Roots in progress

A case study on Dutch semi-lexical verbs

Cora Cavirani-Pots
CRISSP/KU Leuven

Morpho-Syn Syndicate meeting
Penn State University, 10 October 2022

Outline

Introduction: semi-lexicality

- In Dutch, posture verbs like *zitten* 'to sit' can be used to express durative or progressive aspect

Introduction: semi-lexicality

- In Dutch, posture verbs like *zitten* 'to sit' can be used to express durative or progressive aspect

(1) Ik heb **zitten** te lezen.
I have sit to read.
'I have been reading.'

Introduction: semi-lexicality

- The use of these verbs is *semi-lexical*, as they show functional and lexical behaviour at the same time

Introduction: semi-lexicality

- *Functional*: being able to express aspect or modality

Introduction: semi-lexicality

- *Functional*: being able to express aspect or modality
- *Lexical*: being able to select a specific type of complement

Introduction: semi-lexicality

Dutch lexical verbs can select a *te*-complement:

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

Introduction: semi-lexicality

Dutch functional verbs never select a *te*-complement:

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

Introduction: semi-lexicality

Dutch semi-lexical verbs like *zitten* 'to sit' optionally select a *te*-complement:

(4) Hij heeft **zitten** (***te***) ***werken***.
He has sit to work
'He has been working.'

Introduction: semi-lexicality

These verbs furthermore show a high degree of morphosyntactic optionality:

(5) …dat hij heeft moeten **zitten** ***te*** **werken**.
…dat he has must sit to work

(6) …dat hij heeft moeten **zitten** **werken**.
…dat he has must sit work

(7) …dat hij heeft moeten ***te*** **zitten** **werken**.
…dat he has must to sit work
'…that he must have been working.'

Introduction: semi-lexicality

These verbs furthermore show a high degree of morphosyntactic optionality:

(5) …dat hij heeft moeten **zitten** ***te*** **werken**.
…dat he has must sit to work

(6) …dat hij heeft moeten **zitten** **werken**.
…dat he has must sit work

(7) …dat hij heeft moeten ***te*** **zitten** **werken**.
…dat he has must to sit work
'…that he must have been working.'

- This is completely ungrammatical with fully lexical or fully functional verbs

Introduction: main research questions

1. How do we formally analyse elements that show both functional and lexical properties?

Introduction: main research questions

1. How do we formally analyse elements that show both functional and lexical properties?
2. How can we account for the high degree of morphosyntactic optionality displayed by semi-lexical elements?

Introduction: the main proposal in a nutshell

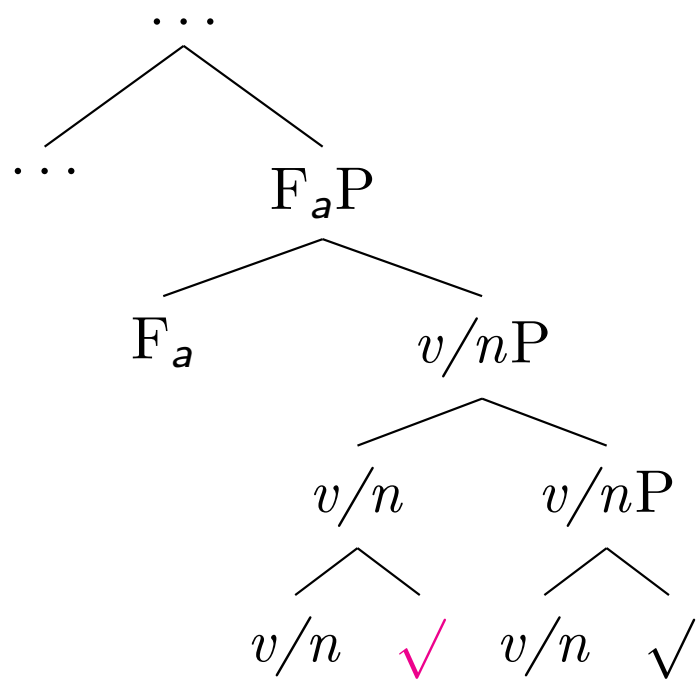
- Semi-lexicality is the result of a root being inserted in the functional domain of another root (Klockmann 2017; Cavirani-Pots 2020; Cavirani-Pots et al. 2021; cf. Song 2019)

Introduction: the main proposal in a nutshell

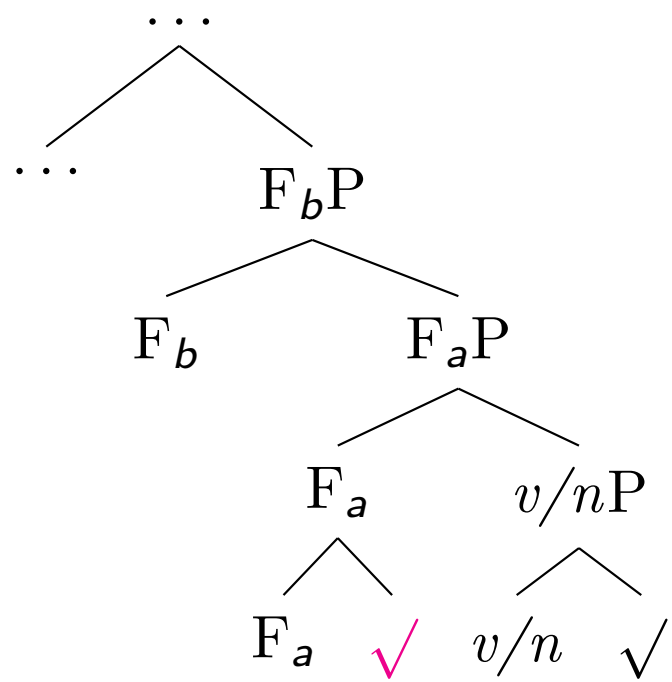
- Semi-lexicality is the result of a root being inserted in the functional domain of another root (Klockmann 2017; Cavirani-Pots 2020; Cavirani-Pots et al. 2021; cf. Song 2019)
- There are two consecutive stages of semi-lexicality (i.e. semi-lexicality is the result of grammaticalisation)

Introduction: the main proposal in a nutshell

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



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



→ The root in pink is the semi-lexically used root.

Introduction: The empirical domain

- I present a case study on two Dutch semi-lexical verbs, *hoeven* 'to need' and *zitten* 'to sit'

Introduction: The empirical domain

- I present a case study on two Dutch semi-lexical verbs, *hoeven* 'to need' and *zitten* 'to sit'
- Both verbs are semi-lexical

Introduction: The empirical domain

- Like *zitten*, *hoeven* shows both functional and lexical behaviour:

Introduction: The empirical domain

- Like *zitten*, *hoeven* shows both functional and lexical behaviour:
  - *Lexical*: it can select a *te*-complement, but does so optionally

Introduction: The empirical domain

- Like *zitten*, *hoeven* shows both functional and lexical behaviour:
  - *Lexical*: it can select a *te*-complement, but does so optionally
  - *Functional*: it can be used to express modality of (the absence of) necessity

Introduction: The empirical domain

- Like *zitten*, *hoeven* shows both functional and lexical behaviour:
  - *Lexical*: it can select a *te*-complement, but does so optionally
  - *Functional*: it can be used to express modality of (the absence of) necessity
- *Hoeven* also shows a high degree of morphosyntactic variation

Introduction: The empirical domain

- I will claim that:

Introduction: The empirical domain

- I will claim that:
- *hoeven* is on its way from the first stage of semi-lexicality to the second

Introduction: The empirical domain

- I will claim that:
- *hoeven* is on its way from the first stage of semi-lexicality to the second
- *zitten* is uniformly in the first stage

Introduction: The empirical domain

- I will claim that:
- *hoeven* is on its way from the first stage of semi-lexicality to the second
- *zitten* is uniformly in the first stage
- This results in different degrees of morphosyntactic optionality between the two verbs

Methodology: design

Large-scale questionnaire study

- The verbs of interest, *hoeven* 'to need' and *zitten* 'to sit' were embedded in non-finite three verb clusters in standard word order

Methodology: design

Hoeven verb cluster

(10) Koen zal niet [hoeven$_1$ *te* gaan$_2$ voetballen$_3$].
Koen will not need.INF to go.INF play.football.INF.
'Koen won't have to go and play football.'

Methodology: design

Hoeven verb cluster

(10) Koen zal niet [hoeven$_1$ *te* gaan$_2$ voetballen$_3$].
Koen will not need.INF to go.INF play.football.INF.
'Koen won't have to go and play football.'

- V1 *hoeven* selects a *te*-infinitive

Methodology: design

Hoeven verb cluster

(10) Koen zal niet [hoeven$_1$ *te* gaan$_2$ voetballen$_3$].
Koen will not need.INF to go.INF play.football.INF.
'Koen won't have to go and play football.'

- V1 *hoeven* selects a *te*-infinitive
- The second verb in the cluster (V2) is a *te*-infinitive

Methodology: design

Zitten verb cluster

(11) Peter zal lang [moeten$_1$ zitten$_2$ *te* wachten$_3$].
Peter will long must.INF sit.INF to wait.INF.
'Peter will have to wait for a long time.'

Methodology: design

Zitten verb cluster

(11) Peter zal lang [moeten$_1$ zitten$_2$ *te* wachten$_3$].
Peter will long must.INF sit.INF to wait.INF.
'Peter will have to wait for a long time.'

- V2 *zitten* selects a *te*-infinitive

Methodology: design

Zitten verb cluster

(11) Peter zal lang [moeten$_1$ zitten$_2$ *te* wachten$_3$].
Peter will long must.INF sit.INF to wait.INF.
'Peter will have to wait for a long time.'

- V2 *zitten* selects a *te*-infinitive
- The lowest verb in the cluster (V3) is a *te*-infinitive

Methodology: design

Goal of the questionnaire study:

- Test whether *te* can appear in a different position than it should appear in based on the selection requirements of *hoeven* and *zitten*

Methodology: design

Goal of the questionnaire study:

- Test whether *te* can appear in a different position than it should appear in based on the selection requirements of *hoeven* and *zitten*
- Different versions of the two verb clusters were included in the questionnaire:

Methodology: design

Goal of the questionnaire study:

- Test whether *te* can appear in a different position than it should appear in based on the selection requirements of *hoeven* and *zitten*

- Different versions of the two verb clusters were included in the questionnaire:
  - the 'correct' version (meeting the selection requirements)

Methodology: design

Goal of the questionnaire study:

- Test whether *te* can appear in a different position than it should appear in based on the selection requirements of *hoeven* and *zitten*

- Different versions of the two verb clusters were included in the questionnaire:
  - the 'correct' version (meeting the selection requirements)
  - *te* occurs on one of the other verbs of the cluster

Methodology: design

Goal of the questionnaire study:

- Test whether *te* can appear in a different position than it should appear in based on the selection requirements of *hoeven* and *zitten*

- Different versions of the two verb clusters were included in the questionnaire:
  - the 'correct' version (meeting the selection requirements)
  - *te* occurs on one of the other verbs of the cluster
  - *te* is absent

Methodology: design

Goal of the questionnaire study:

- Test whether *te* can appear in a different position than it should appear in based on the selection requirements of *hoeven* and *zitten*

- Different versions of the two verb clusters were included in the questionnaire:
  - the 'correct' version (meeting the selection requirements)
  - *te* occurs on one of the other verbs of the cluster
  - *te* is absent
  - *te* occurs twice

Methodology: design

7 different versions of both cluster types:

1. *te*-V1-V2-V3
2. V1-*te*-V2-V3
3. V1-V2-*te*-V3
4. V1-V2-V3
5. *te*-V1-*te*-V2-V3
6. *te*-V1-V2-*te*-V3
7. V1-*te*-V2-*te*-V3

- 14 test items, 25 filler items, 5 practice items

Methodology: procedure

Task

- Judgment task, using a 5-point Likert scale

Methodology: procedure

Task

- Judgment task, using a 5-point Likert scale
- Online written questionnaire, created in LimeSurvey©

Methodology: procedure

Task

- Judgment task, using a 5-point Likert scale
- Online written questionnaire, created in LimeSurvey©
- Test items presented in randomized order, preceded by a practice round (5 practice items, same order for all participants)

Methodology: participants

Participants

- 531 native Dutch speakers completed the questionnaire, 459 were included for analysis:

Methodology: participants

Participants

- 531 native Dutch speakers completed the questionnaire, 459 were included for analysis:
  - 70 participants were excluded due to them having lived abroad for longer than 10% of their lives

Methodology: participants

Participants

- 531 native Dutch speakers completed the questionnaire, 459 were included for analysis:
  - 70 participants were excluded due to them having lived abroad for longer than 10% of their lives
  - 2 participants were excluded due to inconsistent responses to the filler items

Methodology: participants

Participants

- Mean age: 53 (*SD* 12,5; range: 18-99)

Methodology: participants

Participants

- Mean age: 53 (*SD* 12,5; range: 18-99)
- Place of birth: The Netherlands: 361, Belgium: 95 (other: 3)

Methodology: participants

Figure 1: Distribution of included participants

The results: prerequisites

Data preparation

- The verbs of interest are not used in all regions of the language area

The results: prerequisites

Data preparation

- The verbs of interest are not used in all regions of the language area
- 62 participants do not use *hoeven*

The results: prerequisites

Data preparation

- The verbs of interest are not used in all regions of the language area
- 62 participants do not use *hoeven*
- 64 participants do not use *zitten* semi-lexically

The results: prerequisites

Data preparation

- Ratings of 4 and 5 were interpreted as the given test item being grammatical for the participant

The results: prerequisites

Data preparation

- Ratings of 4 and 5 were interpreted as the given test item being grammatical for the participant
- Ratings of 1, 2 and 3 were interpreted as the given test item being ungrammatical for the participant

The results: prerequisites

Terminology

- When *te* occurs in a higher position than required: **high-*te***

The results: prerequisites

Terminology

- When *te* occurs in a higher position than required: **high-*te***
- When *te* occurs in the 'correct' position below the selecting verb: **low-*te***

The results: prerequisites

Terminology

- When *te* occurs in a higher position than required: **high-*te***
- When *te* occurs in the 'correct' position below the selecting verb: **low-*te***
- When *te* is absent: ***te*-drop**

The results: prerequisites

Terminology

- When *te* occurs in a higher position than required: **high-*te***
- When *te* occurs in the 'correct' position below the selecting verb: **low-*te***
- When *te* is absent: ***te*-drop**
- When *te* occurs twice: ***te*-doubling**

The results: prerequisites

Terminology

(12) ...*te* hoeven gaan voetballen. **high-*te***
...to need go play.football.

(13) ...hoeven *te* gaan voetballen. **low-*te***
...need to go play.football.

(14) ...hoeven gaan voetballen. ***te*-drop**
...need go play.football

(15) ...*te* hoeven *te* gaan voetballen. ***te*-doubling**
...to need to go play.football

The results: weighted frequencies

Phenomenon	*hoeven* cluster	*zitten* cluster
High-*te*	19,3%	7,7%
Low-*te*	51,2%	19,6%
Te-drop	22,4%	71,0%
Te-doubling	7,1%	1,7%

Table 1: Weighted frequencies of all phenomena

The results: optionality

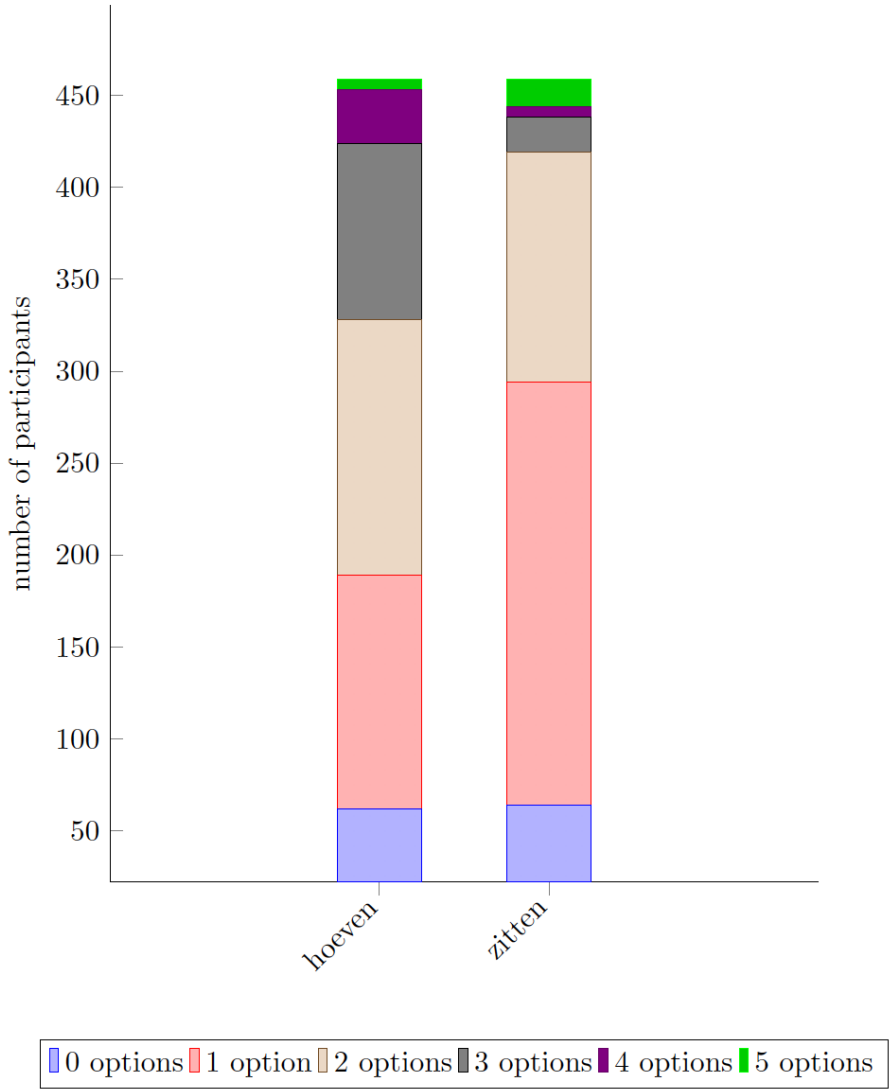


Figure 2: Optionality in both cluster types

The results: summary

- In the *hoeven* cluster, low-*te* is the most frequent, but high-*te* and *te*-drop occur relatively frequently as well

The results: summary

- In the *hoeven* cluster, low-*te* is the most frequent, but high-*te* and *te*-drop occur relatively frequently as well
- In the *zitten* cluster, *te*-drop is by far the most frequent, low-*te* relatively frequent, high-*te* infrequent

The results: summary

- In the *hoeven* cluster, low-*te* is the most frequent, but high-*te* and *te*-drop occur relatively frequently as well
- In the *zitten* cluster, *te*-drop is by far the most frequent, low-*te* relatively frequent, high-*te* infrequent
- In both clusters, *te*-doubling is very rare, virtually non-existent in the *zitten* cluster

The results: summary

- In the *hoeven* cluster, low-*te* is the most frequent, but high-*te* and *te*-drop occur relatively frequently as well
- In the *zitten* cluster, *te*-drop is by far the most frequent, low-*te* relatively frequent, high-*te* infrequent
- In both clusters, *te*-doubling is very rare, virtually non-existent in the *zitten* cluster
- The *hoeven* cluster shows more inter-speaker variation than the *zitten* cluster

The main proposal

Why formalising semi-lexicality

- Semi-lexical items do not behave as 'we expect them to'

The main proposal

Why formalising semi-lexicality

- Semi-lexical items do not behave as 'we expect them to'
- They seem to be neither fully lexical nor fully functional

