Semi-lexicality, or, how to grammaticalise a root

Cora Cavirani-Pots
CRISSP/KU Leuven

Contact Information:
CRISSP, KU Leuven
Email: cora.pots@kuleuven.be

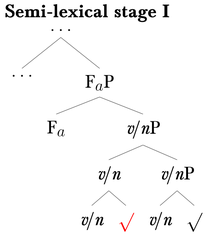
Main proposal: two stages of semi-lexicality

Main question: how do we analyse elements that show both functional and lexical properties, i.e. 'semi-lexical' elements?

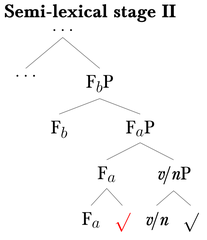
Theoretical assumptions:

1. A lexical item is a featureless root; a functional item is a (bundle of) functional feature(s) (Halle & Marantz 1993; Harley & Noyer 1999; Borer 2005a);
2. Semi-lexicality is the result of a root being inserted in the functional domain of another root (Klockmann 2017; Cavirani-Pots 2020; Cavirani-Pots et al. 2021; see also Song 2019);
3. *v* and *n* a mere categorizers of roots, not introducing any arguments (Kratzer 1996; Lin 2001; Marantz 2005; Bowers 2010; Lohndal 2014; cf. Borer 2005b).

(1) **Semi-lexical stage I**



(2) **Semi-lexical stage II**



→ The red root is the semi-lexically used root.

Semi-lexicality is the result of grammaticalisation:

Lexical > semi-lexical st. I > semi-lexical st. II > functional

Today's empirical domain: Dutch verbs

Lexical verbs select a *te*-complement:

(3) Hij heeft **besloten** ***te werken***.
he has decided to work
'He decided to work.'

Functional verbs never do:

(4) Hij heeft **moeten (*te)** ***werken***.
he has must to work
'He had to work.'

Verbs like *hoeven* 'need' do so optionally:

(5) Hij heeft niet **hoeven (te)** ***werken***.
he has not needed to work
'He didn't need to work.'

→ *Hoeven* adds modality.

Main empirical observation:
Hoeven shows different morphosyntactic behavior compared to both functional and lexical verbs.

Main gist of the analysis
Hoeven is a semi-lexical verb, which is grammaticalising from stage I of semi-lexicality into stage II.

The data

Based on a large-scale questionnaire study (459 speakers):

Low-*te*

(6) Hij zal morgen niet **hoeven**$_1$ ***te werken***$_2$.
he will tomorrow not need to work
'He won't need to work tomorrow.' 46,8%

High-*te*

(7) Hij zal morgen niet ***te*** **hoeven**$_1$ ***werken***$_2$.
he will tomorrow not to need work
'He won't need to work tomorrow.' 19,3%

***Te*-drop**

(8) Hij zal morgen niet **hoeven**$_1$ ***werken***$_2$.
he will tomorrow not need work
'He won't need to work tomorrow.' 22,4%

***Te*-doubling**

(9) Hij zal morgen niet ***te*** **hoeven**$_1$ ***te werken***$_2$.
he will tomorrow not to need to work
'He won't need to work tomorrow.' 7,1%

High degree of intraspeaker variation: 139 speakers allow 2 versions, 96 speakers allow 3 versions, 29 speakers allow 4 versions.

Prerequisites for the analysis

Three theoretical prerequisites:

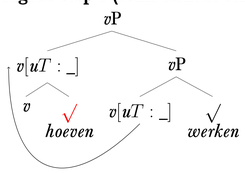
1. Every verbal head enters the derivation with a unvalued [*u*T]-feature; every functional verbal head has a valued [*i*T]-feature corresponding to its semantic interpretation (e.g. *v*: [*u*T:_], Mod: [*i*T:Mod], Wurmbrand 2012);
2. Only *v* can spell out *te*, and only when its [*u*T]-feature has been valued for [irrealis];
3. Verbal feature valuation in Germanic is the result of Reverse Agree (Wurmbrand 2012).

The analysis: *Hoeven* in semi-lexical stage I

Root *hoeven* is used to add modality of (absence of) necessity; Dutch has no formal means yet to add this functional information.

In stage I (cf. (1)), the semi-lexical root is merged with a verbaliser, and then inserted low in the functional structure of the other root.

