

International Morphology Meeting, Budapest, 2022 September 1

Forward-formation and back-formation in a Relational Morphology perspective

MARTIN HASPELMATH

Max Planck Institute for Evolutionary Anthropology (Leipzig)

haspelmath@gmail.com

1. Why should back-formation be a problem?

back-formation: *sculpt-or* → *(to) sculpt*
 televis-ion → *(to) televise*
 tape-record-er → *(to) tape-record*
 pease/pea-s → *(a) pea*
 tamale-s → *(a) tamale*
 etc.

Back-formation is a problem only if

- **forward-formation** is the default (see §3)
- morphosyntax is conceived of in **procedural** terms (see §4)
- it is unclear how the **inventorium** (“lexicon”) relates to **productive** regularities (see §5)

In a Relational Morphology (RM) perspective (§2), there is no problem. But then the next question is:

- why is back-formation so **rare**? (see §7)

2. Relational Morphology (and “constructionist approaches” in general)

In the RM conceptual infrastructure (Jackendoff & Audring 2020a,b),

- **forward-formation** is just one of the possibilities (see §3)
- morphosyntax is conceived of in **declarative** terms (see §4)
- **productivity** is an idiosyncratic property of specific schema slots, not an architectural property of a component of grammar (see §5)

I regard RM not as a “theory”, but as a general **conceptual infrastructure** that contains all the necessary ingredients:

- schemas of various types (semantic, morphosyntactic, ...)
- links of various types
- the **motivating** and the **generative/productive** role of schemas

Most importantly, RM eliminates two traditional stereotypes:

- that rules are “processes”
- that “the lexicon” is separate from the “rule components”

Process metaphors permeate our ways of talking, and sometimes also our ways of thinking, about grammatical patterns. But it has long been known that talking in terms of processes can be problematic:

Hockett’s (1954) alternative was **item-and-arrangement** (a kind of declarative approach).

But **arrangements** require **items**, so they suggest “**morpheme-based**” modes of description – which are also known to be problematic (e.g. Anderson 2015).

RM solves the “morpheme” problem by abandoning the **lexicon vs. grammar** stereotype, and replacing it with the “**Parallel Architecture**”: schemas or constructions of various levels of generality.

3. Forward-formation is just one of the possibilities

3.1. Is forward-formation more “natural”?

• Giving special attention to **back-formation** makes good sense in a context where **forward-formation** is regarded as the default – but why should this be so?

- in everyday life, addition and subtraction are probably equally common

RM can deal with mismatches easily, e.g.

(Jackendoff & Audring 2020b: 6)

- | | |
|---------------------------|--|
| (12) a. <i>assassin</i> : | |
| Semantics: | [PERSON WHO [MURDERS POLITICIAN] ₁₁] ₁₀ |
| Morphosyntax: | N ₁₀ |
| Phonology: | /ə'sæsən/ ₁₀ |
| b. <i>assassinate</i> : | |
| Semantics: | [MURDER POLITICIAN] ₁₁ |
| Morphosyntax: | [V N ₁₀ aff ₁₂] ₅ |
| Phonology: | /ə'sæsən ₁₀ ejt _{12/11} |

But forward-formation also conforms to the **diagrammatic stereotype**:

- adding a formative implies adding a meaning
- new words are created by creating a composite form

- (1) base derivative
 (a) *fair* → *unfair*
 (b) *drink* → *drinkable*

Iacobini (2000: 866):

“The **direction of derivation** can ... be identified with a semantic and morphophonological growth which acts upon a base to form a derived word. The derivative is distinguished from the base because of its greater semantic specificity and more complex morphological structure, which usually reveal themselves in the addition of a morphological element (typically an affix)...”

Štekauer (2015: 340) suggests that forward-formation is the normal case because it is “**natural**”:

According to the theory of natural morphology the most natural process in inflectional and derivational morphology is captured by the principle of diagrammaticity, i.e. *constructional iconicity* (cf. Mayerthaler 1981; Dressler 2005; Dressler, Mayerthaler, Panagl and Wurzel 1987). This principle requires that a new meaning be represented by a corresponding form. Therefore, suffixation and prefixation processes are most natural, while conversion is not natural, and subtraction processes, including backformation, are counter-natural.

