

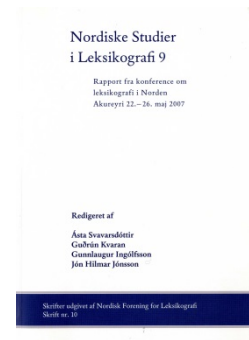
# NORDISKE STUDIER I LEKSIKOGRAFI

Titel: "Movement towards Structure" - Basic Verbs in Learners' Lexicons

Forfatter: Silvie Cinková

Kilde: Nordiska Studier i Lexikografi 9, 2008, s. 77-88  
Rapport fra Konference om leksikografi i Norden, Akureyri 22.-26. maj 2007

URL: <http://ojs.statsbiblioteket.dk/index.php/nsil/issue/archive>



© Nordisk forening for leksikografi

## Betingelser for brug af denne artikel

Denne artikel er omfattet af ophavsretsloven, og der må citeres fra den. Følgende betingelser skal dog være opfyldt:

- Citatet skal være i overensstemmelse med „god skik“
- Der må kun citeres „i det omfang, som betinges af formålet“
- Ophavsmanden til teksten skal krediteres, og kilden skal angives, jf. ovenstående bibliografiske oplysninger.

## Søgbarhed

Artiklerne i de ældre Nordiske studier i leksikografi (1-5) er skannet og OCR-behandlet. OCR står for 'optical character recognition' og kan ved tegngenkendelse konvertere et billede til tekst. Dermed kan man søge i teksten. Imidlertid kan der opstå fejl i tegngenkendelsen, og når man søger på fx navne, skal man være forberedt på at søgningen ikke er 100 % pålidelig.

SILVIE CINKOVÁ

## “Movement towards Structure” Basic Verbs in Learners’ Lexicons

I denna artikel beskrivs en syntaktiskt grundlagd semantisk skiftning hos den svenska verbala pseudosamordningen *hålla på och* som används som progressivetsmarkör. I vissa kontexter får den en närmast motsatt betydelse, nämligen ‘hela tiden’. Sökningar i den svenska PAROLE-korpusen samt i en liten svensk-tjeckisk parallellkorpus visar att den nya betydelsen framför allt är kopplad till negativa imperativyttranden men att det samtidigt finns en spridningstendens till andra syntaktiska strukturer. Undersökningens resultatet jämförs med innehållet i några svenska referensverk och relateras till Paul Hoppers (1987) idé om ”emergent grammar”. Den observerade semantiska skiftningen förklaras med en allmän mekanism som Heine, Claudi och Hünnemeyer (1991) har definierat och kallat ”context-induced reinterpretation“.

### 1 Emergent Grammar

#### 1.1 Introduction

Very common lexical verbs, especially verbs of motion and physical action are in many languages used to denote abstract events and relations. This results in semantic generalization, which in extreme cases makes these verbs (semi-)auxiliary or modal verbs. In numerous other cases they undergo less apparent semantic shifts that are too context-dependent to be noted by grammars, but at the same time they are too vague to be recognized as separate lexical units in lexicons. Drawing on the ideas of *grammaticalization* and *context-induced reinterpretation*, this paper describes an ongoing semantic shift in the Swedish auxiliary verb *hålla på* in its use as progressivity marker, suggesting that cases of context-induced reinterpretation should be systematically identified and observed by lexicographers in general.

#### 1.2 “Movement towards Structure”

The title contains a quotation borrowed from Hopper (1987). Hopper does not conceive grammar as an abstract set of rules filled by words at given communicational occasions, but he rather believes that grammar is what emerges from words, collocations and phrases: “Structure, or regularity, comes out of discourse and is shaped by discourse as much as it shapes discourse in an ongoing pro-

cess” (Hopper 1987). Making no difference between what is traditionally called ‘grammaticalization’ and ‘lexicalization,’ he encourages linguists “to study the whole range of repetition in discourse” and to “seek out those regularities which promise interest as incipient sub-systems”. This in effect requires searching large corpora for formulations that are considered especially useful by many speakers, and therefore they are used (and hence re-assessed and re-shaped) in many different contexts. To fit the immense variety of contexts (i.e. to increase their cross-textual consistency), the formulations extend their collocability, which results in their increasing vagueness.