The main proposal

Why formalising semi-lexicality

- Semi-lexical items do not behave as 'we expect them to'
- They seem to be neither fully lexical nor fully functional
- The 'in between' syntactic behaviour of these elements is problematic for their integration in a theory of linguistic categories

The main proposal

Why formalising semi-lexicality

- Semi-lexical items do not behave as 'we expect them to'
- They seem to be neither fully lexical nor fully functional
- The 'in between' syntactic behaviour of these elements is problematic for their integration in a theory of linguistic categories
- The number of nouns, verbs and adjectives which behave semi-lexically makes it hard to set them aside as exceptions (Ross 1972, Emonds 1985, Van Riemsdijk 1998, Vos 1999, Corver and Van Riemsdijk 2001)

The main proposal

Why formalising semi-lexicality

- Formalising semi-lexicality gives us a tool to formalise the early steps of grammaticalisation

The main proposal

Why formalising semi-lexicality

- Formalising semi-lexicality gives us a tool to formalise the early steps of grammaticalisation
- E.g. the stage of functional item towards affix is much more theoretically fleshed out than the step from lexical to functional

The main proposal

Why formalising semi-lexicality

- Formalising semi-lexicality gives us a tool to formalise the early steps of grammaticalisation
- E.g. the stage of functional item towards affix is much more theoretically fleshed out than the step from lexical to functional
- Formalising semi-lexicality is furthermore important for our theory of the syntax-lexicon interface

The main proposal

Roots and features

- An important question in the semi-lexicality debate is how semi-lexicality should be analysed in terms of roots and syntactic features

The main proposal

Roots and features

- An important question in the semi-lexicality debate is how semi-lexicality should be analysed in terms of roots and syntactic features
- De Belder (2011) argues that semi-lexicality surfaces when a functional item is inserted in a root position (semi-lexicality is syntactic)

The main proposal

Roots and features

- An important question in the semi-lexicality debate is how semi-lexicality should be analysed in terms of roots and syntactic features
- De Belder (2011) argues that semi-lexicality surfaces when a functional item is inserted in a root position (semi-lexicality is syntactic)
- Klockmann (2017) argues that semi-lexicality is the result of a root that bears one or more features in the lexicon (semi-lexicality is lexical)

The main proposal

Theoretical assumptions

- Semi-lexicality is the result of early grammaticalisation (Haider 2001, Hagemijer 2001, Klockmann 2017)

The main proposal

Theoretical assumptions

- Semi-lexicality is the result of early grammaticalisation (Haider 2001, Hagemijer 2001, Klockmann 2017)
- 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)

The main proposal

Theoretical assumptions

- Semi-lexicality is the result of a root being inserted in the functional domain of another root (Klockmann 2017; Cavirani-Pots 2020; see also Song 2019)

The main proposal

Theoretical assumptions

- Semi-lexicality is the result of a root being inserted in the functional domain of another root (Klockmann 2017; Cavirani-Pots 2020; see also Song 2019)
- 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)

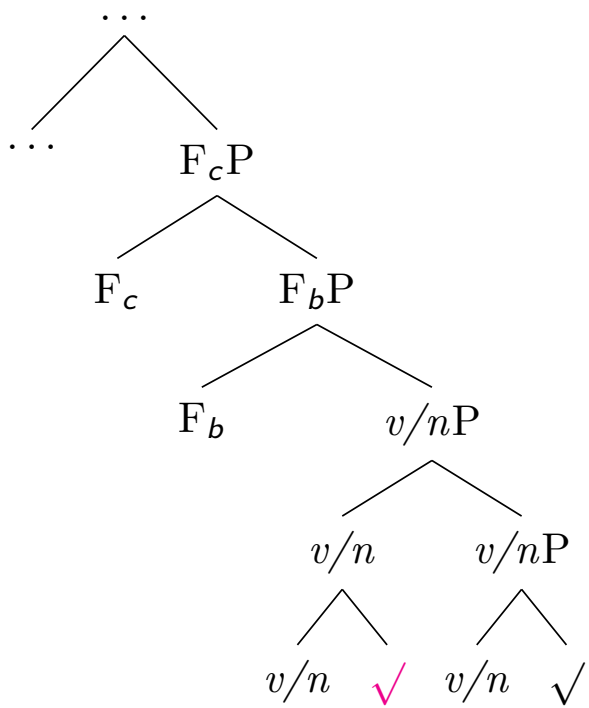
The main proposal

Main proposal

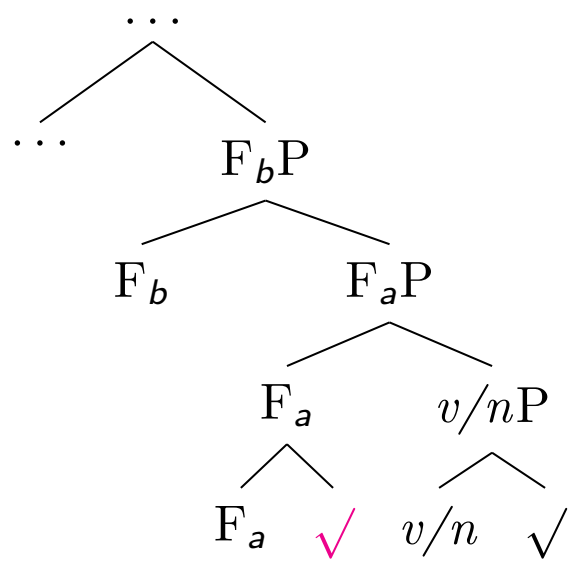
- There are two stages of semi-lexicality, which are early steps on a grammaticalisation path

The main proposal

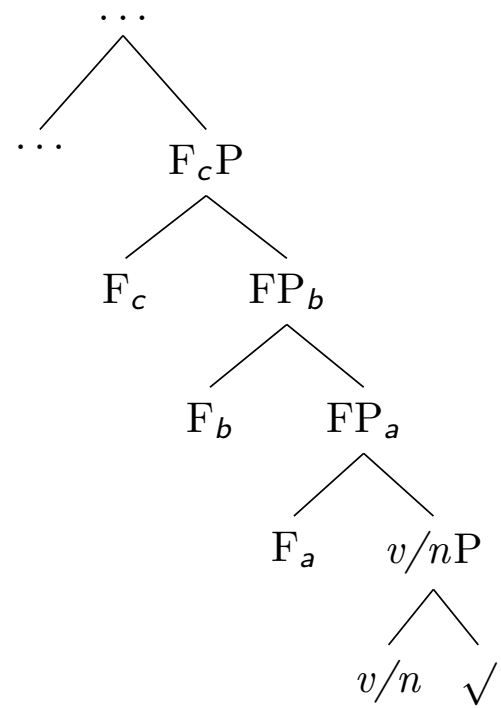
Semi-lex stage I



Semi-lex stage II



Functional stage



→ The root in pink is the semi-lexically used root

The main proposal

- Revision of standard grammaticalisation path (cf. Hopper & Traugott 1993)

(16) $\mathrm{A}_{\textit{lexical}} > \mathrm{B}_{\textit{functional}} / \mathrm{A}_{\textit{lexical}}\ (> \mathrm{B}_{\textit{functional}})$

The main proposal

- Revision of standard grammaticalisation path (cf. Hopper & Traugott 1993)

Stage	Vocabulary items
Stage 0	$A_{lexical}$
Stage I	$A_{lexical}$ + semi-lexical use_{stageI} of $A_{lexical}$
Stage II	$A_{lexical}$ + semi-lexical $use_{stageII}$ of $A_{lexical}$
Stage III	$B_{functional}$ (+ $A_{lexical}$)

The case study: prerequisites

Three theoretical prerequisites, regarding:

1. the featural specification of the verbal domain in Germanic
2. the spell out of *te*
3. the direction of Agree

The case study: prerequisites

The featural specification of the verbal domain (I)

- I assume the featural specification of the verbal domain in Germanic as proposed by Wurmbrand (2012)

The case study: prerequisites

The featural specification of the verbal domain (I)

- I assume the featural specification of the verbal domain in Germanic as proposed by Wurmbrand (2012)
- There is a four-way split among interpretable, uninterpretable, valued and unvalued features (Pesetsky and Torrego 2007, Bošcović 2009)

The case study: prerequisites

The featural specification of the verbal domain (II)

- Every verbal head enters the derivation with an unvalued [uT]-feature

The case study: prerequisites

The featural specification of the verbal domain (II)

- Every verbal head enters the derivation with an unvalued [uT]-feature
- Every functional verbal head has a valued [iT]-feature corresponding to its semantic interpretation

The case study: prerequisites

The featural specification of the verbal domain (II)

- Every verbal head enters the derivation with an unvalued [uT]-feature
- Every functional verbal head has a valued [iT]-feature corresponding to its semantic interpretation
  - E.g.: v bears a [uT:_]-feature, Mod bears an [uT:_]-feature and [iT:Mod]-feature

The case study: prerequisites

The featural specification of the verbal domain (II)

- Every verbal head enters the derivation with an unvalued [uT]-feature
- Every functional verbal head has a valued [iT]-feature corresponding to its semantic interpretation
  - E.g.: v bears a [uT:_]-feature, Mod bears an [uT:_]-feature and [iT:Mod]-feature
- At PF, the valuation of [uT:_]-feature on v is what is morphologically realised on the given verb (see also Stechow 2003 et seq, and Gronn and Stechow 2011)

The case study: prerequisites

The spell out of *te*

- I assume that v can spell out *te* when the [uT:_]-feature on v has been valued for [irrealis]

The case study: prerequisites

The spell out of *te*

- I assume that v can spell out *te* when the [uT:_]-feature on v has been valued for [irrealis]
- I.e. *te* is not an independent functional head in narrow syntax, but the spell out of a feature on v when the right feature valuation has taken place

The case study: prerequisites

The spell out of *te*

- I assume that v can spell out *te* when the [uT:_]-feature on v has been valued for [irrealis]
- I.e. *te* is not an independent functional head in narrow syntax, but the spell out of a feature on v when the right feature valuation has taken place
- *Te* started out as an marker of irrealis clauses (IJbema 2001); I therefore assume that *te* is still associated with this feature

The case study: prerequisites

The direction of Agree

- I follow Wurmbrand (2012) in assuming that verbal feature valuation in Germanic is the result of Reverse Agree

The case study: prerequisites

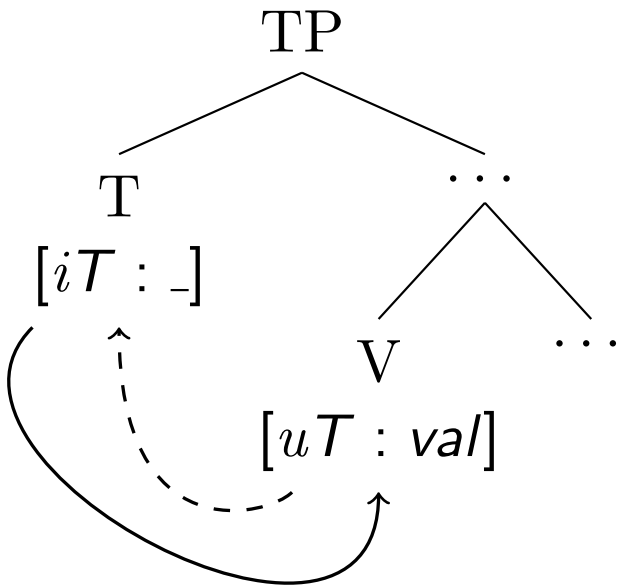
The direction of Agree

- I follow Wurmbrand (2012) in assuming that verbal feature valuation in Germanic is the result of Reverse Agree
- I.e. feature probing is upwards, valuation is downwards

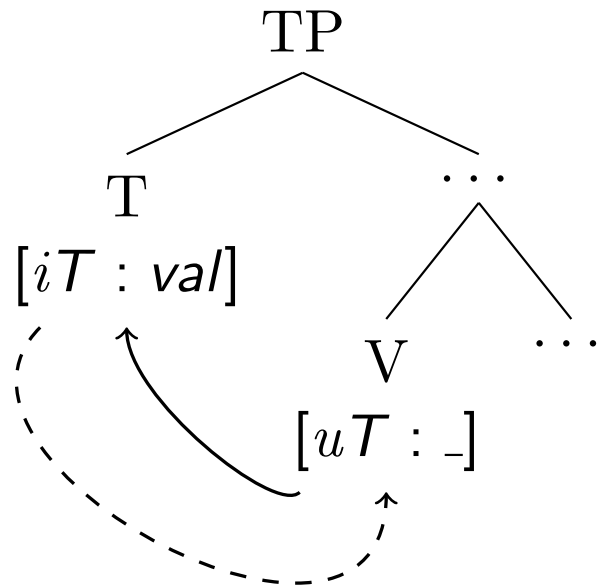
The case study: prerequisites

The direction of Agree

Standard Agree



Reverse Agree



The case study: the analysis

- A first step of the analysis: *hoeven* is grammaticalising from stage I of semi-lexicality to stage II

The case study: the analysis

Tests for semi-lexicality

- Wurmbrand (2001) has three tests for the status of semi-lexically used restructuring verbs:

The case study: the analysis

Tests for semi-lexicality

- Wurmbrand (2001) has three tests for the status of semi-lexically used restructuring verbs:
  1. they occur in IPP form when embedded under a perfective auxiliary

The case study: the analysis

Tests for semi-lexicality

- Wurmbrand (2001) has three tests for the status of semi-lexically used restructuring verbs:
  1. they occur in IPP form when embedded under a perfective auxiliary
  2. they do not allow extraposition

The case study: the analysis

Tests for semi-lexicality

- Wurmbrand (2001) has three tests for the status of semi-lexically used restructuring verbs:
  1. they occur in IPP form when embedded under a perfective auxiliary
  2. they do not allow extraposition
  3. they establish a thematic relation with the subject; they do not allow weather-*it* subjects

The case study: the analysis

Tests for semi-lexicality

- Wurmbrand (2001) has three tests for the status of semi-lexically used restructuring verbs:
  1. they occur in IPP form when embedded under a perfective auxiliary
  2. they do not allow extraposition
  3. they establish a thematic relation with the subject; they do not allow weather-*it* subjects
- Both *hoeven* and *zitten* pass the first and second test

Case study: the analysis

The IPP effect

(17) dat ze dat niet heeft **gehoeven**.
that she that not has need.PTCP
'that she didn't need that.'

(18) dat ze niet heeft {**hoeven**/***gehoeven**} (te) werken.
that she not has need.INF/need.PTCP to work
' that she didn't need to work.'

(19) dat ze niet heeft {**zitten**/***gezeten**} (te) werken.
that she not has sit.INF/sit.PTCP to work
'that she hasn't been working.'

Case study: the analysis

Blocking of extraposition

(20) dat ze **besluit** [de koek te eten].
that she decides the cookie to eat
'that she decides to eat the cookie.'

(21) *dat Frans niet **hoeft** [de koek te eten].
that Frans not need the cookie to eat
'that Frans doesn't need to eat the cookie.'

(22) *dat Frans **zit** [de koek te eten].
that Frans sit the cookie to eat
'that Frans is eating the cookie.'

Case study: the analysis

- *Hoeven* and *zitten* show different results on the third test

Case study: the analysis

Blocking of weather-*it* subjects

(23) ***Het* hoeft** niet te sneeuwen.
it need not to snow
'It doesn't need to snow.

(24) ****Het* zit** niet te sneeuwen.
it sit not to snow
'It is not snowing.'

Case study: the analysis

Hoeven in both the first and second stage

- *Hoeven* shares the capacity to occur with weather-*it* subjects with functional verbs (auxiliary verbs and modals)

Case study: the analysis

Hoeven in both the first and second stage

- *Hoeven* shares the capacity to occur with weather-*it* subjects with functional verbs (auxiliary verbs and modals)
- I take this as an indication that *hoeven* is grammaticalising from the first stage of semi-lexicality to the second stage of semi-lexicality

Case study: the analysis

Hoeven in both the first and second stage

- Further support for this assumption is given by Van de Velde (2017)

Case study: the analysis

Hoeven in both the first and second stage

- Further support for this assumption is given by Van de Velde (2017)
- Over the last 50 years, *hoeven* has shown a rapid increase in selecting a bare rather than a *te*-infinitive, while acquiring a more modal interpretation

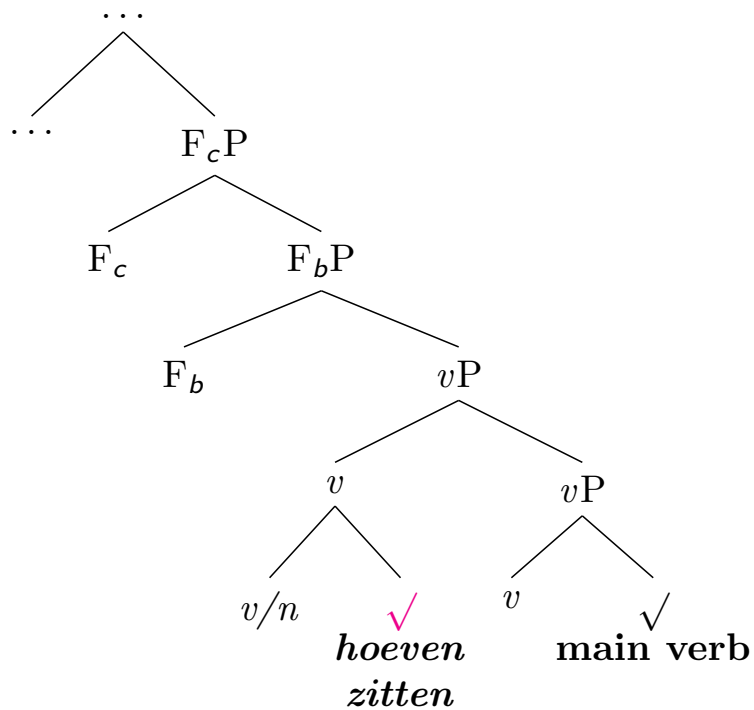
Case study: the analysis

Hoeven in both the first and second stage

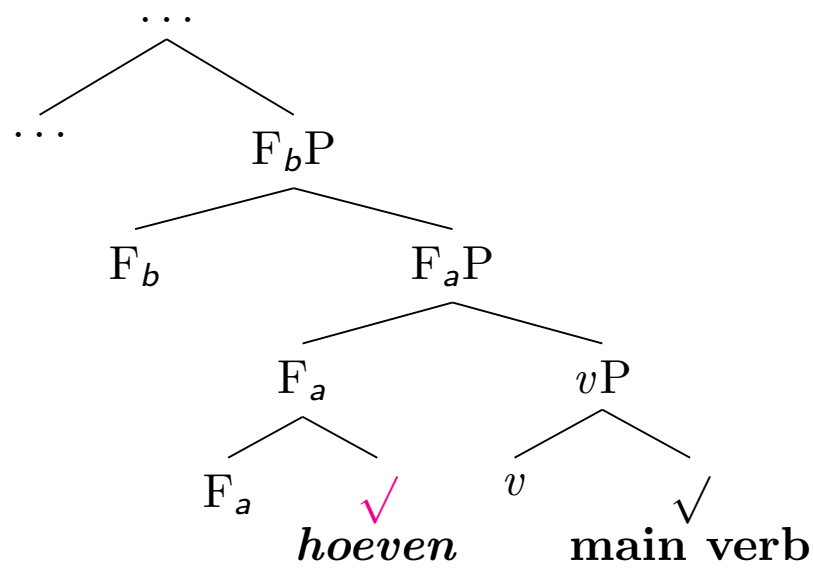
- Further support for this assumption is given by Van de Velde (2017)
- Over the last 50 years, *hoeven* has shown a rapid increase in selecting a bare rather than a *te*-infinitive, while acquiring a more modal interpretation
- Given that *zitten* blocks weather-*it* subjects and does not seem to change in its morphosyntactic behaviour, I assume that this verbs is uniformly in the first stage of semi-lexicality

Case study: the analysis

Semi-lex stage I



Semi-lex stage II



Case study: the analysis

- *Recall*: the morphosyntactic behaviour of *hoeven* was tested with the following test item

(25) Koen zal niet [hoeven$_1$ *te* gaan$_2$ voetballen$_3$].
Koen will not need.INF to go.INF play.football.INF.
'Koen won't have to go and play football.'