This makes intuitive sense, and the diagrammaticity principle has rarely been questioned (but see Haspelmath 2008).

A related idea:

The Monotonicity Hypothesis (Koontz-Garboden 2009; 2012)

“Word formation operations do not remove operators from lexical semantic representations.”

- (6.1) a. *red* “the state of redness”
 b. *redd-en* “a change into the state of redness”
 c. *redd-en-ed* “a state of redness brought about by a change into the state of redness”

Another author who has attributed formal complexity to diagrammatic iconicity is Givón (1991):

The quantity principle (a) (Givon 1991: 87)

“A larger chunk of information will be given a larger chunk of code.”

e.g.

- (3) a. *act* ⇒ *act-ive* ⇒ *act-iv-ate* ⇒ *act-iv-at-ion*

3.2. The diagrammatic stereotype is wrong

While it may intuitively be evident that “more meaning” is expressed by “more form”, this is actually wrong!

Grammatical (and also lexical) forms serve to **counter expectations on the part of the addressee**, not to “convey meanings”. This can be seen from contrasts like those in (A)-(F), where the asymmetry of coding can go in both directions (Haspelmath 2021).

(A)		INDICATIVE		IMPERATIVE	
	Spanish	<i>canta</i>	‘(he) sings’	que <i>cante</i>	‘let him sing’
		<i>canta-s</i>	‘you sing’	<i>canta!</i>	‘(you) sing!’
(B)		NONCAUSAL		CAUSAL	
	Russian	<i>kipet’</i>	‘boil (intr.)’	<i>kipja-tit’</i>	‘make boil’
		<i>slomat’-sja</i>	‘break (intr.)’	<i>slomat’</i>	‘break (tr.)’
(C)		AGENT (ERG)		PATIENT (ACC)	
	Dyirbal	<i>ngadya</i>	‘I (ERG)’	<i>ngaygu-na</i>	‘me (ACC)’
		<i>yarra-ngu</i>	‘man (ERG)’	<i>yarra</i>	‘man (ACC)’
(D)		UNIPLEX		MULTIPLEX	
	Welsh	<i>cath</i>	‘cat’	<i>cath-od</i>	‘cats’
		<i>moron-en</i>	‘carrot’	<i>moron</i>	‘carrots’
(E)		ATTRIBUTIVE		PREDICATIVE	
	English	<i>happy</i> (children)		(they) are <i>happy</i>	
		<i>play-ing</i> (children)		(they) <i>play</i>	
(F)		UNPOSSESSED		POSSESSED	
	Koyukon	<i>tel</i>	‘socks’	(se-)tel- e’	‘my socks’
		k’e-tee’	‘head’	(se-)tee’	‘my head’

The marking depends on which grammatical meaning is **less expected** (and needs to be countered), not on “greater semantic specificity”, or on the presence of “operators in the semantic representation”.

Givón is quit right to add another formulation of his principle (Givón 1991: 87):

The quantity principle (b)

“Less predictable information will be given a larger chunk of code.”

4. Morphosyntax is conceived of in declarative terms

(Jackendoff & Audring 2020b: 7)

a fundamental point of PA/RM is that schemas are stored in the extended lexicon, right alongside of words. Both consist of pieces of linguistic structure – **stored declarative knowledge** – and both involve interface links that connect their levels, as well as relational links to other lexical items. The consequence for processing is that all principles of lexical activation and lexical access apply to schemas in the same way as they apply to words. This is not possible in traditional accounts, in which the lexicon and the grammar are quite distinct.

Forward-formation vs. lateral formation:

In addition to forward-formation and back-formation, there is also **lateral formation**, as in (1).

- (1) German
- | | | | |
|-------------------|---|-------------------|--------------------------------|
| <i>Belgien</i> | → | <i>belgisch</i> | |
| <i>Pamphylien</i> | → | <i>pamphylich</i> | |
| <i>Oltenien</i> | → | <i>oltenisch</i> | (Oltenia: a region in Romania) |

Is this a combination of subtraction and addition?

