Eloise. 1990. Grammatical relations and coindexing in inverse systems. In K. Dziwirek, P. Farrell & E. Mejias-Bikandi (eds.) *Grammatical relations: A cross-theoretical perspective*, 227-246. Stanford, CA: Stanford Linguistics Association.
- Jelinek, Eloise & Richard A. Demers. 1994. Predicates and pronominal arguments in Straits Salish. *Language* 70: 697-736.
- Klaiman, M.H. 1992. Inverse languages. *Lingua* 88: 227-261.
- Klaiman, M.H. 1993. The relationship of inverse voice and head-marking in Arizona Tewa and other Tanoan languages. *Studies in Language* 17: 343-370.
- Kroskrity, Paul V. 1985. A holistic understanding of Arizona Tewa passives. *Language* 61:306-328.
- Stiebels, Barbara. 2000a. Linker inventories, linking splits and lexical economy. In Barbara Stiebels & Dieter Wunderlich (eds.) *Lexicon in Focus*, 211-245. Berlin: Akademie Verlag.
- Stiebels, Barbara. 2000b. *Typologie des Argumentlinkings: Ökonomie und Expressivität*. Habilitationsschrift, Universität Düsseldorf.
- Wolfart, H. Christoph & Janet F. Carroll. 1981. *Meet Cree. A guide to the language*. Lincoln: Univ. of Nebraska Press.

Anhebung der 'definit-Objekt'-Markierung und des Präverbs im Ungarischen

(13) Indefinit-Objekt vs. definit-Objekt:

- a. András egy könyv-et kér.
András a book-ACC want.INDEF.3SG
'Andreas will ein Buch'
- b. Tamás a könyv-et kér-i.
Tamás the book-ACC want-DEF.3SG
'Thomas will das Buch'

(14) Indefinit-Objekt vs. definit-Objekt wird am höheren Verb markiert (Clause-Union):

- a. Anna könyv-et akar olvas-ni.
Anna book-ACC want.INDEF.3SG read-INF
'Anna will Bücher lesen'
- b. Zugleich auch Präverb-Anhebung:
Anna el akar-ja olvas-ni a könyv-et.
Anna PV want-DEF.3SG read-INF the book-ACC
'Anna will das Buch lesen'
- c. Präverb-Anhebung wird blockiert durch ein anderes Element in Fokusposition:
Anna most akar-ja el-olvas-ni a könyv-et.
Anna now want-DEF.3SG PV-read-INF the book-ACC
'Anna will das Buch jetzt lesen'
- d. Anna el fog-ja akar-ni olvas-ni a könyv-et
Anna PV will-DEF.3SG want-INF read-INF the book-ACC
'Anna wird das Buch lesen wollen'

Klitikanhebung im Italienischen:

- (15) a. Anna vuole presenta-re ella a lui.
Anna want.3SG introduce-INF PRON.3FSG DAT PRON.3SG
- b. Anna vuole presenta-r- glie- lo.
Anna want.3SG introduce-INF- 3SG.IO 3SG.DO
- c. Anna glie- lo vuole presenta-re.
Anna 3SG.IO 3SG.DO want.3SG introduce-INF
'Anna will sie ihm vorstellen'

Possessorklitik-Anhebung im Bulgarischen:

- (16) a. Vidja-x interesna-ta i statija.
saw-1SG interesting-DEF 3FSG.DAT article
- b. Vidja-x i interesna-ta statija.
saw-1SG 3FSG.DAT interesting-DEF article
'Ich sah ihren interessanten Artikel'