1

THE PHONOLOGY OF SYNTHETIC COMPOUNDS

CRISSP SEMINAR JUNE 7, 2023

HARRY VAN DER HULST, UNIVERSITY OF CONNECTICUT

2 OUTLINE OF THE TALK

1. What is a synthetic compound?
2. Many different analyses
3. Obstacles for the suffixation analysis
4. Theoretical preliminaries
   1. Morphology by itself
   2. Different phonological hierarchies
5. Empirical support for the suffixation analysis
6. What is '?'
7. Ambivalent constructions are separable
   1. Evidence from denominal SCs
   2. Evidence from separable verbs ('samenkoppelingen')
8. Does 'non-native' suffixation proof me wrong?
9. Conclusions and prospects

3 THE POINT OF THIS TALK

Within Generative Grammar, synthetic compounds have been a fruitful subject for analysis and discussion for many decades.

Arguments to support specific theories have focused on either syntactic or semantic considerations.

What I add this debate today: phonological considerations.

4 1 DEVERBAL SYNTHETIC COMPOUND

Consider the word *beer drinker*. For a long time, words of this type have been called *synthetic compounds* and at least three types of analyses have been defended:

(1) a. [[beer drink] er] 'the suffixation analysis' (adhering to binarity)

b. [[beer] [[drink] er]] 'the compound analysis' (adhering to binarity)

c. [[beer][drink][er]] 'the ternary analysis'

Beer drinker is called a deverbal synthetic compound

5 DENOMINAL SYNTHETIC COMPOUND

Now consider the word *broadshouldered*. For a long time, words of this type have also been called *synthetic compounds* and again at least three types of analyses have been defended:

(2) a. [[broad shoulder] d] 'the suffixation analysis' (adhering to binarity)

b. [[broad] [[shoulder] d]] 'the compound analysis' (adhering to binarity)

c. [[broad][shoulder][d]] 'the ternary analysis'

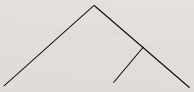
Broad shouldered is called a denominal synthetic compound

6 THREE DIFFERENT STRUCTURES

suffixation structure

beer drink er

compound structure



beer drink er

ternary structure

beer drink er

I will here focus on the first two (binary) structural analyses.

(Note: Compounds also called 'root' compounds, but not 'root' in the sense of Distributed Morphology, DM.)

7 WHAT IS A SYNTHETIC COMPOUND?

The Dutch term for synthetic compound is samenstellende afleiding.

Afleiding means 'derivation', while samenstelling means 'compound'. The present participle form (used as an adjective) samenstellende indicates that 'in the act of suffixation a compound is created'.

This traditional terminology seems to imply a suffixation analysis.

One could perhaps say that the English term synthetic compound (SC) suggests a similar analysis.

8 THE PRIMARY DIAGNOSTIC

In both traditional grammars and later works, a diagnostic property of SCs is that in many cases one or both parts do not exist as actual words (which does not imply that these parts are not possible words):

Suffixation analysis: *broad shoulder, *beer drink

Compound analysis: *shouldered, *nem-er

Beer drink is a special case because NV compounds are said to NOT form a possible productive compound type (in English and Dutch, at least). Existing cases such as *to baby sit* are explained as backformations.

9 2 MANY DIFFERENT ANALYSES

While SCs have been identified in many traditional works, a trend setting proposal for analyzing deverbal SCs was the LI paper by Tom Roeper and Muffy Siegel from 1978.

Many other proposals followed in works by Margaret Allen, Rudolf Botha, Lisa Selkirk and various Dutch morphologists (Willem Meys, Michael Moortgat, Geert Booij and Ariane van Santen, …).

A critical discussion of these proposals (up to 1985) is offered in Jack Hoeksema's 1985 dissertation on SCs.

10 SEMANTICS SUPPORT THE SUFFIXATION ANALYSIS?

The semantic scope argument

Several linguists have supported the suffixation analysis because *beer drinker* means 'someone who *drinks beer*' and *broad-shouldered* means 'someone with *broad shoulders*'.