No: Grammatical patterns can **generally** be understood as consisting of **constructions** with **taxonomic** and/or **lateral/horizontal** associations

In all constructionist approaches, “derivational rules” are replaced by **declarative schemas**, which require no directionality.

(see also Becker 1993; Bochner 1993; Plag 2003; Booij 2010; Diessel 2019, etc.)

Becker (1993): **cross-formations**

(Becker also talks about “back- and forth-formation”, but not seriously)

Gaeta & Montermini (2022):

“recent research trends in morphology in various theoretical frameworks have shifted the focus from **purely derivational rules** to lexical / derivational **networks or paradigms**. As a consequence, the very role of **directionality** in word-formation (and more generally in linguistics) has been challenged”

- e.g. Plag (2003: 187): “schema-based model” (similar to Haspelmath 2002)
 Booij (2010): “paradigmatic word-formation”
 (cf. Štekauer 2015: §2.4-5)

On the direction of derivation:

The “direction of derivation” in conversion cannot be determined consistently (cf. Grestenberger & Kastner 2022); in many cases, semantic criteria do not yield a clear result.

There are also non-conversion pairs such as the following, where the “semantic direction of derivation” is unclear:

<i>beauty</i> ? ‘state of being beautiful’	–	<i>beauti-ful</i> ? ‘exhibiting beauty’
<i>linguist</i> ? ‘practitioner of linguistics’	–	<i>linguist-ics</i> ? ‘the science of a linguist’
<i>warm</i> ? ‘higher in temperature than average’	–	<i>warm-er (than X)</i> ? ‘higher in temperature (than X)’
Icelandic <i>kvelja</i> ‘torment, make someone suffer’	–	Icelandic <i>kvelja-st</i> ‘suffer’ (? ‘torment oneself?’) (see Anderson 2020)

5. The inventorium vs. productive regularities

Back-formation is a traditional problem because

- **lexicon** and **grammar** are conceived of as different components
- the lexicon is thought to contain **existing words**
- **productivity** is thought of as the default for grammatical rules (so it must be restricted when a rule seems less than fully productive)

a fundamental point of PA/RM is that **schemas are stored in the extended lexicon**, right alongside of words. Both consist of pieces of linguistic structure – stored declarative knowledge – and both involve interface links that connect their levels, as well as relational links to other lexical items. The consequence for processing is that all principles of lexical activation and lexical access apply to schemas in the same way as they apply to words. This is not possible in traditional accounts, in which the **lexicon and the grammar are quite distinct**.

In Relational Morphology (and related conceptualizations),

- **lexicon** and **grammar** are **NOT** different components
- **existing words** and **potential words** do not have different status (potential words could be used at any point)

- **productivity** is an idiosyncratic feature of slots in constructions, not a feature of a particular component

Audring (2022: §4.1):

Should productivity be considered the norm, and is it the task of a theory to explain the restrictions we find? Or is productivity itself the phenomenon that needs explaining?

From a usage-based perspective, the latter approach makes better sense. If grammatical schemas are first and foremost generalizations over words stored in memory, then the productive use of such schemas to form new words is, in fact, an upgrade. What causes this upgrade in the mind of a language user is not as well understood as we would like.

Word-formation is not really different from analogy in this approach (as also in Becker 1990; 1993); compare also Bauer (2022):

<i>exhibition</i>	<i>edition</i>	<i>komme</i>	<i>schwimme</i>
<i>exhibit</i>	<i>edit</i>	<i>kommst</i>	<i>schwimmst</i>
<i>exhibitor</i>	<i>editor</i>	<i>kommt</i>	<i>schwimmt</i>

It is true that „analogy” is typically thought of as affecting individual items, while a “rule” affects multiple or many items, but if **productivity is a separate issue anyway**, then we need not distinguish between analogy and rules.

Three additional points:

(1) The **inventorium** is the set of all forms, constructions and idioms that must be stored in order to use a language – Bloomfield’s (1933: 274) “lexicon (as a list of basic irregularities)”.