Case study: the analysis

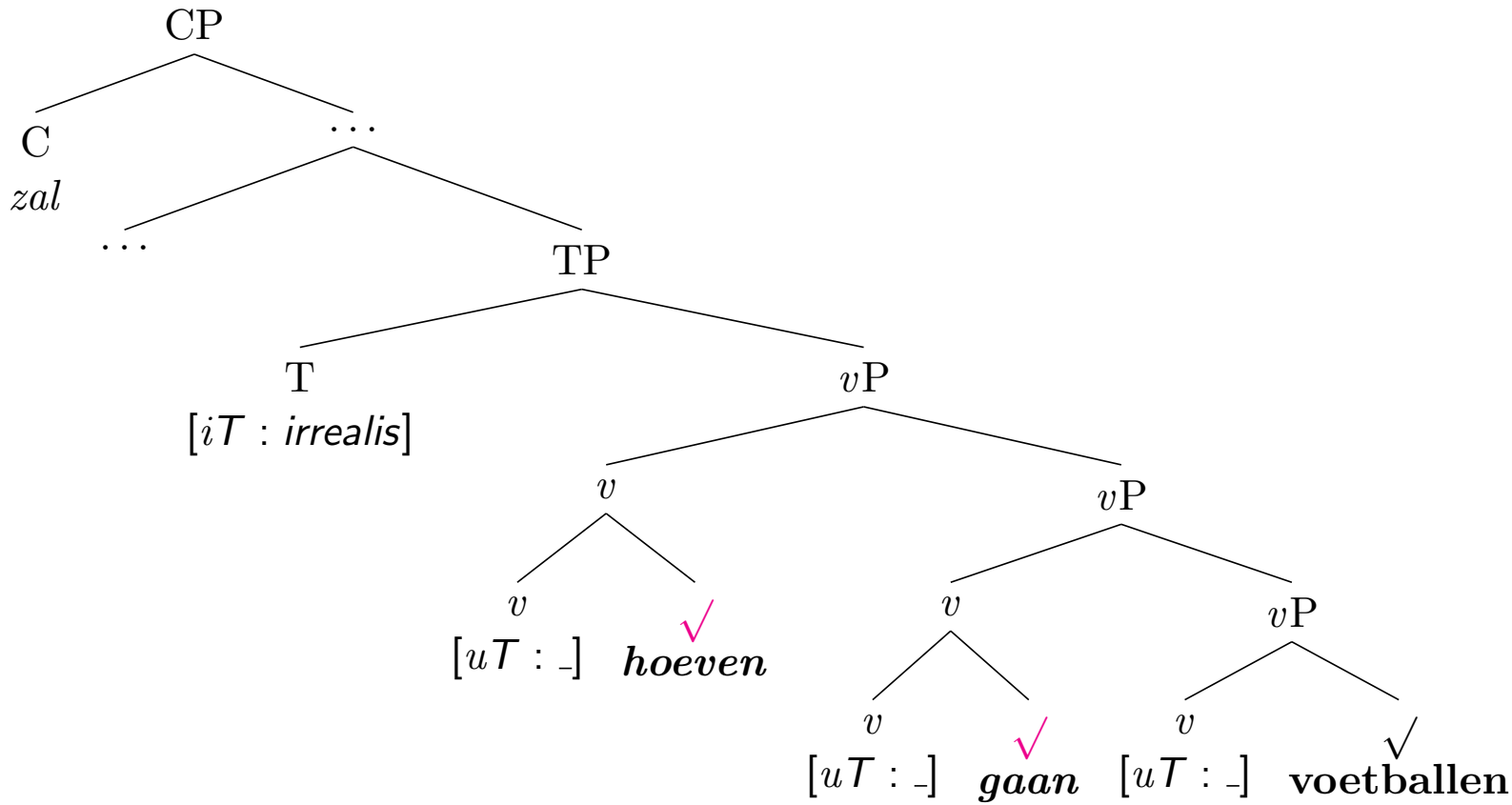
- *Recall*: the morphosyntactic behaviour of *hoeven* was tested with the following test item

(25) Koen zal niet [hoeven$_1$ *te* gaan$_2$ voetballen$_3$].
Koen will not need.INF to go.INF play.football.INF.
'Koen won't have to go and play football.'

- Applying the three tests for semi-lexicality, V2 *gaan* 'go' is semi-lexical as well, and blocks weather-*it* subjects (i.e. semi-lexical stage I)

Case study: the analysis

(26) **Structure of *hoeven* cluster (sem-lex stage I)**



The case study: the analysis

Recap of the data

- In the *hoeven* cluster, low-*te* is the most frequent, but high-*te* and *te*-drop occur relatively frequently as well

The case study: the analysis

Recap of the data

- In the *hoeven* cluster, low-*te* is the most frequent, but high-*te* and *te*-drop occur relatively frequently as well
- *te*-doubling is very infrequent

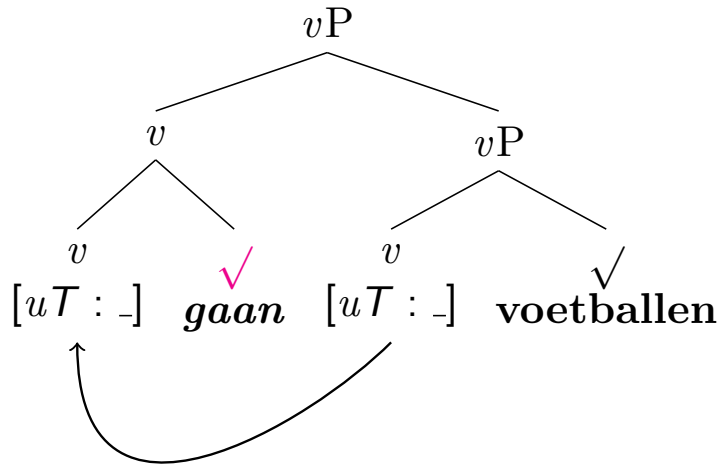
The case study: the analysis

Recap of the data

- In the *hoeven* cluster, low-*te* is the most frequent, but high-*te* and *te*-drop occur relatively frequently as well
- *te*-doubling is very infrequent
- The *hoeven* cluster shows a high degree of intra-speaker variation

Case study: the analysis

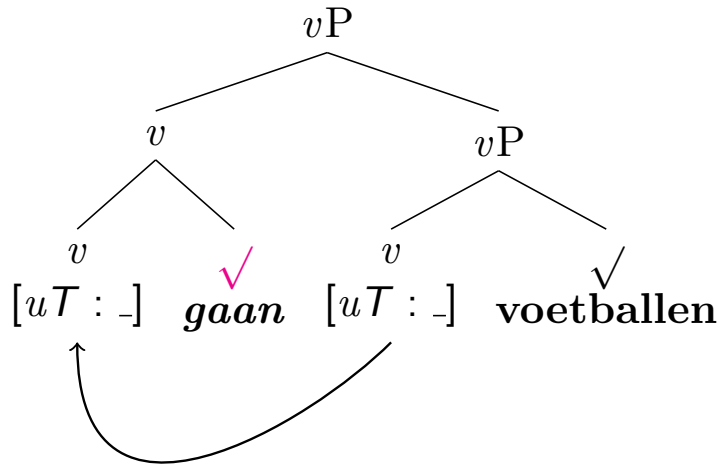
(27) **Agree step I (sem-lex stage I)**



- $[uT : \_]$ on the lower v probes up and Agrees with the $[uT : \_]$ on the second v

Case study: the analysis

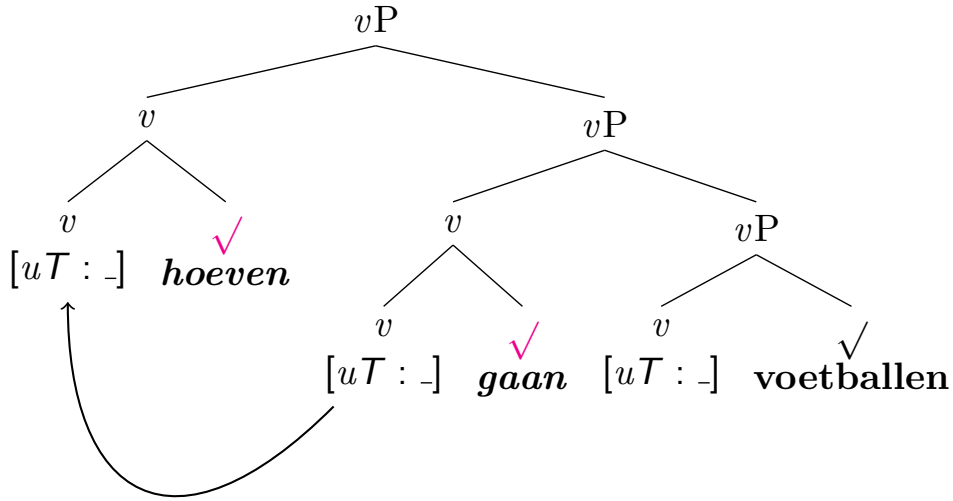
(27) **Agree step I (sem-lex stage I)**



- $[uT : \_]$ on the lower v probes up and Agrees with the $[uT : \_]$ on the second v
- No valuation is possible, but a feature link is established (cf. Pesetsky & Torrego 2007; Haegeman & Lohndal 2010)

Case study: the analysis

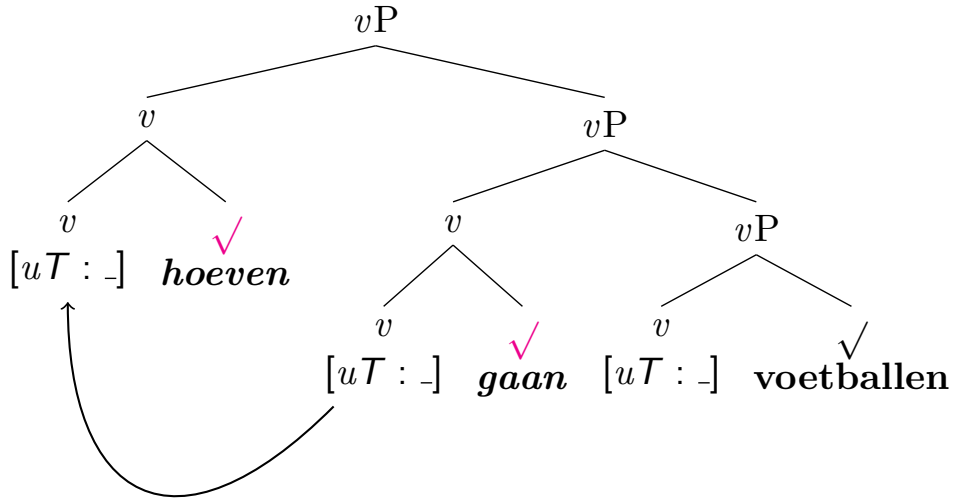
(28) **Agree step II (sem-lex stage I)**



- $[uT : \_]$ on the middle v probes up and Agrees with the $[uT : \_]$ on the highest v

Case study: the analysis

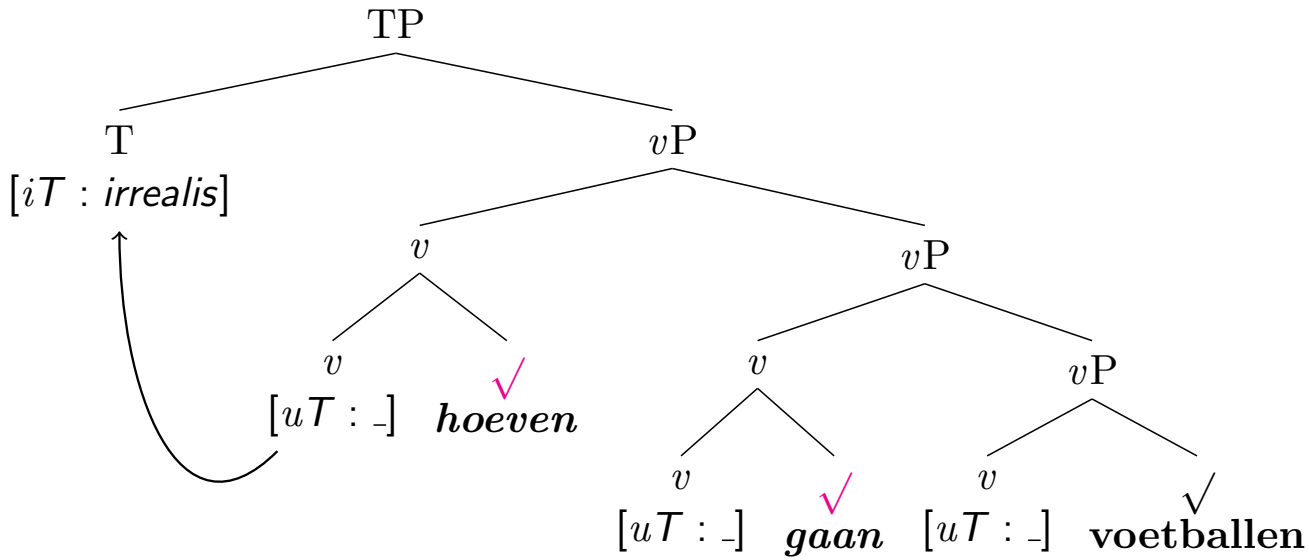
(28) **Agree step II (sem-lex stage I)**



- $[uT : \_]$ on the middle v probes up and Agrees with the $[uT : \_]$ on the highest v
- No valuation is possible, but a feature link is established (cf. Pesetsky & Torrego 2007; Haegeman & Lohndal 2010)

Case study: the analysis

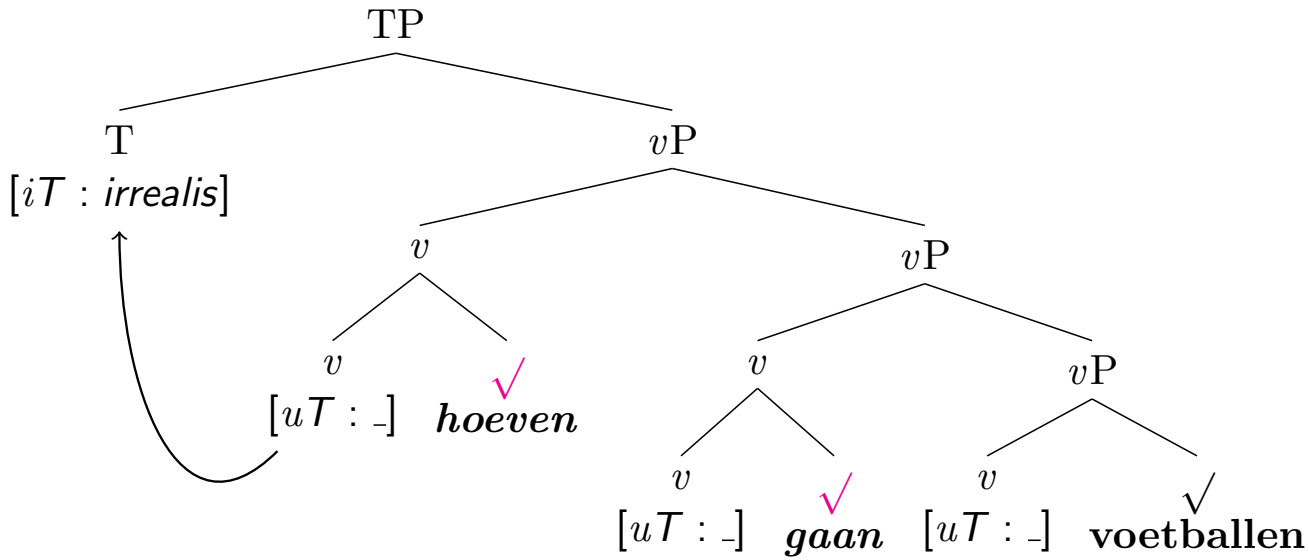
(29) **Agree step III (sem-lex stage I)**



- T is Merged, and comes with an [*i*T]-feature valued for [irrealis] (due to modal *zal* 'will', which will surface in V2 position)

Case study: the analysis

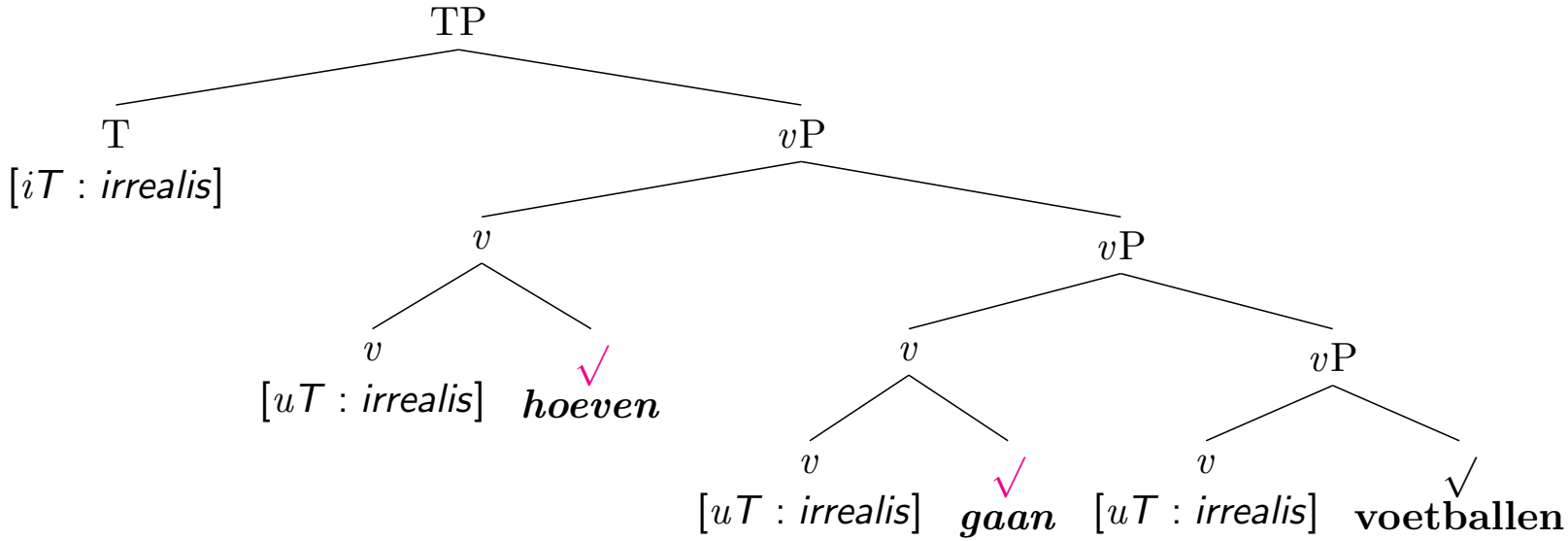
(29) **Agree step III (sem-lex stage I)**