### 1.3 Context-Induced Reinterpretation

Heine, Claudi and Hünemeyer (1991) as well as Lakoff (1987), Lakoff and Johnson (1980) and others have listed a hierarchy of semantic domains, such as SPACE – TIME etc. Many grammatical structures have developed by a metaphorical transfer from a more concrete semantic domain to a more abstract one. For instance, the English volitional-future marker *to be going to* (as in *Peter is going to write a letter*) is a case of metaphorical use of the verb *go*, which primarily denotes a spatial change, in the temporal domain. However, the process of metaphorical shift is a continuous one. Many uses are ambiguous as to which semantic domain they relate to. To follow Heine, Claudi and Hünemeyer’s example with *to be going to*, the following continuum can be observed (cf. Heine, Claudi and Hünemeyer 1991):

- |   |                               |
|---|-------------------------------|
| (1) <i>Henry is going to town.</i>                          | SPATIAL MOVEMENT              |
| (2) <i>Are you going to the library?</i>                    | SPATIAL MOVEMENT/INTENTION    |
| (3) <i>No, I am going to eat.</i>                           | INTENTION/SPATIAL MOVEMENT?   |
| (4) <i>I am going to do my very best to make you happy.</i> | INTENTION/NO SPATIAL MOVEMENT |
| (5) <i>The rain is going to come.</i>                       | PREDICTION                    |

For such indistinct semantic shifts Heine, Claudi and Hünemeyer (1991) introduce the term *context-induced reinterpretation*. It is described in three idealized stages:

- I. A form F has a focal meaning A. It comes in use with a context-specific side meaning B.
- II. F can be used as B only in unambiguous contexts.
- III. F/B becomes conventionalized and F becomes polysemous.

The case of *to be going to* has reached stage III. Thus a sentence like

- (6) *Peter is going to leave the hospital on Monday.*

has become ambiguous, expressing either Peter’s intention to leave the hospital or the speaker’s prediction about Peter’s future, and other contextual clues must be added to disambiguate it.

## 2 Verbs in the Vocabulary

### 2.1 Swedish ‘Lexical Profile’

Nouns and verbs are believed to be the only universal parts of speech in human languages; i.e. all known human languages have verbs and nouns, though their syntactic profile in general might be very different from that of verbs and nouns in Indoeuropean languages. Verbs as well as nouns are considered to be open parts of speech, which obtain new members during the development of any given language. Frequency analyses of various European languages show an interesting fact: there are significantly fewer verbs than nouns. As Hanks (2003) puts it: “... it seems almost as if all the other parts of speech (verbs and function words) are little more than repetitive glue holding the names in place”. For Swedish, Viberg (1990) notes that there are about 8.5 times fewer verbs than nouns. This implies that verbs, which by their definition denote relations between entities, must have the ability to fit many more different contexts than nouns do. Besides that, almost one half (45.5 %) of Swedish verb occurrences is represented by the 20 most frequent verbs. Almost every second verb usage belongs then to the top twenty! The most frequent verbs (which are not just auxiliary and modal verbs) must evidently have the ability to fit a wide range of different contexts, and some of their usages really “promise interest as incipient sub-systems” (Hopper 1987).

### 2.2 Basic Verbs

Not specifically for Swedish, many of the most frequently used verbs are lexical (full) verbs that are stylistically neutral and denote basic-level categories of events, i.e. events identified and classified with just the level of granularity that is first and intuitively best perceived when acquiring the common knowledge or a foreign language. (For instance, *go* expresses a basic-level event, while *march* would be too specific and *move* too general.) For the purpose of this paper, verbs denoting basic-level categories of events will be called basic verbs.