(see Haspelmath 2022 for this term)

The inventorium is roughly what Jackendoff & Audring call “extended lexicon” – but I distinguish between the **inventorium** (the set of unpredictable elements of a language) and a speaker’s **mentalicon** (the set of stored elements in a speaker’s “mental lexicon”).

(2) The inventorium contains **no fully regular complex forms**, while each speaker’s mentalicon contains a lot of fully regular complex forms (even many completely regular inflected forms).

(3) This means that there is no difference between **existing words** and **potential words** in the inventorium – and the difference between stored forms and non-stored forms in each speaker’s mentalicon is trivial.

So what is “back-formation” then?

6. Defining back-formation (as a general concept)

Is back-formation a **purely diachronic process** or **can it be a synchronic** word-formation process?

“Once back-formation has occurred, it becomes invisible to speakers; linguistically naïve contemporary speakers have no reason to think, for example, that *peddle* was derived from *peddler*, rather than the other way around.” (Bauer et al. 2013: 20)

– but this also applies to forward-formation!

Marchand (1969), Kiparsky (1982) and others: only **diachronic**
Plag (2003) and others: **also synchronic**

This question could be asked of any unproductive regularity, e.g.

<i>Spain</i>	–	<i>Span-iard</i>	
<i>Cyprus</i>	–	<i>Cypr-iot</i>	purely diachronic?

Gaeta & Montermini (2022):

“It is not clear what we should consider genuine cases of back-formation, by contrast with ‘canonical’ derivation... Is back-formation a well-defined set of phenomena? How to determine its boundaries and content?”

If we define **derivation** as the creation of a regular noncompound word that is not created by inflection (see Haspelmath 2023), then we might simply say that

- forward-formation is derivation in which **an affix is added**
- back-formation is derivation in which **an affix is removed**
- conversion is derivation in which **nothing** is added or removed

However, these definitions would be trivial, and not really useful, because without a distinction between **existing** and **potential** words, regular back-formation includes cases such as the following:

<i>tallness</i>	→	<i>tall</i>
<i>organization</i>	→	<i>organize</i>
<i>child-like</i>	→	<i>child</i>

De facto, it seems, the usual notion of back-formation is defined as

“derivation by affix removal using a rule of **very low productivity**, or **idiosyncratic creation** of a new word by means resembling a rule”

e.g.	<i>surveillance</i>	→	<i>surveille</i>
	<i>tape-recorder</i>	→	<i>tape-record</i>
	<i>enthusiasm</i>	→	<i>enthuse</i>

7. Why is back-formation not more common?

If back-formation is defined as the creation by a speaker of a word that they had not used before, then it is quite common, at least in inflection:

wugs	–	a wug
minions	–	a minion

Typically, inflectional back-formation is not noticed

e.g. *Have you already **mentalized** the novel term inventorium?*

*Speakers **mentalize** new forms more easily if they are at least semi-transparent.*

Unproductive formations are much more readily noticed than productive ones, and as a result, they are often thought to be “existing words”.

Otherwise, back-formation is noticed

- if it semantically special (buttle from butler)
- if it is formally irregular (enthuse from enthusiasm)
- if it results in unusual compounds (babysit from babysitting)
- if it is from loanwords (tamale from tamales)

Thus, we should probably say that back-formation is common, but it is rarely noticed, and back-formed words do not often become part of the inventorium.

References

- Anderson, Stephen. 2020. Semantically subtractive morphology. In Körtvélyessy, Livia & Štekauer, Pavol (eds.), *Complex words: Advances in morphology*, 36–54. Cambridge: Cambridge University Press.
- Anderson, Stephen R. 2015. The morpheme: Its nature and use. In Baerman, Matthew (ed.), *The Oxford handbook of inflection*, 11–33. Oxford: Oxford University Press.
- Audring, Jenny. 2022. Advances in morphological theory: Construction Morphology and Relational Morphology. *Annual Review of Linguistics* 8. 39–58. (doi:<https://doi.org/10.1146/annurev-linguistics-031120-115118>)
- Bauer, Laurie. 2022. Thinking about back-formation (informal workshop contribution).
- Bauer, Laurie & Lieber, Rochelle & Plag, Ingo. 2013. *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.
- Becker, Thomas. 1990. *Analogie und morphologische Theorie*. München: Fink.
- Becker, Thomas. 1993. Back-formation, cross-formation, and ‘bracketing paradoxes’ in paradigmatic morphology. In Booij, Geert & van Marle, Jaap (eds.), *Yearbook of Morphology 1993*, 1–25. Dordrecht: Springer. (DOI: 10.1007/978-94-017-3712-8_1)
- Bloomfield, Leonard. 1933. *Language*. New York: H. Holt and Company.
- Bochner, Harry. 1993. *Simplicity in generative morphology*. Berlin: Mouton de Gruyter.
- Booij, Geert. 2010. *Construction morphology*. Oxford: Oxford University Press.
- Diessel, Holger. 2019. *The grammar network*. Cambridge: Cambridge University Press.