- T is Merged, and comes with an [iT]-feature valued for [irrealis] (due to modal *zal* 'will', which will surface in V2 position)
- The [uT]-feature on the highest v probes up and Agrees with [iT] on T

Case study: the analysis

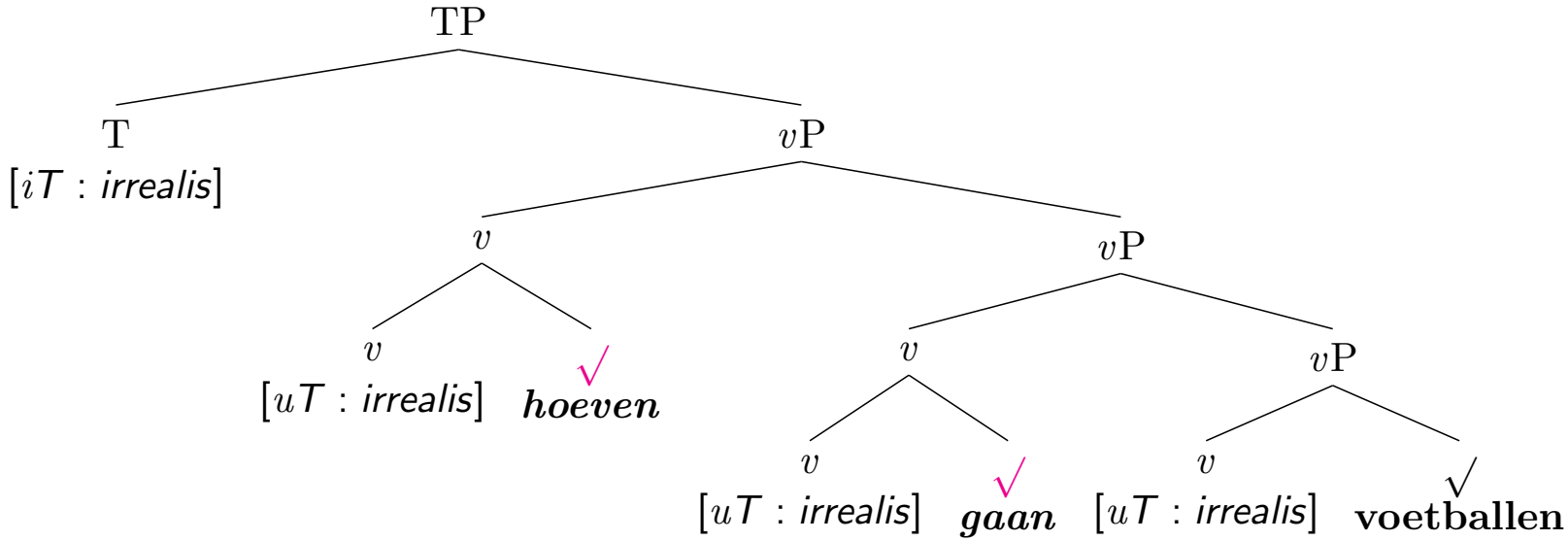
(30) **Valuation (sem-lex stage I)**



- Since the three [uT]'s in the structure Agreed before and formed a feature chain, they all get valued for [irrealis]

Case study: the analysis

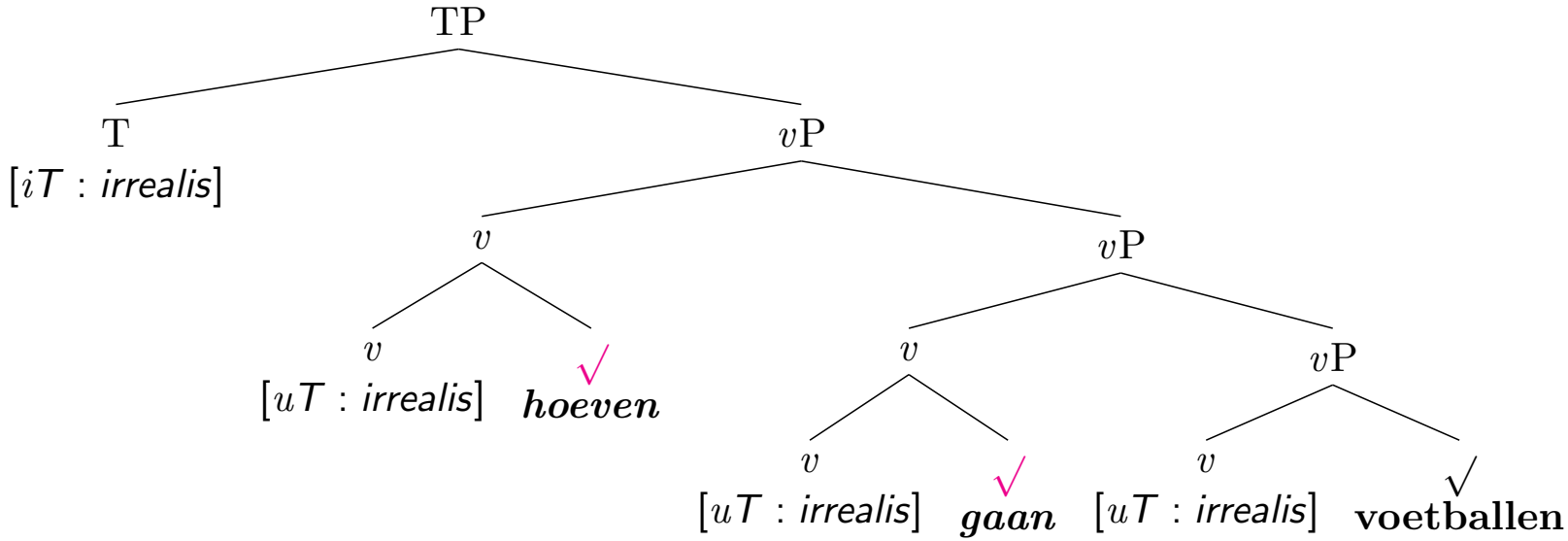
(30) **Valuation (sem-lex stage I)**



- Since the three [uT]'s in the structure Agreed before and formed a feature chain, they all get valued for [irrealis]
- *Recall*: spelling out *te* is possible when the [uT]-feature on v has been valued for [irrealis]

Case study: the analysis

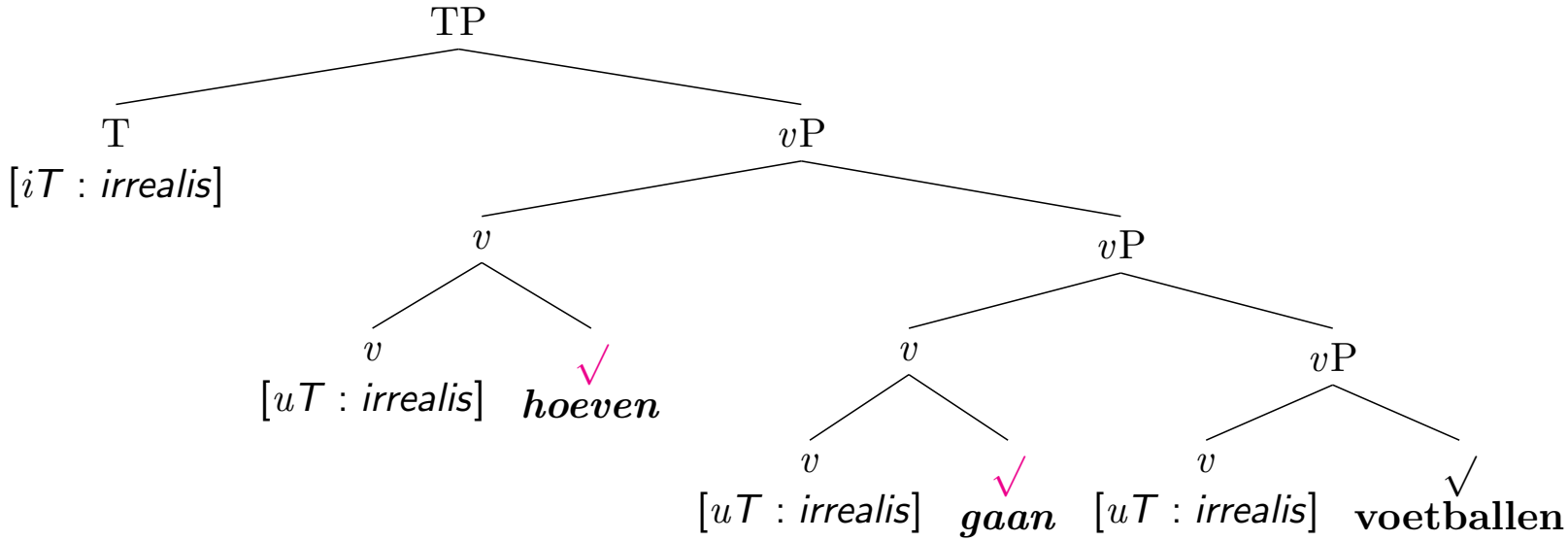
(31) **Valuation (sem-lex stage I)**



- In this cluster, *te* can therefore be spelled out on all verbs

Case study: the analysis

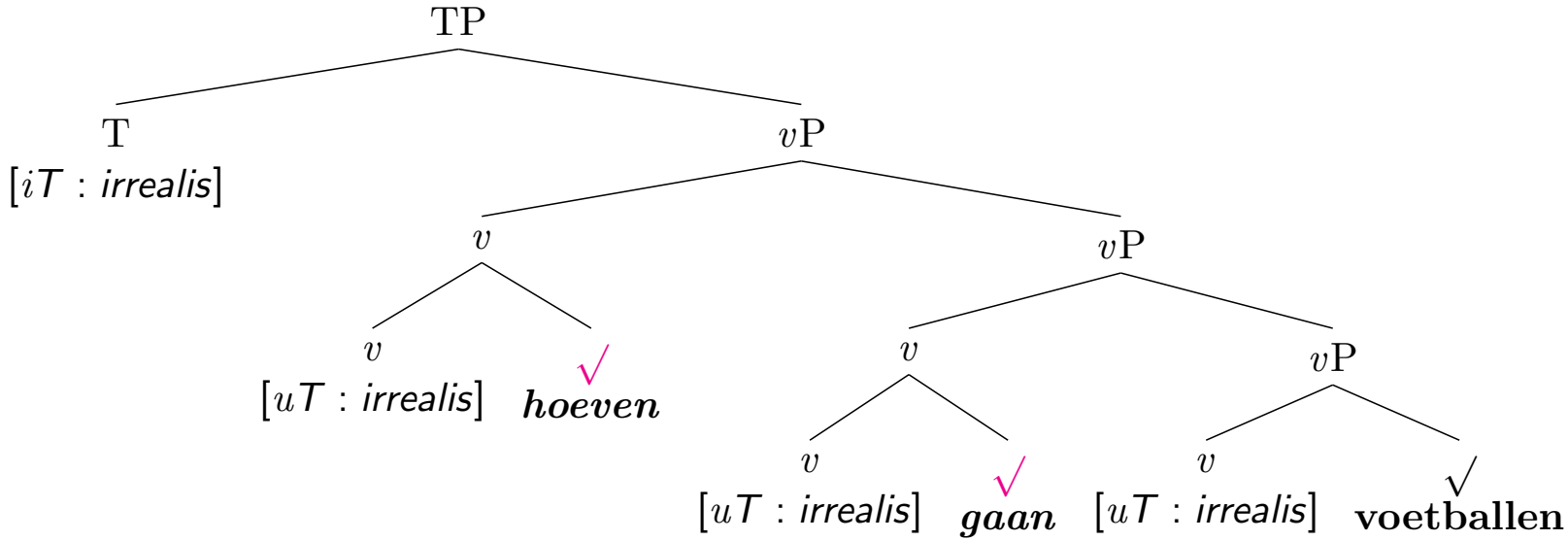
(31) **Valuation (sem-lex stage I)**



- In this cluster, *te* can therefore be spelled out on all verbs
- I assume that *te*-tripling is excluded by haplology

Case study: the analysis

(31) **Valuation (sem-lex stage I)**



- In this cluster, *te* can therefore be spelled out on all verbs
- I assume that *te*-tripling is excluded by haplology
- High-*te*, low-*te* and *te*-doubling are expected to occur, which is indeed the case

Case study: the analysis

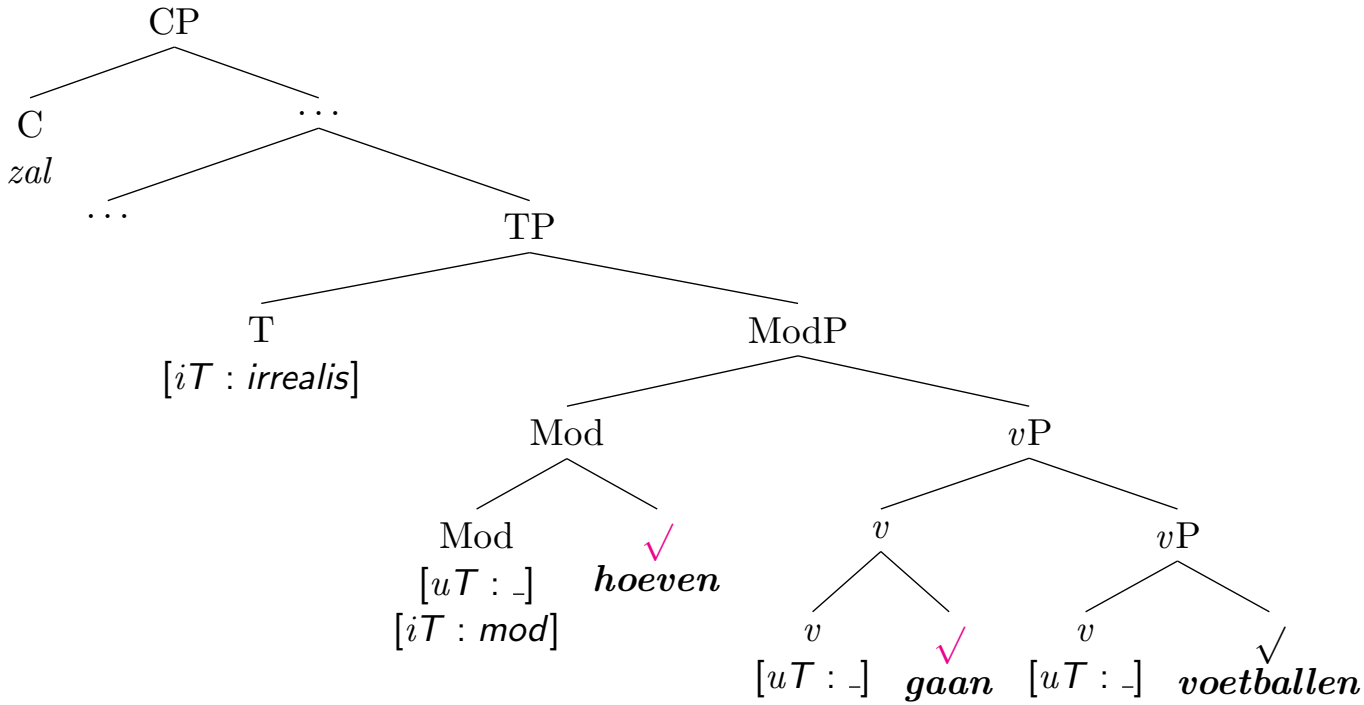
- I assume that *te*-doubling is the least preferred option because it is more effortful or redundant than spelling out one feature of a feature chain

Case study: the analysis

- I assume that *te*-doubling is the least preferred option because it is more effortful or redundant than spelling out one feature of a feature chain
- The high degree of intra-speaker optionality is exactly what is expected in such a configuration: syntax doesn't care

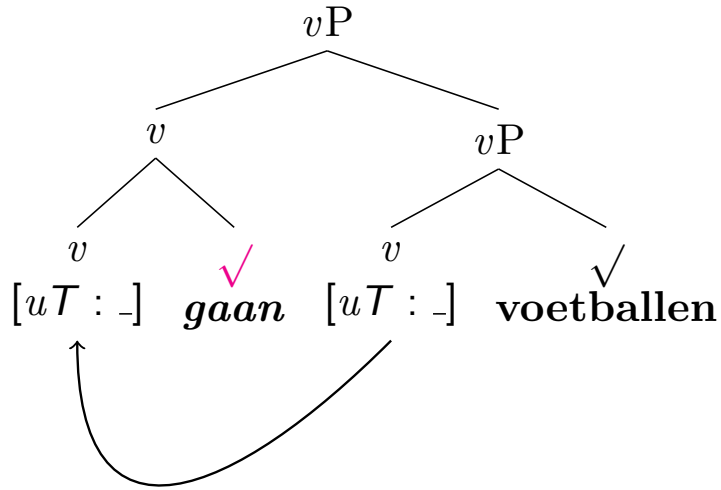
Case study: the analysis

(32) **Structure of *hoeven* cluster (sem-lex stage II)**



Case study: the analysis

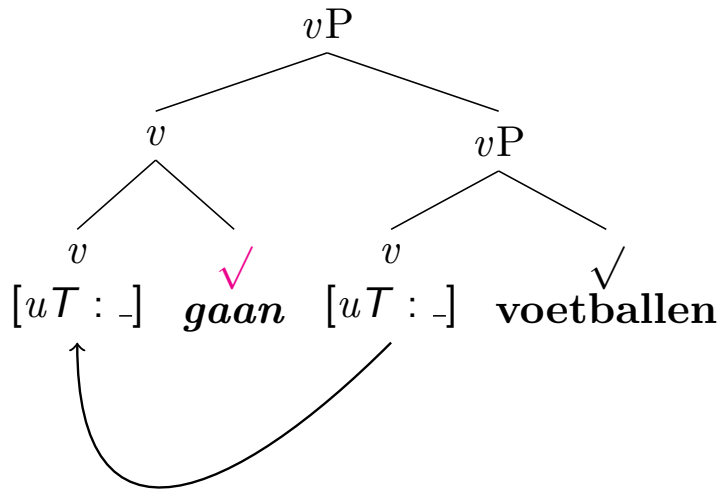
(33) **Agree step I (sem-lex stage II)**



- $[uT : \_]$ on the lower v probes up and Agrees with the $[uT : \_]$ on the second v

Case study: the analysis

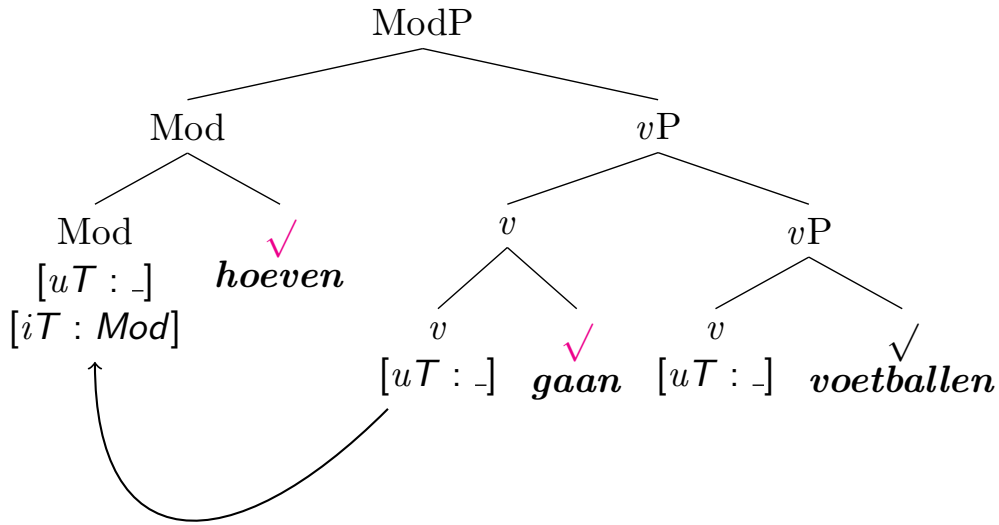
(33) **Agree step I (sem-lex stage II)**