Basic verbs primarily denote motion/position, like *stå* ‘stand’, *ställa* ‘put vertically’, *komma* ‘come’, *gå* ‘go’, or actions of physical control, such as *få* ‘get’, *ta* ‘take’, *ge* ‘give’ or *hålla* ‘hold, keep’. They often undergo semantic shifts to denote abstract events and states; e.g. they often act as support verbs (*fatta beslut* – lit. ‘grasp a decision’, *ställa en fråga* – lit. ‘put a question’, *ge/få/ha besvär* ‘give/get/have problems’).

Sometimes metaphorical uses of basic verbs gain such a significant cross-textual consistency that they develop into auxiliary verbs. To name but a few, *have*

and its equivalents have in many languages developed into perfect tense markers and the English *go* as well as the Swedish *komma* express different types of future tenses. With Hopper's commitment in mind, any frequent uses of basic verbs enabled by minor meaning generalizations deserve attention as they are unpredictable for foreign learners while they are neither universal enough to be noticed by grammars, nor do they form collocations stable enough to be put on a list of idioms in a lexicon. The following case study is tracing the context-induced reinterpretation (see 1.2) in a well-established grammaticalized use of *hålla* 'hold, keep'.

### 3 The Case of *hålla på och*

#### 3.1 Language Resources

My reasoning on an ongoing semantic shift within the construction *hålla på och* is based on the evidence from two corpora:

- 1) the Swedish PAROLE corpus, which comprises about 20 million tokens and has morphosyntactic tagging. It belongs to the Språkbanken corpora set maintained by the Gothenburg University (PAROLE).
- 2) the Swedish-Czech parallel corpus belonging to the Czech National Corpus, which is being built within a larger project of parallel corpora (INTERCORP). It has no tagging and comprises approx. 1.5 million tokens at the moment. The texts are mostly taken from fiction, with Swedish and Czech being the source as well as the target language respectively.

#### 3.2 *Hålla på* in Lexicons and Grammar

The corpus evidence was first confronted with *Svenskt språkbruk* (SS), a large Swedish monolingual dictionary, which is partly designed as a production dictionary.

The lexicon entry *hålla på* in SS captures:

- 1) the lexical meaning of the construction *hålla på med* – 'to be busy with', 'to pursue', 'to be engaged in' etc. (*Han höll på med deklARATIONEN när det ringde på dörren.*) (...)
- 2) information on the constructions *hålla på att/hålla på och/hålla på med att* as progressivity markers, even with the implicit warning that the last only has marginal use (*även*). (*Han höll på att stryka när telefonen ringde. Jag håller på och diskar. Han höll just på med att deklarerera när det ringde. Landskapet håller på att förändras.*) (...)
- 3) information on the construction *hålla på att* in the sense 'to be on the verge of'.
- 4) "pragmatic phrases/idioms" – *Så här kan du inte hålla på!* (lit. 'this way can you not hold on!') with the explanation "sägs när ngn el.ngra inte kan

fortsätta på samma vis längre” (...) (‘said when someone cannot go on in the same way any more’)

Another large monolingual dictionary, *Norstedts stora svenska ordbok* (NSSO), provides roughly the same information, except that it does not list the idiom *Så här kan du inte hålla på!*.

The largest Swedish grammar, *Svenska Akademiens grammatik* (SAG), makes remarks on the different distribution of the three constructions *hålla på att/hålla på och/hålla på med att*. It notes that *hålla på att/hålla på med att* used to be the only correct form until the late 20th century, though the competing *hålla på och* had occurred centuries ago. Nowadays *hålla på och* is also recognized as standard. Yet *hålla på att* is strongly preferred with inanimate subjects while *hålla på med att* is going out of use. The construction *hålla på och* can only be used as a progressivity marker (except in verbs denoting states) while *hålla på att* can, when used with telic (perfective) verbs (i.e. verbs denoting events with an inherent terminal point), also be used in the sense ‘to be on the verge of’, e.g.

(7) *Jag höll på att svimma* ‘I nearly fainted’.