- Gaeta, Livio & Montermini, Fabio. 2022. Back-formation in a new theoretical universe (workshop description, IMM Budapest). (<http://www.nytud.hu/imm20/>)
- Givón, T. 1991. Markedness in grammar: Distributional, communicative and cognitive correlates of syntactic structure. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"* 15(2). 335–370. (doi:10.1075/sl.15.2.05giv)
- Grestenberger, Laura & Kastner, Itamar. 2022. Directionality in cross-categorical derivations. *Glossa: a journal of general linguistics*. Open Library of Humanities 7(1). (doi:10.16995/glossa.8710)
- Haspelmath, Martin. 2002. *Understanding morphology*. London: Arnold. (<https://zenodo.org/record/1236482>)
- Haspelmath, Martin. 2008. Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive linguistics* 19(1). 1–33.
- Haspelmath, Martin. 2021. Explaining grammatical coding asymmetries: Form-frequency correspondences and predictability. *Journal of Linguistics* 57(3). 605–633. (doi:10.1017/S0022226720000535)
- Haspelmath, Martin. 2022. Against lexicalization (ans what to replace it with). London: UCL. (doi:<https://zenodo.org/record/6408756>)
- Haspelmath, Martin. 2023. Inflection and derivation as traditional comparative concepts. *Linguistics (to appear)*.
- Hockett, Charles F. 1954. Two models of grammatical description. *Word* 10(2–3). 210–234.
- Iacobini, Claudio. 2000. (title). In Booij, Geert E. & Lehmann, Christian & Mugdan, Joachim (eds.), *Morphology: An international handbook on inflection and word-formation (Volume 1)*. Berlin: Walter de Gruyter.
- Jackendoff, Ray & Audring, Jenny. 2020a. *The texture of the lexicon: Relational Morphology and the Parallel Architecture*. Oxford: Oxford University Press.
- Jackendoff, Ray & Audring, Jenny. 2020b. Relational Morphology: A cousin of Construction Grammar. *Frontiers in Psychology* 11. (<https://www.frontiersin.org/articles/10.3389/fpsyg.2020.02241>)
- Kiparsky, Paul. 1982. Lexical Morphology and Phonology. In Yang, In-Seok (ed.), *Linguistics in the Morning Calm: Selected Papers from SICOL-1981*, 3–91. Seoul: Hanshin.
- Koontz-Garboden, Andrew. 2009. Anticausativization. *Natural Language & Linguistic Theory* 27. 77–138. (doi:10.1007/s11049-008-9058-9)
- Koontz-Garboden, Andrew. 2012. The Monotonicity Hypothesis. In McNally, Louise & Demonte, Violeta (eds.), *Telicity, change, and state: A cross-categorical view of event structure*, 139–161. Oxford: Oxford University Press.
- Marchand, Hans. 1969. *The categories and types of present-day English word-formation: A synchronic-diachronic approach*. Munich: C.H. Beck.
- Plag, Ingo. 2003. *Word-formation in English*. Cambridge: Cambridge University Press.
- Štekauer, Pavol. 2015. Backformation. In Müller, Peter O. & Ohnheiser, Ingeborg & Olsen, Susan & Rainer, Franz (eds.), *Word-Formation*, 340–352. Berlin: De Gruyter Mouton. (doi:10.1515/9783110246254-016)