- $[uT : \_]$ on the lower v probes up and Agrees with the $[uT : \_]$ on the second v
- No valuation is possible, but a feature link is established

Case study: the analysis

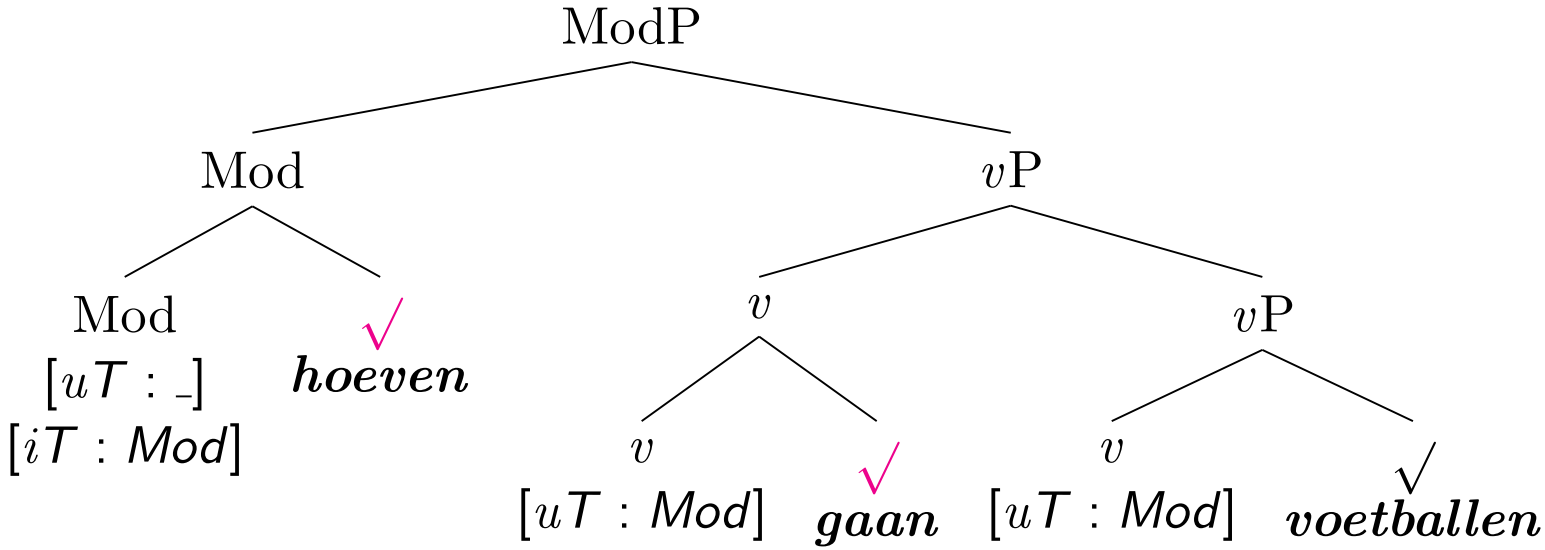
(34) **Agree step II (sem-lex stage II)**



- $[uT : \_]$ on the middle v probes up and Agrees with the $[iT : Mod]$ on Mod

Case study: the analysis

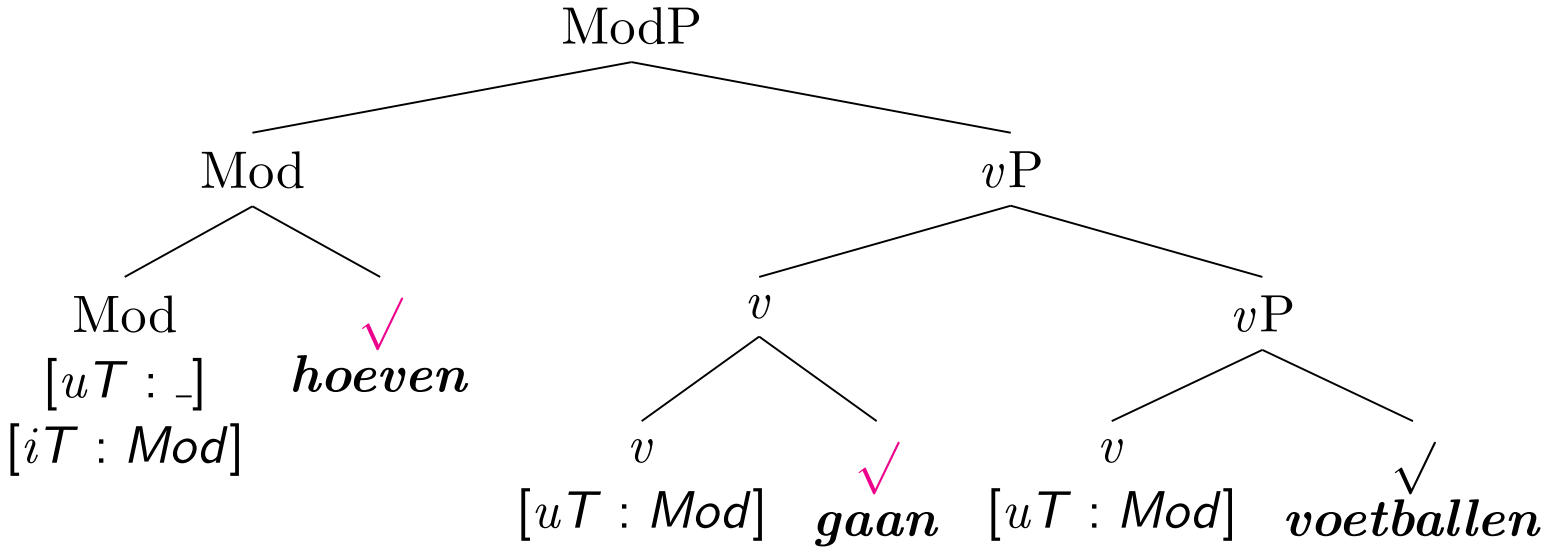
(35) **Valuation step I (sem-lex stage II)**



- Both [uT] get valued for [Mod]

Case study: the analysis

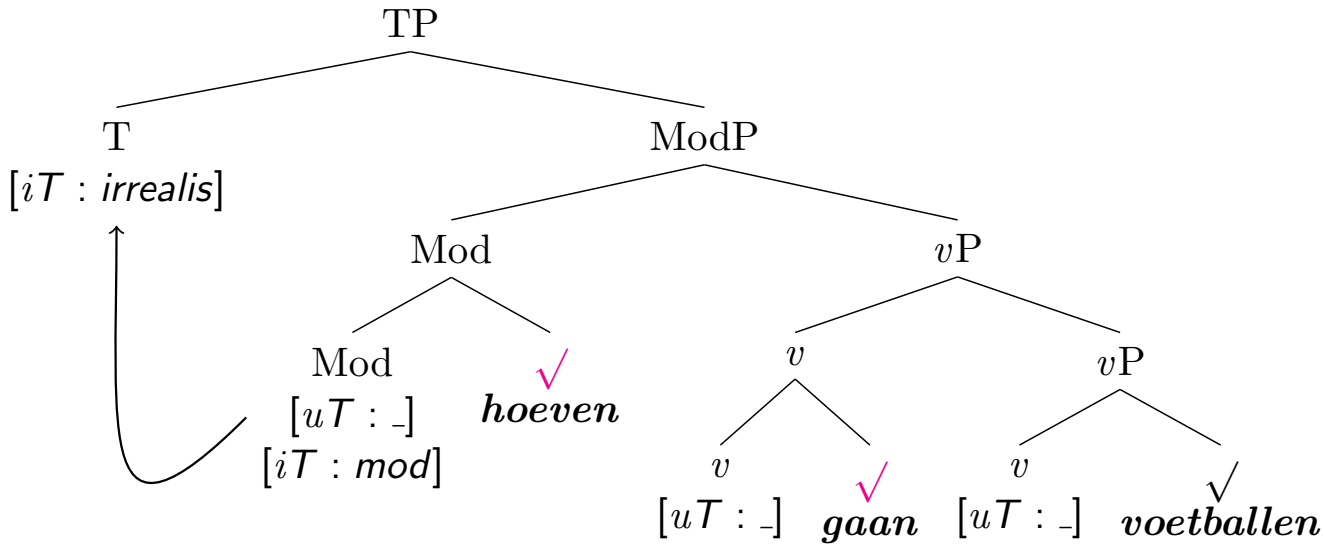
(35) **Valuation step I (sem-lex stage II)**



- Both [uT] get valued for [Mod]
- This valuation results in both verbs being spelled out as a bare infinitive

Case study: the analysis

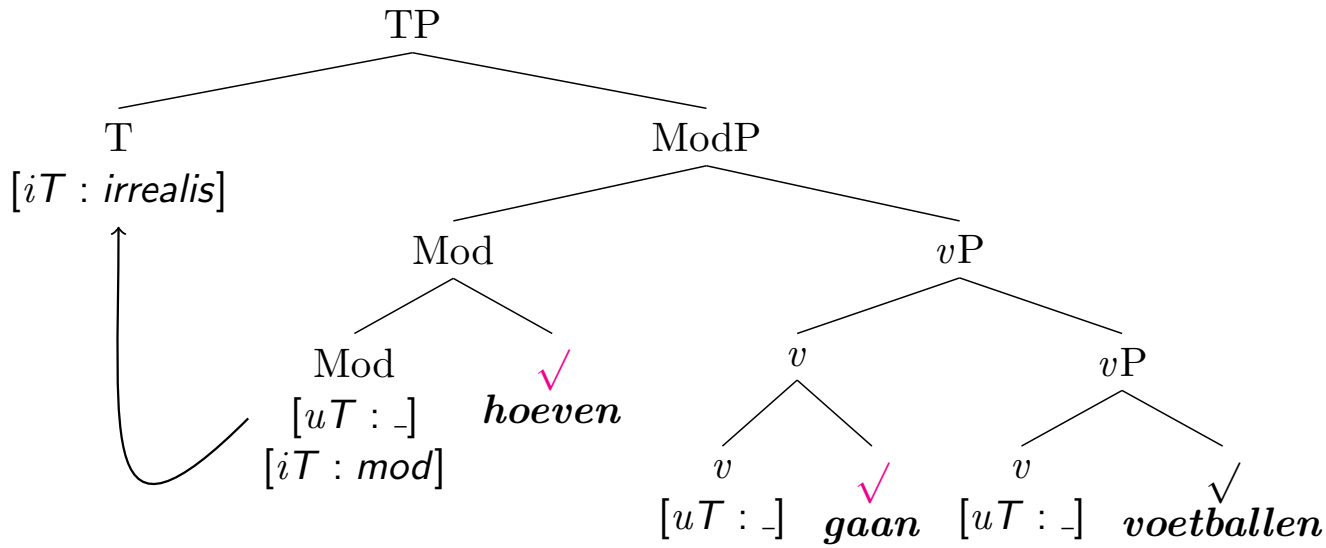
(36) **Agree step III (sem-lex stage II)**



- T is merged, and comes with an [*i*T:irrealis] feature (due to *zal* 'will' in V2)

Case study: the analysis

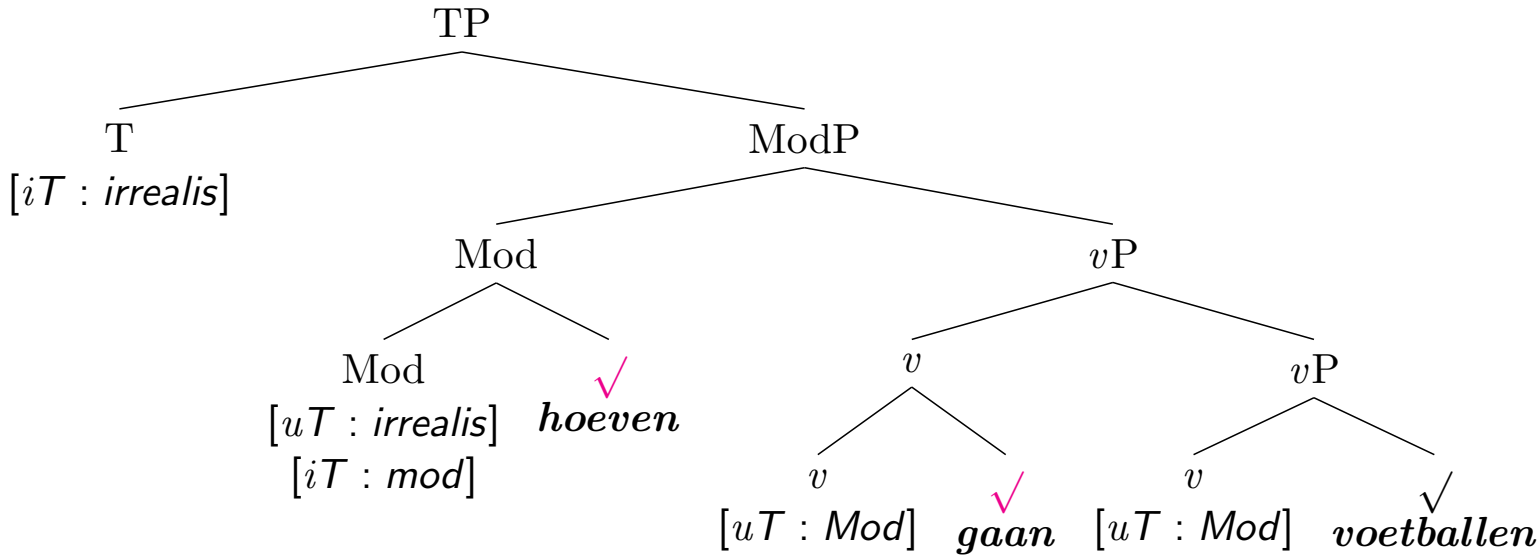
(36) **Agree step III (sem-lex stage II)**



- T is merged, and comes with an [*i*T:irrealis] feature (due to *zal* 'will' in V2)
- [*u*T:_] on Mod probes up and Agrees with [*i*T:irrealis] on T

Case study: the analysis

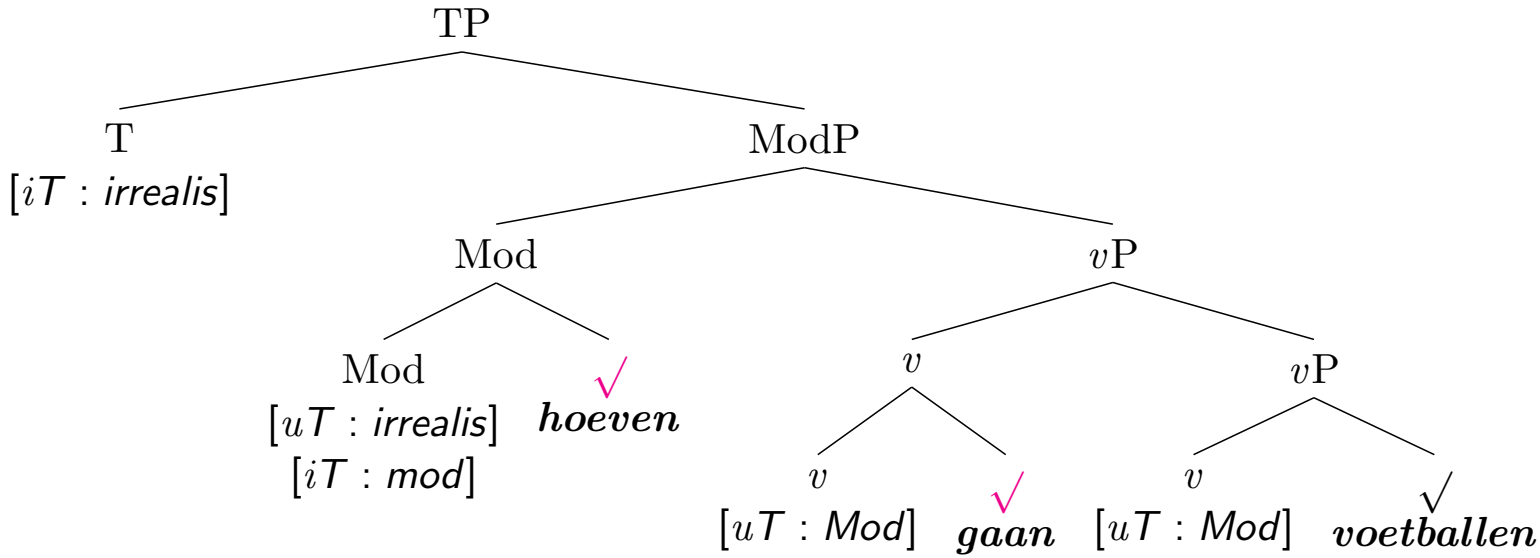
(37) **Valuation step II (sem-lex stage II)**



- *Recall*: *te* can only be spelled out if a [uT]-feature on v has been valued for [irrealis]

Case study: the analysis

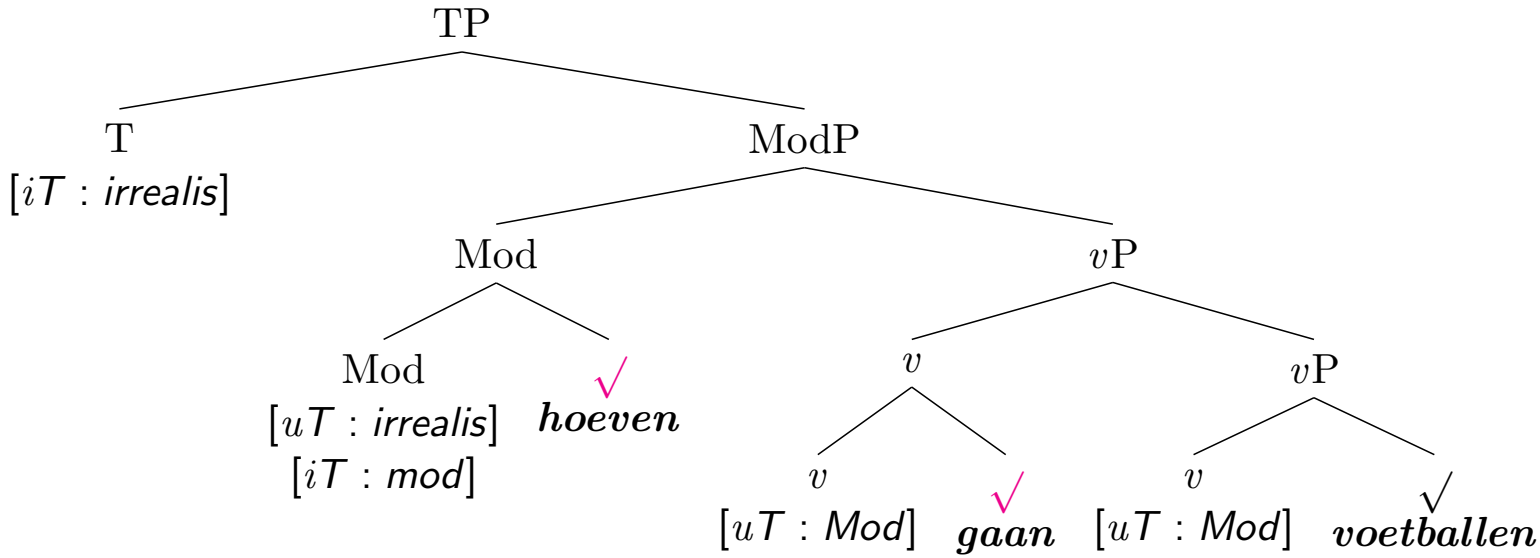
(37) **Valuation step II (sem-lex stage II)**