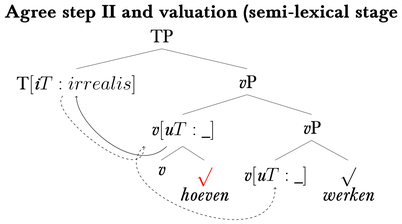
(10) **Agree step I (semi-lexical stage I)**



[*u*T] on the lower *v* probes up and Agrees with [*u*T] on the higher *v*.

No valuation is possible, but a feature link is established (cf. Pesetsky & Torrego 2007; Haegeman & Lohndal 2010).

(11) **Agree step II and valuation (semi-lexical stage I)**



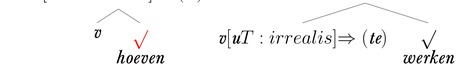
T is Merged, and comes with an [*i*T]-feature valued for [irrealis] (due to modal *zal* 'will' in V2 (cf. (6-9)).

[*u*T] on the higher *v* probes up and Agrees with [*i*T] on T.

Since the two [*u*T]'s in the structure Agreed before and formed a feature chain, they both get valuated for [irrealis].

Based on both [*u*T]-features present in the structure being valuated for [irrealis], we expect *te* to be spelled out twice: i.e. ***te*-doubling**.

(12) **Spell out (semi-lexical stage I)**



→ Why is *te*-doubling not the only option (and in fact the least common option)?

Proposal: spelling out only one of the valued [*u*T]-features of the feature chain of both *v*'s suffices at PF (and is fact preferred).

Given that both the higher *v* position and the lower *v* position can be used to spell out [*u*T:irrealis], we find both **high-*te*** and **low-*te***.

Furthermore, since the choice for spell out is arbitrary, we expect a high degree of intraspeaker variation, which is indeed the case.

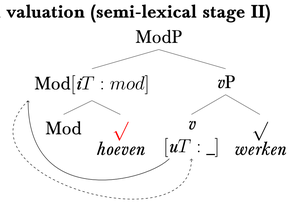
→ What about *te*-drop?

The analysis: *Hoeven* in semi-lexical stage II

Van de Velde (2017) shows that over the last 50 years, *hoeven* is rapidly occurring more without *te* than with *te*.
V → *Hoeven* is grammaticalising into stage II of semi-lexicality.

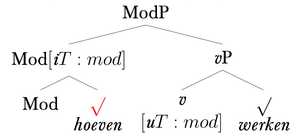
In stage II (cf. (2)), *hoeven* is merged with a functional head Mod, and then inserted in the functional projection.

(13) **Agree and valuation (semi-lexical stage II)**



The [*u*T]-feature on *v* probes up, and finds the [*i*T:Mod]-feature on Mod: the [*u*T]-feature on *v* gets valued for *Mod*.

(14) **Spell out (semi-lexical stage II)**



[*u*T:Mod] on *v* cannot be spelled out as *te*: *werken* 'work' gets spelled out as a bare infinitive.

I.e. in the second stage of semi-lexicality *hoeven* uniformly shows ***te*-drop**.

Conclusion

High degrees of morphosyntactic variation and optionality of semi-lexical items can be accounted for by assuming two stages of semi-lexicality, with different underlying syntactic structures.

Outlook

In Cavirani-Pots (2020), I extend this analysis to the semi-lexical use of *zitten* 'sit'. This verb adds aspectual information (progressive/durative aspect) and shows a high degree of variation in the presence and position of *te*:

(15) Hij zal wel weer **(*te*) zitten**$_1$ **(*te*) *werken***$_2$.
he will aff again to sit to work
'He is probably working again.'

→ The variation regarding *te* can be explained if we assume *zitten* is grammaticalising from stage I of semi-lexicality into stage II.

Selected references

•**Cavirani-Pots**, C. 2020. *Roots in Progress. Semi-lexicality in the Dutch and Afrikaans verbal domain.* Amsterdam: LOT Dissertation Series. •**Cavirani-Pots**, C., **De Belder**, M. and **Klockmann**, H. 2021. Semi-lexicality: Syntax or lexicon? Talk to be presented at GLOW 44 (April 15 2021). •**Klockmann**, H. 2017. *The design of semi-lexicality. Evidence from Case and Agreement in the nominal domain.* Utrecht: LOT Dissertation Series. •**Wurmbrand**, S. 2012. Parasitic participles. Evidence for the theory of verb clusters. *Taal en Tongval* 64: 129–156.

Acknowledgments

This research has been facilitated by KU Leuven Internal Fonds. I'm grateful to the CRISSP members, Susi Wurmbrand, Theresa Biberauer, Sjef Barbiers and Liesbeth Augustinus for comments on (earlier versions of) this work.