### 3.3 Analyzing the Corpus Evidence

The corpora basically confirm the statements given by the grammar and the two dictionaries. They have returned a few outstanding concordances, however, for which the reference books give no explanation. It is to be stressed that all the titles are comprehensive and up-to-date, and that they have been willingly chosen as the ultimate choice for every advanced foreign learner. The most striking concordance comes from the parallel corpus:

(8) ... *proto se taky náš farář musel v jednom tahu modlit, aby nebyl tak zlej...*  
lit. ... and that’s why also our priest **had to pray all the time** in order not to be that vicious...

... *därför måste också vår präst hålla på och be stup i ett, så att han inte skulle vara så elak...*

lit. ... and that’s why our priest had to also **hold on and pray all the time** in order not to be that vicious...

The sentence comes from a Czech novel (written in a low-standard spoken Czech), translated into Swedish. The Czech original uses a simple infinitive of the reflexive verb *modlit se* ‘pray’ and a temporal adverbial (*v jednom tahu*). The Swedish translation, however, puts the infinitive *be* ‘pray’ in syntactic coordination with *hålla på*.

The general understanding of the progressive tense suggests that it is to be used under one of the following conditions:

- 1) the speaker focuses on one given moment, making a “snapshot” of the situation, as in:
- (9) *Jag håller på och diskar.*  
I am washing up.
- 2) discourse backgrounding: the speaker reports on an event or a state that provided the background for a more important event, e.g.:
- (10) *Han höll på att stryka när telefonen ringde.*  
He was ironing when the phone rang.

The use of a progressive marker in a sentence like (8) is unexpected. The sentence (8) may express anything but a situation snapshot. Moreover, it combines the progressive tense with a semantically incompatible adverbial that expresses constancy. Neither can the progressive tense be used for discourse backgrounding as the context does not provide any event which could serve as background for the priest’s praying. The translator said that he had used *håller på och* both to indicate the low language standard and as intensification of *all the time*. Neither of these two additional functions of *hålla på och* has been explicitly mentioned by SAG, SS or NSSO. SS, though, contains an idiom that can be paraphrased with “you cannot go on like this any more”. It possibly shares the semantic feature of constancy with the sentence (8).

A closer search in the PAROLE corpus gave more than 40 occurrences of *hålla på och* in contexts that suggested the presence of the constancy feature. Most of them actually fitted the syntactic pattern of what SS had classified as idiom, e.g.:

- (11) *Ni ska inte hålla på och larva er sådär, för jag har ingenting att skämmas för.*  
You shouldn’t keep acting like this because I have nothing to be ashamed of.

However, some occurrences lack the imperative:

- (12) *Men i längden så kan vi ju inte hålla på att bara försvara oss.*  
But in the long run we cannot just keep defending ourselves.

Some occurrences even lack both the imperative and the negation:

- (13) *I princip tyckte hon att det verkade botten att hålla på och knega mellan nio och fem.*  
Actually she meant that it appeared miserable to keep working from nine to five.

The construction *hålla på och* in the constancy sense could appear in questions,

- (14) *Det vet jag inte heller. Varför ska du hålla på och fråga så där?*  
I don't know myself. Why should you keep inquiring like this?

and even one positive declarative sentence like (8) can be found in PAROLE:

- (15) *“Du förnekar det fortfarande. Det är otroligt.”*  
”Det är otroligt att du fortfarande håller på och ältar det. Jag gillade henne aldrig.”  
“You are *still* denying it. It is incredible.”  
“It is incredible that you still keep agonizing over that. I never liked her.”

In accordance with the translator's statement, the “constancy-*hålla på*” mainly occurred in direct speech or in free indirect speech, which both imitate spoken language. With the reservation that both the corpora used are written-language corpora it is likely that the constancy-*hålla på* is a construction typical of or (at the moment) confined to spoken language. Another conclusion that can be drawn from the concordance is that the construction mainly occurs in sentences with low facticity. The example sentences (8) and (15) suggest that high facticity may require a disambiguating adverbial as no positive declarative sentence without a disambiguating adverbial was found either in the parallel corpus or in PAROLE.