- *Recall*: *te* can only be spelled out if a [*u*T]-feature on *v* has been valued for [irrealis]
- I.e. [*u*T:irrealis] on Mod cannot be spelled out as *te* on *hoeven*

Case study: the analysis

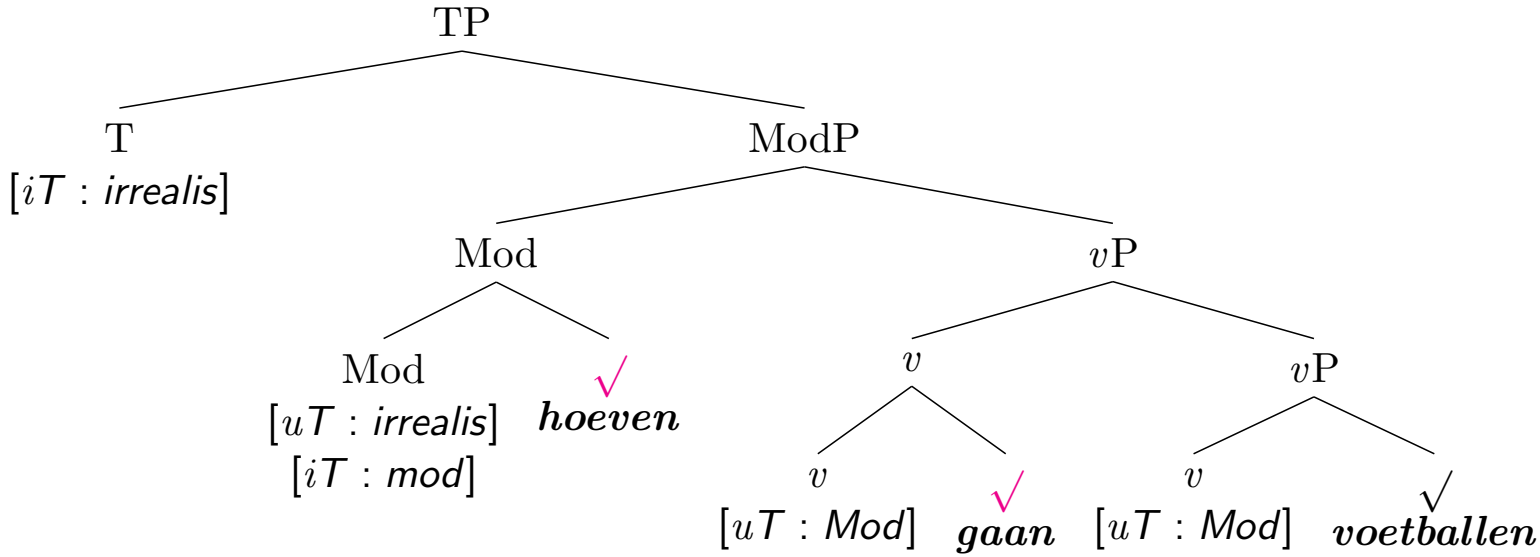
(37) **Valuation step II (sem-lex stage II)**



- *Recall*: *te* can only be spelled out if a [*u*T]-feature on *v* has been valued for [irrealis]
- I.e. [*u*T:irrealis] on Mod cannot be spelled out as *te* on *hoeven*
- This results in *hoeven* being spelled out as a bare infinitive

Case study: the analysis

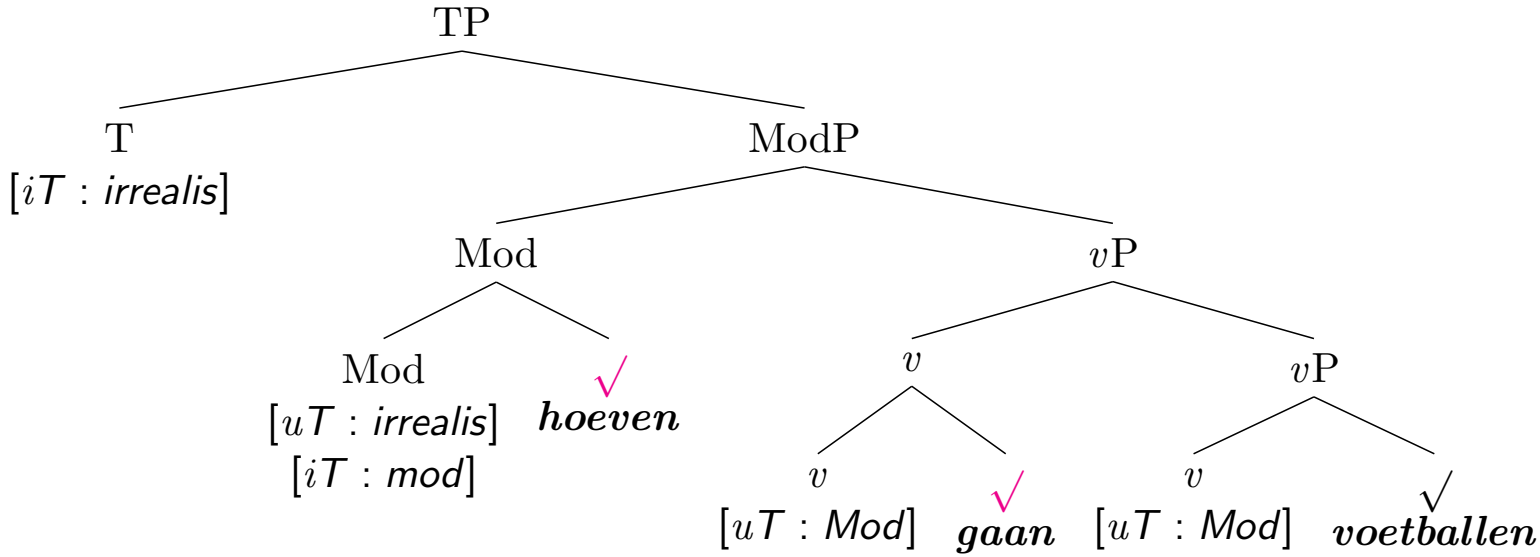
(38) **Valuation step II (sem-lex stage II)**



- I.e. when *hoeven* is in the second stage of semi-lexicality, I expect *te*-drop to occur across the board

Case study: the analysis

(38) **Valuation step II (sem-lex stage II)**



- I.e. when *hoeven* is in the second stage of semi-lexicality, I expect *te*-drop to occur across the board
- *Hoeven*'s grammaticalisation is still ongoing: high degrees of optionality (both inter- and intra-speaker) are expected

Case study: the analysis

- *Recall*: the morphosyntactic behaviour of *zitten* was tested with the following test item

(39) Peter zal lang [moeten$_1$ zitten$_2$ *te* wachten$_3$].
Peter will long must.INF sit.INF to wait.INF.
'Peter will have to wait for a long time.'

Case study: the analysis

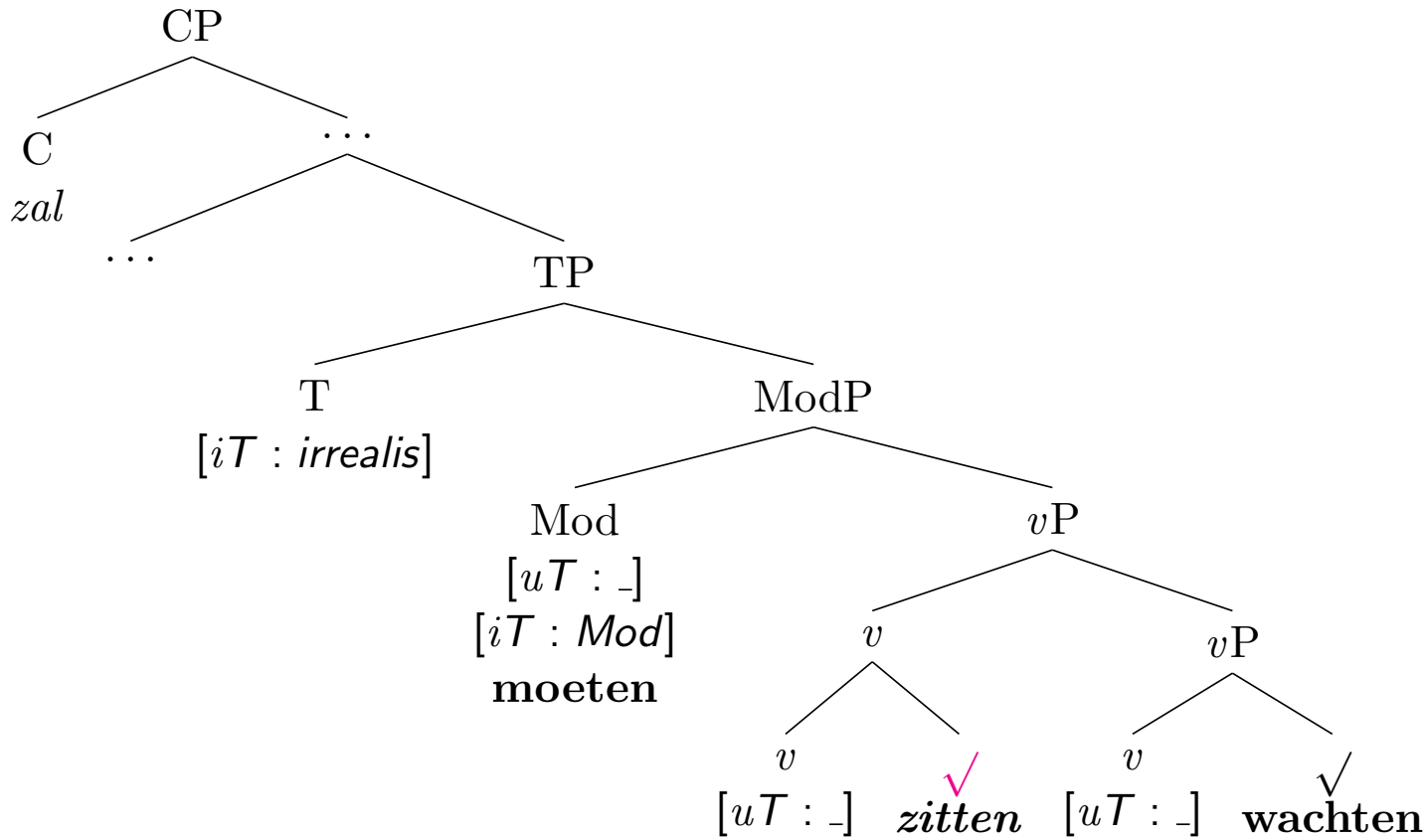
- *Recall*: the morphosyntactic behaviour of *zitten* was tested with the following test item

(39) Peter zal lang [moeten$_1$ zitten$_2$ *te* wachten$_3$].
Peter will long must.INF sit.INF to wait.INF.
'Peter will have to wait for a long time.'

- Applying the three tests for semi-lexicality, V1 *moeten* 'must' is used as a functional verb

Case study: the analysis

(40) **Structure of *zitten* cluster (sem-lex stage I)**



The case study: the analysis

Recap of the data

- In the *zitten* cluster, *te*-drop is by far the most frequent, low-*te* relatively frequent, high-*te* infrequent

The case study: the analysis

Recap of the data

- In the *zitten* cluster, *te*-drop is by far the most frequent, low-*te* relatively frequent, high-*te* infrequent
- *te*-doubling is virtually non-existent

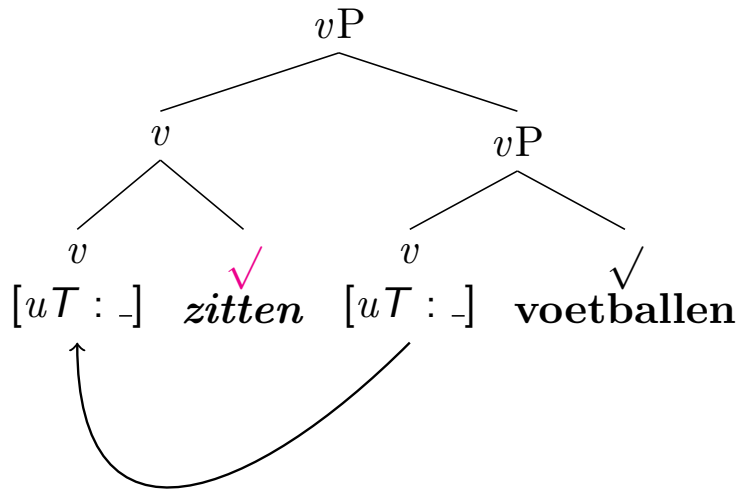
The case study: the analysis

Recap of the data

- In the *zitten* cluster, *te*-drop is by far the most frequent, low-*te* relatively frequent, high-*te* infrequent
- *te*-doubling is virtually non-existent
- The *zitten* cluster shows a lower degree of intra-speaker variation compared to the *hoeven* cluster

Case study: the analysis

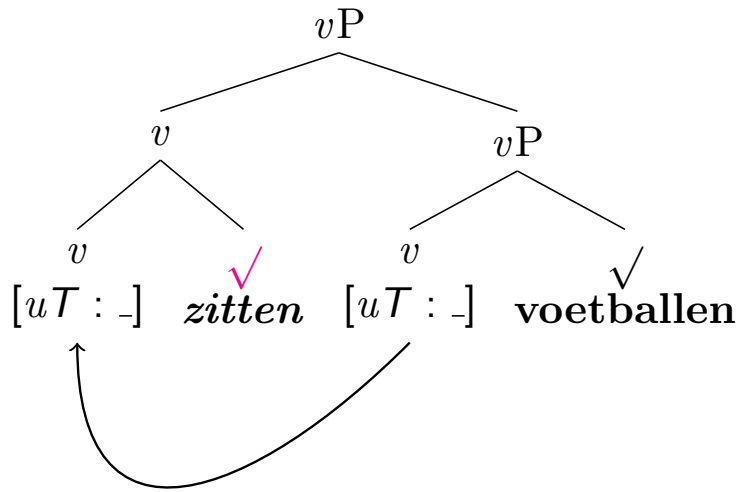
(41) **Agree step I (sem-lex stage I)**



- $[uT : \_]$ on the lower v probes up and Agrees with the $[uT : \_]$ on the second v

Case study: the analysis

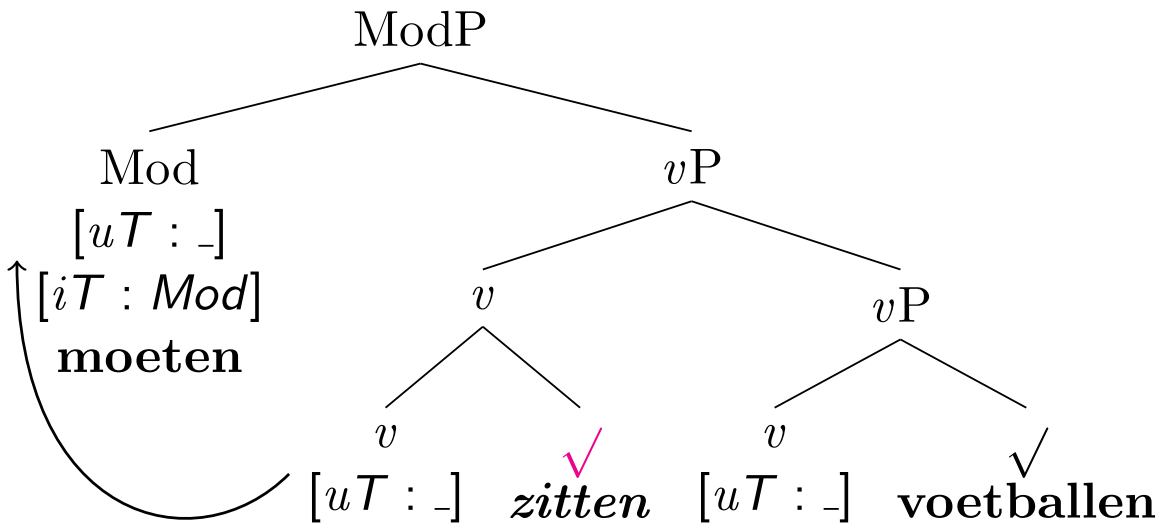
(41) **Agree step I (sem-lex stage I)**



- $[uT : \_]$ on the lower v probes up and Agrees with the $[uT : \_]$ on the second v
- No valuation is possible, but a feature link is established

Case study: the analysis

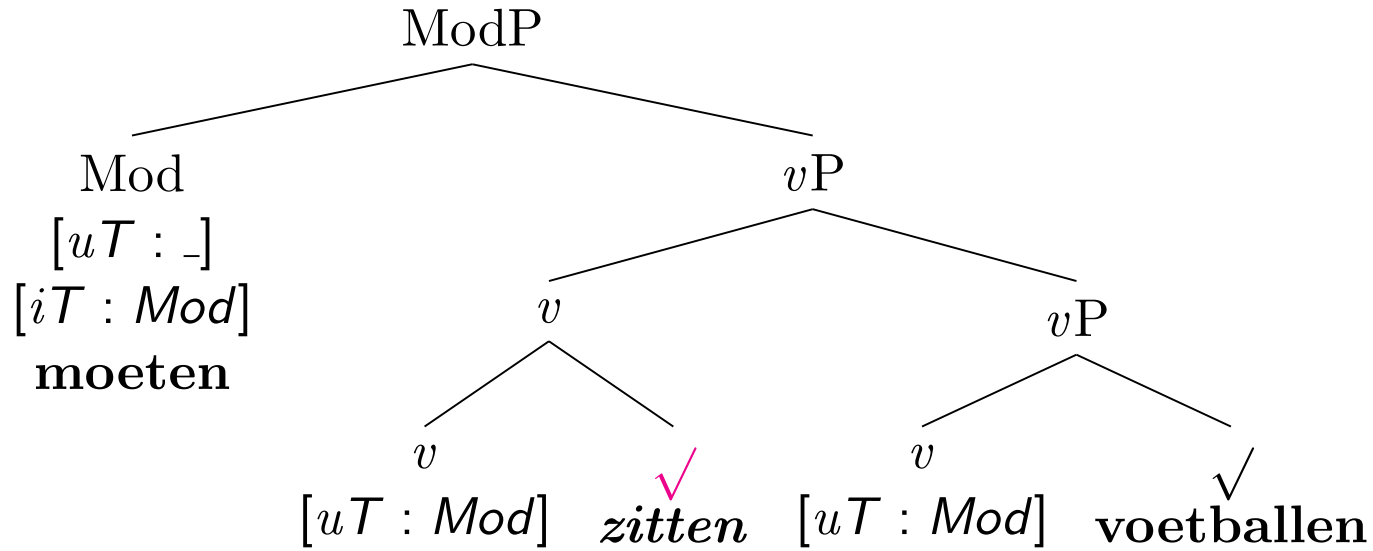
(42) **Agree step II (sem-lex stage I)**



- [uT : _] on the second v probes up and Agrees with [iT : *Mod*] on Mod

Case study: the analysis

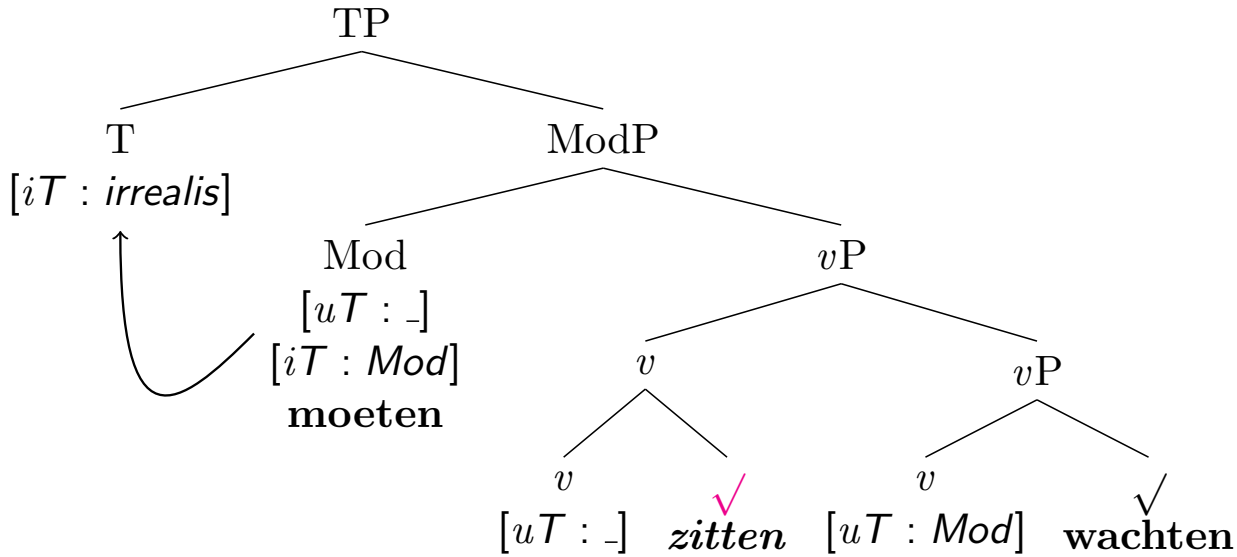
(43) **Valuation step I (sem-lex stage I)**



- Both *zitten* and the lexical verb will be spelled out as a bare infinitive due to [*uT* : *Mod*] on both *v*'s

Case study: the analysis

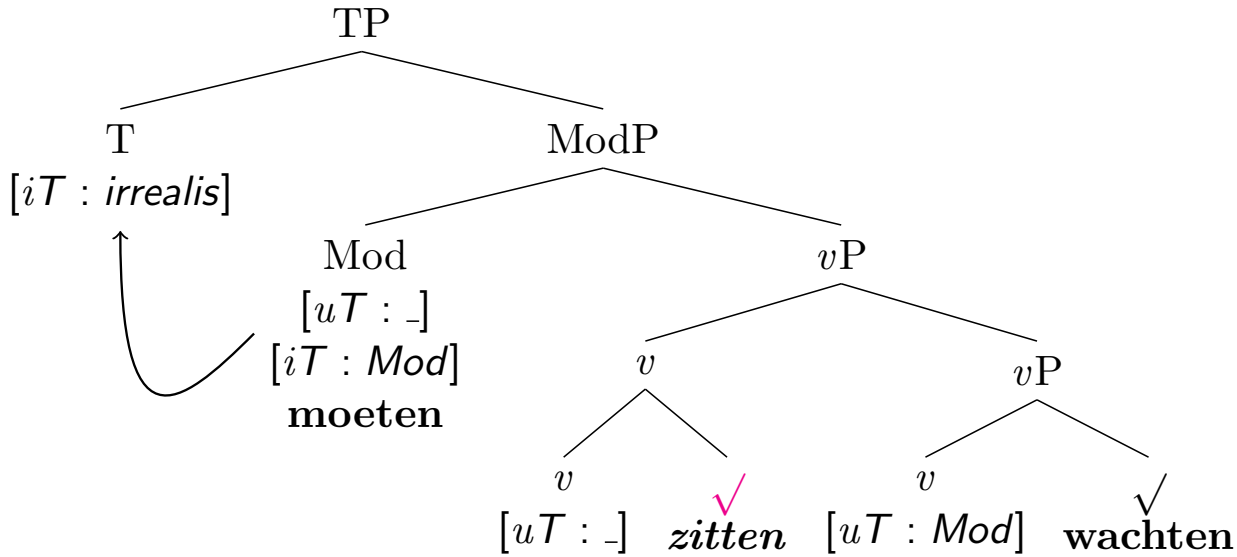
(44) **Agree step III (sem-lex stage I)**



- T is Merged, and comes with an [*i*T]-feature valued for [irrealis] (due to modal *zal* 'will', which will surface in V2 position)

Case study: the analysis

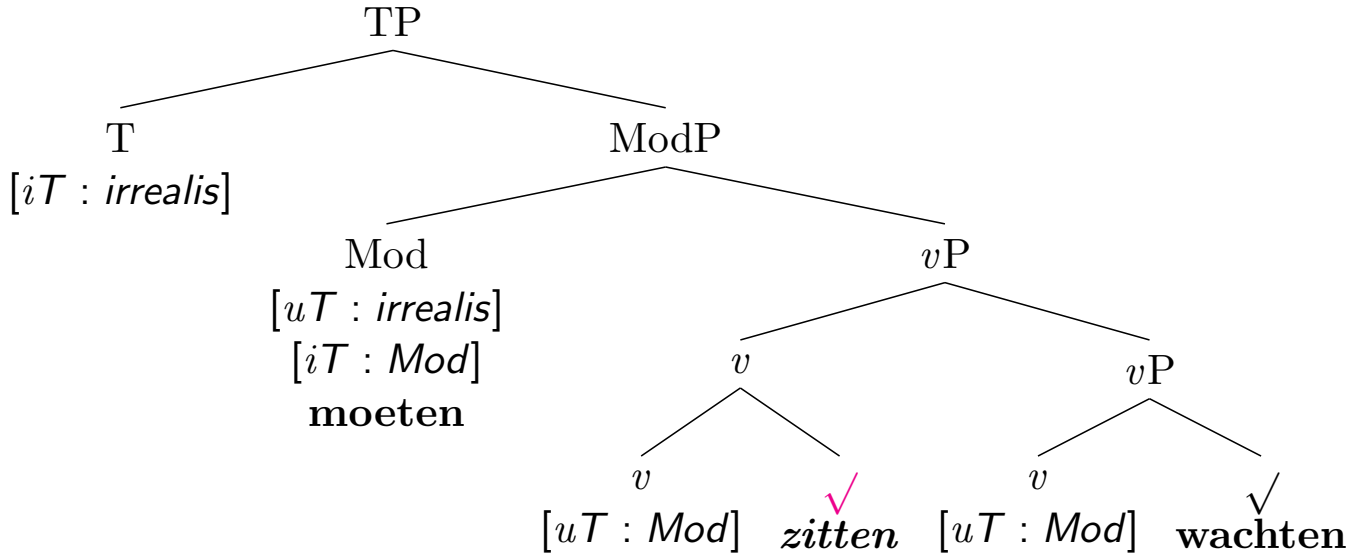
(44) **Agree step III (sem-lex stage I)**



- T is Merged, and comes with an [*i*T]-feature valued for [irrealis] (due to modal *zal* 'will', which will surface in V2 position)
- [*u*T] on Mod probes up and Agrees with [*i*T:irrealis] on T

Case study: the analysis

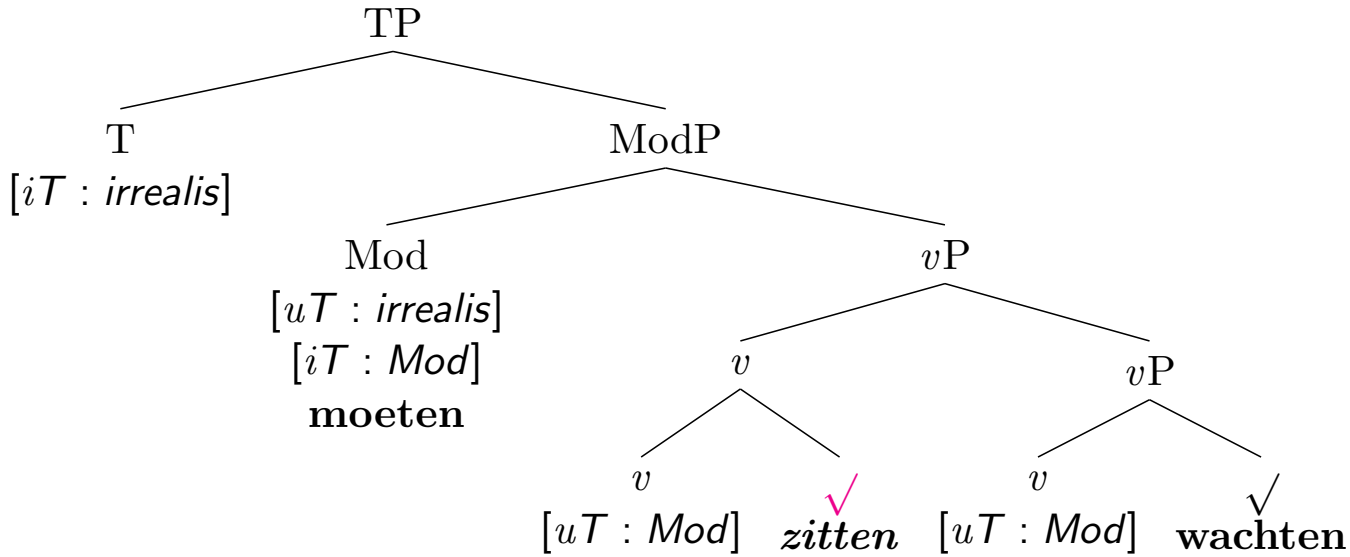
(45) **Valuation step II (sem-lex stage I)**



- *Recall*: only a v with an [uT:irrealis]-feature can spell out *te*

Case study: the analysis

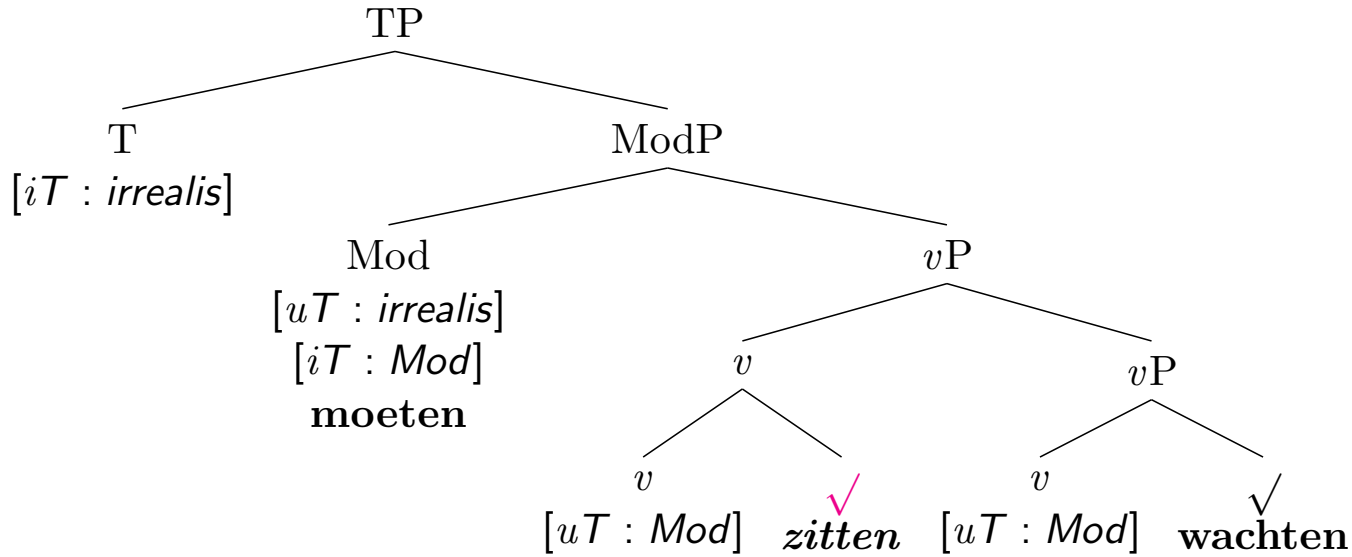
(45) **Valuation step II (sem-lex stage I)**



- *Recall*: only a v with an [uT:irrealis]-feature can spell out *te*
- Mod in this cluster must therefore be spelled out as a bare infinitive

Case study: the analysis

(45) **Valuation step II (sem-lex stage I)**



- *Recall*: only a v with an [uT:irrealis]-feature can spell out *te*
- Mod in this cluster must therefore be spelled out as a bare infinitive
- I.e. this structure predicts *te*-drop, which is indeed by far the most frequent pattern (71%)

Case study: the analysis

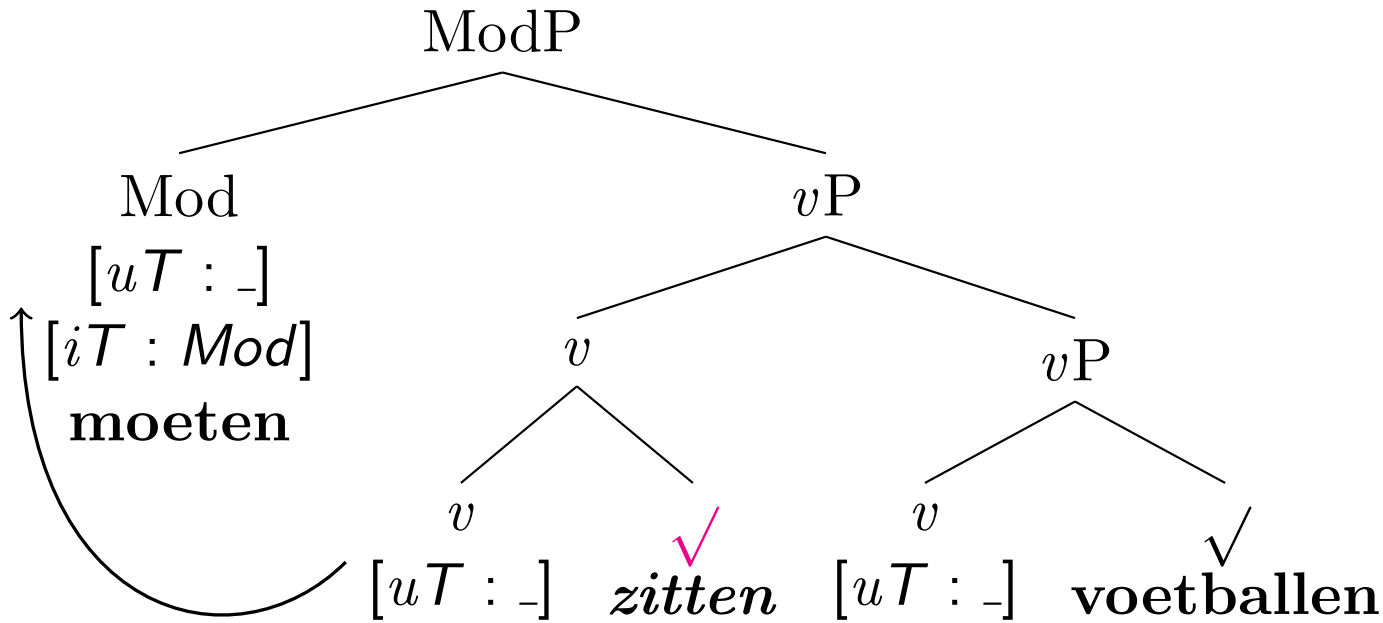
- However, low-*te* (relatively infrequent) and *te*-raising (very infrequent) occur as well

Case study: the analysis

- However, low-*te* (relatively infrequent) and *te*-raising (very infrequent) occur as well
- For those infrequent patterns, I propose they are the result of an alternative way of Agree having taken place at Agree step II

Case study: the analysis

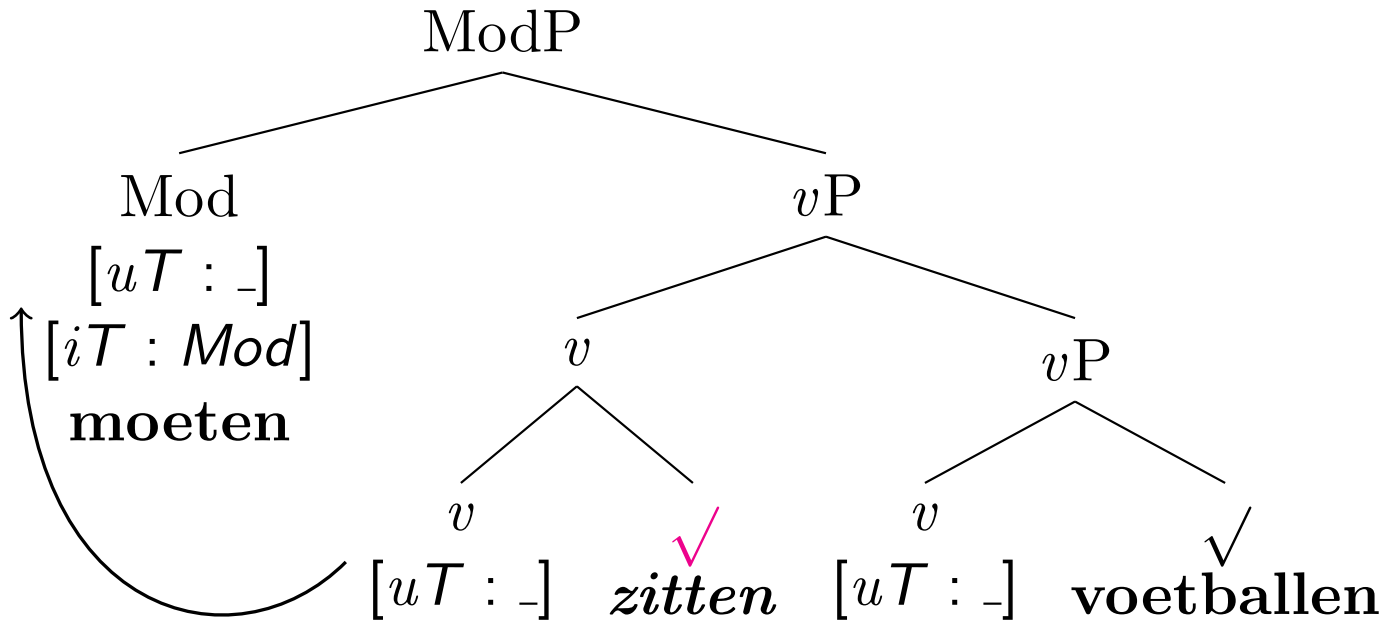
(46) **Alternative Agree step II (sem-lex stage I)**



- For some speakers, the [uT]-feature on the second v can also Agree with the unvalued [uT]-feature on Mod rather than [iT]-feature