How is it that a progressive construction has acquired precisely the opposite meaning? The progressive *hålla på* is the default interpretation of *hålla på* with atelic verbs. It appears in positive as well as in negative declarative clauses, questions etc. in all tenses. Unlike that, the constancy-*hålla på* seems to appear almost exclusively in negations, questions and infinitives. It is the negation that might give the clue for this semantic shift. In a negated progressive sentence, it is not just a single moment of the given event that is negated, but it is the entire event. For instance, the sentence

- (16) *De håller på att bråka.*  
They are fighting.

focuses just one moment of the ongoing event. The same goes for the progressive tense as discourse backgrounder:

- (17) *De höll på att bråka när jag kom.*  
They were fighting when I arrived.

The negation of a sentence predicate, however, says that the entire event does not take place (at the moment of reference) rather than that a single moment of the event does not take place at the moment of reference. This is best perceived in imperative; for instance, by saying:



- (18) *Don't be doing something!*

the speaker necessarily means:

- (19) *Stop doing what you have been doing just long enough to annoy me!*

Implicitly, the event really must have been taking place.

More to say, the relation between progressivity and facticity also works the other way round: when the constancy-*hålla på* is employed in a negative utterance with an event, it suggests that the given event is taking place and should be stopped. This implication has been effectively exploited for rhetorical purposes: at many occasions, especially when reading a text or listening to a speech, the audience naturally has no way of immediately being certain that the mentioned event is taking place at all, mainly when the commented event has been underspecified and expressed by a more or less evaluative paraphrase. By employing the constancy-*hålla på* the speaker adds some kind of asserting modality. A random search in Google returned a few interesting examples of this type:

- (20) *Vi ska inte hålla på att keynesianskt försöka mota konjunkturer.*  
We are not supposed to keep catching conjunctures in the Keynesian fashion.
- (21) *De latinamerikanska, asiatiska och afrikanska staterna ska inte hålla på och blanda sig i USA's och Europas affärer hela tiden!*  
The South American, Asian and African states should not get permanently involved in the USA's and Europe's business!

#### 3.4 Constancy-*hålla på* and Context-Induced Reinterpretation

The close association of the constancy reading of *hålla på* with negation, together with its sparse occurrence in positive utterances, points to the context-induced reinterpretation (see 1.2) as the explanation of this spectacular semantic shift of *hålla på* from progressivity into constancy, which could even have been supported by the semantics of the related construction *hålla på med något* ('to pursue something'). The focal sense "progressivity" came in use in negative utterances with the context-specific side meaning 'stop doing'. As shown above, the constancy sense of *hålla på* is unambiguous in negative sentences, and thus convenient to use. Therefore it has gradually become a truly grammatical means of forming a new semantic shading of event negation. From there the constancy-*hålla på* has been spreading even to contexts in which it is ambiguous. Additional disambiguators must be added to make sure that progressivity, which is the primary sense of *hålla på*, remains inactive. This is achieved with adverbials semantically incompatible with the progressivity sense, such as *all the time*, *permanently*, etc.

### 3.5 Constancy-*hålla på* and Reference Books

The constancy-*hålla på* has neither been mentioned in SAG, nor in NSSO. It can only be traced in the idiom

- (22) *Så här kan du inte hålla på!*  
lit. This way can you not hold on!

captured by SS. Capturing this construction is evidence of corpus-based approach applied to this entry. Unlike in SAG and NSSO, the semantically deviant use of *hålla på* was identified among the enormous amount of the progressive uses. However, the few occurrences of constancy-*hålla på* that were not in negative imperatives were ignored as marginal, and it was described as bound to negative imperative and classified as an idiom (or a *pragmatic phrase*, to stick to SS’s terminology). This is likely to be too strong a statement, though. Idioms are supposed to have their fixed patterns. Representing the constancy-*hålla på* as an idiom implies that only the elements expressed with dummies (indefinite pronouns etc.) can be modified. There was no way for the user to find out that this construction can appear outside the negative imperative.