Case study: the analysis

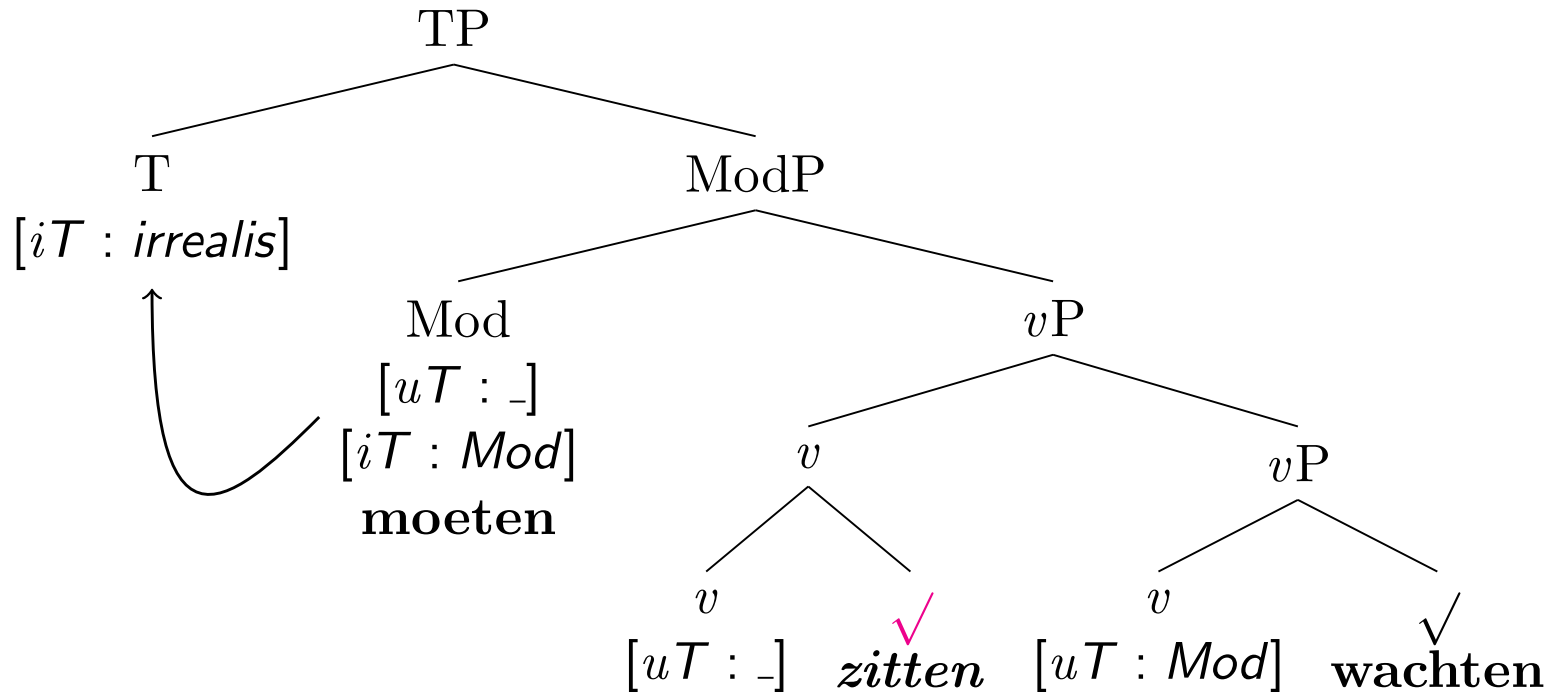
(46) **Alternative Agree step II (sem-lex stage I)**



- For some speakers, the [uT]-feature on the second v can also Agree with the unvalued [uT]-feature on Mod rather than [iT]-feature
- No valuation takes place, but the feature chain of the two [uT]-features gets extended to the [uT]-feature on Mod

Case study: the analysis

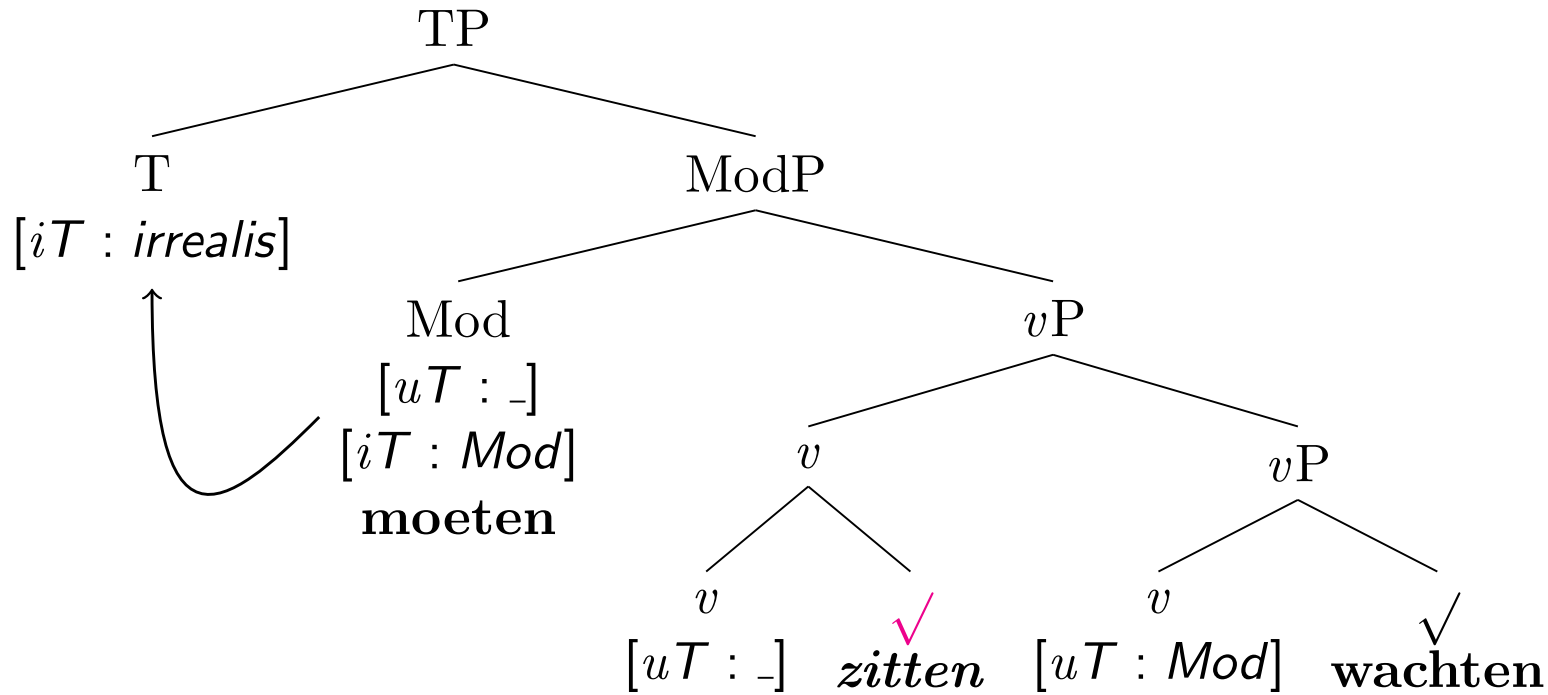
(47) **Agree step III (sem-lex stage I)**



- The next step proceeds as usual

Case study: the analysis

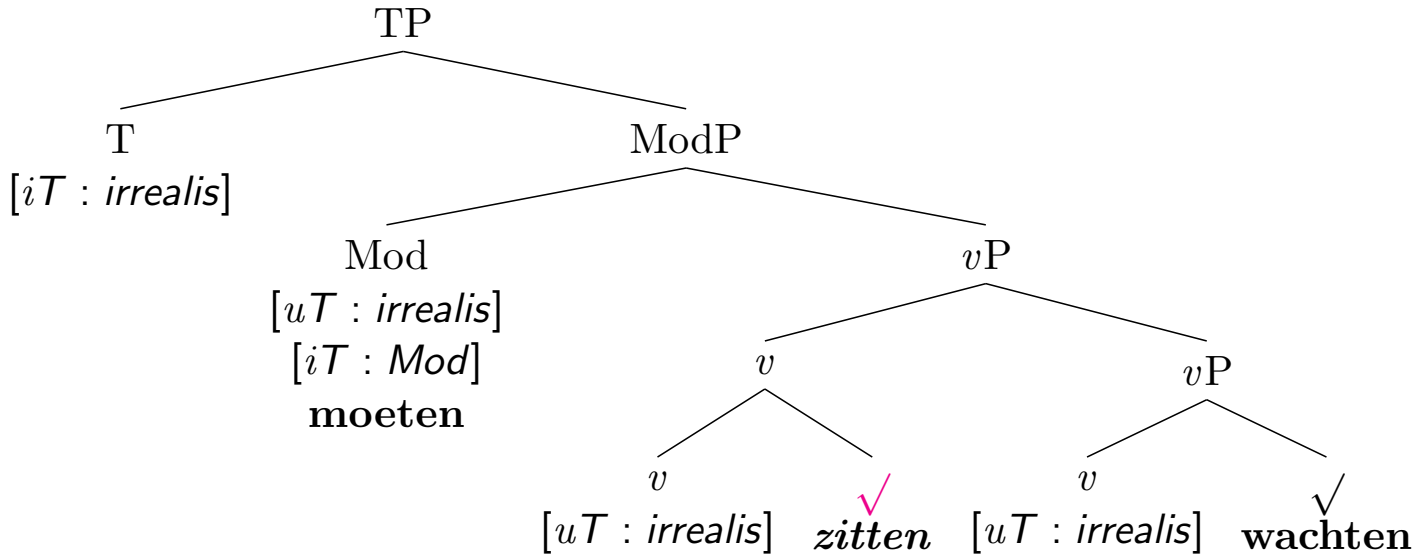
(47) **Agree step III (sem-lex stage I)**



- The next step proceeds as usual
- The [uT]-feature on Mod probes up and Agrees with [iT:irrealis] on T

Case study: the analysis

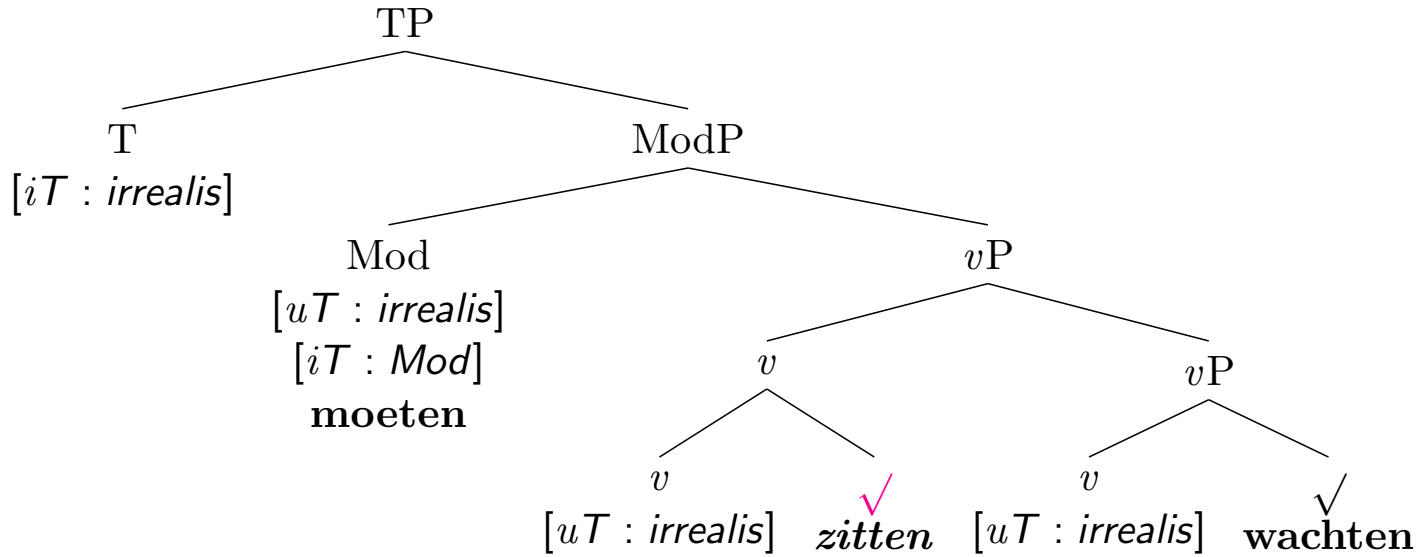
(48) **Alternative valuation (sem-lex stage I)**



- As a result of the feature chain between the three [uT]-features, they all get valued for [irrealis]

Case study: the analysis

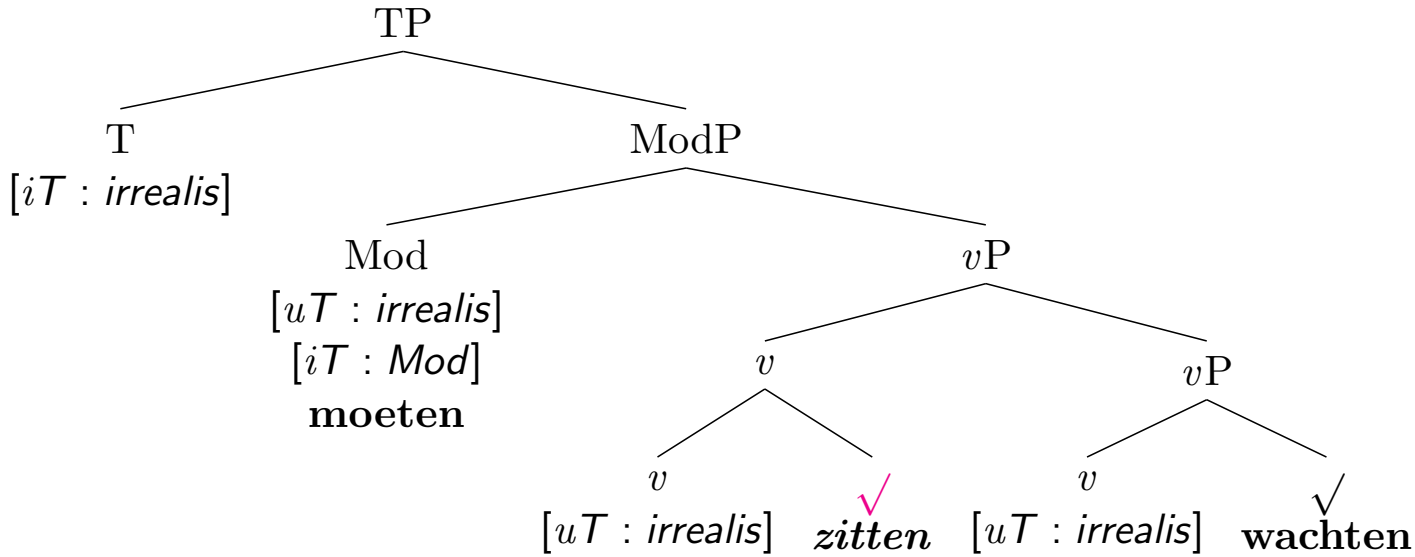
(48) **Alternative valuation (sem-lex stage I)**



- As a result of the feature chain between the three [uT]-features, they all get valued for [irrealis]
- *Te* can be spelled out when v has a [uT:irrealis]-feature: we expect high-*te*, low-*te* and *te*-doubling to be able to occur

Case study: the analysis

(48) **Alternative valuation (sem-lex stage I)**



- As a result of the feature chain between the three [uT]-features, they all get valued for [irrealis]
- *Te* can be spelled out when v has a [uT:irrealis]-feature: we expect high-*te*, low-*te* and *te*-doubling to be able to occur
- This is indeed the case, though with low to very low frequencies

Case study: the analysis

- Given the alternative Agree step, it is expected that these patterns occur much less frequently than *te*-drop

Case study: the analysis

- Given the alternative Agree step, it is expected that these patterns occur much less frequently than *te*-drop

Case study: the analysis

- Given the alternative Agree step, it is expected that these patterns occur much less frequently than *te*-drop
- Furthermore, low degree of intra-speaker variation is expected, as I assume that not even all speakers to allow the alternative way to Agree

Conclusion and outlook

- I have proposed a novel analysis to account for semi-lexical elements

Conclusion and outlook

- I have proposed a novel analysis to account for semi-lexical elements
- This analysis keeps the devision between roots and functional features in the lexicon

Conclusion and outlook

- I have proposed a novel analysis to account for semi-lexical elements
- This analysis keeps the devision between roots and functional features in the lexicon
- But proposes that semi-lexicality is syntactic, due to roots sometimes being able to occur in the functional domain of another root

Conclusion and outlook

- I have proposed a novel analysis to account for semi-lexical elements
- This analysis keeps the devision between roots and functional features in the lexicon
- But proposes that semi-lexicality is syntactic, due to roots sometimes being able to occur in the functional domain of another root
- I have argued for two stages of semi-lexicality, which are consecutive steps on a grammaticalisation path

Conclusion and outlook

- I have illustrated the analysis with a case study on two Dutch verbs, *hoeven* and *zitten*

Conclusion and outlook

- I have illustrated the analysis with a case study on two Dutch verbs, *hoeven* and *zitten*
- I have argued that *hoeven* is grammaticalising from the first to the second stage of semi-lexicality, which results in a high degree of morphosyntactic optionality

Conclusion and outlook

- I have illustrated the analysis with a case study on two Dutch verbs, *hoeven* and *zitten*
- I have argued that *hoeven* is grammaticalising from the first to the second stage of semi-lexicality, which results in a high degree of morphosyntactic optionality
- For *zitten*, I have argued it is uniformly in the first stage of semi-lexicality, which results in a lower degree of morphosyntactic optionality

Conclusion and outlook

- In my dissertation, I have shown that the main proposal can also be applied to cases of pseudocoordination in Afrikaans (Cavirani-Pots 2020)

Conclusion and outlook

(49) Ek het **sit/staan/lê/loop** **(*en*) lees.**
I have sit/stand/lie/walk and read.
'I have been reading.'

Conclusion and outlook

(49) Ek het **sit/staan/lê/loop** **(*en*)** **lees.**
I have sit/stand/lie/walk and read.
'I have been reading.'

- These verbs are semi-lexical

Conclusion and outlook

(49) Ek het **sit/staan/lê/loop** **(*en*)** **lees.**
I have sit/stand/lie/walk and read.
'I have been reading.'

- These verbs are semi-lexical
- *Functional:* they can indicate progressive or andative aspect

Conclusion and outlook

(49) Ek het **sit/staan/lê/loop** **(*en*)** **lees.**
I have sit/stand/lie/walk and read.
'I have been reading.'

- These verbs are semi-lexical
- *Functional:* they can indicate progressive or andative aspect
- *Lexical:* they are not (completely) compatible with all lexical verbs, thus retaining their own lexical semantics

Conclusion and outlook

- I have shown that they also show a high degree of morphoysyntactic variation

Conclusion and outlook

- I have shown that they also show a high degree of morphoysyntactic variation
- Higher degrees of morphosyntactic variation correlate with higher degrees of semantic bleaching

Conclusion and outlook

- I have shown that they also show a high degree of morphoysyntactic variation
- Higher degrees of morphosyntactic variation correlate with higher degrees of semantic bleaching
- I have argued that this is an indication for a shift from stage I to stage II

Conclusion and outlook

- I am working on two other data sets on which I want to apply the semi-lexicality proposal

Conclusion and outlook

- I am working on two other data sets on which I want to apply the semi-lexicality proposal
  1. Elative compounds (*bereleuk* lit. 'bear nice'), in which the first element seems to have developed from a noun into a semi-lexical element, or 'affixoide', but is not a prefix (yet)

Conclusion and outlook

- I am working on two other data sets on which I want to apply the semi-lexicality proposal
  1. Elative compounds (*bereleuk* lit. 'bear nice'), in which the first element seems to have developed from a noun into a semi-lexical element, or 'affixoide', but is not a prefix (yet)
  2. Verbal pseudo compounds (*achtervolgen* lit. 'behind follow'), in which the first element seems to have developed from a preposition, adjective or noun into a semi-lexical element or 'affixoide', but is not a prefix (yet)