## 4 Lexicographical Relevance of Context-Induced Reinterpretation

The approved practice of today’s corpus-based lexicography has been to sort out and to display prototypical uses of lexical units found in a large corpus. One reading of a lexical unit in a corpus-based lexicon is often determined by its syntagmatic behavior as well as by its most typical collocates (here it was the negative imperative and the constancy-*hålla på*). Uses of a lexical unit that are supposed to be confined to certain collocations and syntactic patterns are to be presented as phraseological units, while those that show tendencies for syntagmatic variability should be rather presented as separate readings. Heine, Claudi and Hünemeyer’s multilingual evidence for context-induced reinterpretation suggests a standard way to examine how a marginal syntactic pattern of a potential new reading of a lexical unit is associated with the prototypical one. The context-induced reinterpretation helps us to be alert to regular syntactic transformations.

## 5 Conclusion

Large corpora and tools for collocation analysis alone enable us to identify many lexical patterns we would not be aware of when using only introspection, which makes the lexicographical description of lexical units more comprehensive than ever. So far, the approved lexicographical practice in the corpus-based description of lexical units has been to pick out the prototypical uses out of the corpus evidence, sorting others aside as exploitations of the prototypes or simply oddi-

ties. This approach is mostly efficient but it carries the risk of ignoring systematic spreading of lexical patterns out of limited contexts towards general usability. Including the idea of context-induced reinterpretation into the lexicographical reasoning makes it easier to trace structural tendencies in the discourse in the Hopperian sense.

### Acknowledgements:

This work has been supported by the Grant Agency of the Czech Republic (GA-CR 405-06-0589/2006-2008) and by the Czech Ministry of Education (joint project 0021620823/2005-2011).

### Lexicons, Grammars and Corpora:

INTERCORP = Intercorp: Swedish-Czech Parallel Corpus. Joint Research Project 2005-2011, Nr. 0021620823, Czech Ministry of Education. <https://trnka.ff.cuni.cz/ucnk/intercorp/>.

NSSO = Allén, Sture m.fl. (red.). *Norstedts stora svenska ordbok*. Stockholm: Språkdata & Norstedts Förlag 1995.

PAROLE = Språkbanken. <http://spraakbanken.gu.se>.

SAG = Teleman, Ulf m.fl.: *Svenska Akademiens grammatik*. Stockholm: Svenska Akademien 1999.

SS = Clausén, Ulla m.fl. (red.): *Svenskt språkbruk. Ordbok över konstruktioner och fraser*. Stockholm: Svenska Språknämnden & Norstedts Ordbok 2003.

### Bibliography

- Hanks, Patrick 2003: *Norms and Exploitations: Corpus, Computing, and Cognition in Lexical Analysis*. MIT Press, forthcoming. Manuscript, obtained 2003 from the author.
- Heine, Bernd, Ulrike Claudi and Friederike Hünemeyer 1991: *Grammaticalization. A conceptual Framework*. Chicago and London: The University of Chicago Press.
- Hopper, Paul 1987: Emergent Grammar. I: *Berkeley Linguistics Society* 13, 139–157.
- Lakoff, George 1987: *Women, Fire and Dangerous Things. What categories reveal about the mind*. Chicago: Chicago University Press.
- Lakoff, George, and Mark Johnson 1980: *Metaphors We Live By*. Chicago: Chicago University Press.
- Viberg, Åke 1990: Svenskans lexikala profil. *Svenskans beskrivning* 17, 391–408.

Silvie Cinková  
 research assistant, f. 1975  
 Institute of Formal and Applied Linguistics  
 Faculty of Mathematics and Physics  
 Charles University in Prague

Malostranske nam. 25  
CZ-118 00 Praha 1  
cinkova@ufal.mff.cuni.cz