Especially in de deverbal SCs, it has been argued that the noun is an argument of the verb (theme, agent, …)

11 THE COMPOUND ANALYSIS IS ALWAYS 'POSSIBLE'

But many others argue that one can get the semantics right when the compound analysis is adopted.

Beer (in *beer drinker*) being an argument can be explained with a theory of *inheritance*, but some reject this, saying that the argument reading, falling within the scope of the relation R ('N1 has something to do with N2), is simply the most natural reading.

Broad (in *broad shouldered*) fills in a 'semantic gap' that a word like 'shouldered' necessarily has because just 'having shoulders' is uninformative. Saying that someone is *shouldered is strange because everyone has shoulders. Such words have* a 'semantic gap' that is filled in morphologically or semantically. E.g., saying that some is *muscled* means 'having a lot of muscles'.

See van Ariane van Santen's clear defense of this view. Also see Jack Hoeksema who calls the scope argument 'naïve' and proposes a formal semantics that takes care of the scope issue.

Additionally, the compound analysis is seemingly supported by the phonological structure argument: *drink* and *-er* form a 'phonological word'. Same for *shoulder* and *-d*. I return to that point later.

12 THE COMPOUND ANALYSIS PREVAILS?

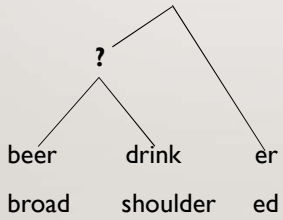
The compound analysis has predominated the earlier proposals, with Botha being the most explicit proposal of the suffixation analysis, with suffixation to (some kind of) phrasal structures.

Some approaches are difficult to classify (e.g., Roeper & Siegel, Distributed Morphology).

I will provide evidence from Dutch, involving the location of word accent, that the suffixation analysis is correct in at least some cases.

13 3 OBSTACLES FOR THE SUFFIXATION ANALYSIS

What is '?'



A predominant intuition has been that it is some kind of phrase (Marchand, Botha). There are not many proponents of the idea that '?' is a (root) compound. (beer drink is in fact an impossible compound verb). Heidi Harley calls it a 'root phrase'…

What if '?' is at neither a (compound) word nor a phrase…. (hold that thought)

14 PREVENTING A TERMINOLOGICAL CONFUSION…

The 'suffixation' analysis of SCs possibly involves the notion of a 'compound' as the base for affixation but this should not be confused with what I call the 'compound analysis' of SCs which treats the alleged SC as a whole as an ordinary compound.

That said, I will argue in my suffixation approach to SCs the base that suffixes attach to is not a compound (nor a phrase).

…..Got that?

15 LEVEL ORDERING ARGUES AGAINST THE SUFFIXATION ANALYSIS?

- Level ordering: affixation to compounds and phrases is not allowed.
- Clear examples of affixation of compounds:
  - Geestdrift > geestdriftig 'enthusiastic'
  - Hartstocht > harstochtelijk 'passionate'
- Clear examples of affixation to phrases:
  - hete lucht ballon 'hot air ballon'
  - zwart geld circuit 'black money circuit
- Phrases as left members in compounds:
  - No phrase constraint
  - First sister principle
  - Nuclear stress rules

16 2 THEORETICAL PRELIMINARIES: 2.1 IS MORPHOLOGY SYNTAX?

Many linguists who have contributed more recently to the 'synthetic compound debate': Heidi Harley, Hagit Borer, Anna Maria di Sciullo, Peter Ackema and Ad Neeleman, Chiara Melloni, Paul Kiparsky, Dieter Wunderlich and more…

A major issues that divides linguists (apart from whether they pursue some of form of the suffixation analysis or compound analysis) is whether morphology is distinct from syntax?

Despite the name *Distribution* ***Morphology***, DM does not recognize a distinction.

Most others in my list support the notion that morphology and syntax are distinct modules, but with various overlaps.

17 'MORPHOLOGY BY ITSELF'

Or: 'Morphology from below'

I regard morphology and syntax as different modules, which means that, despite many analogies, morphology does not need to full apparatus of syntax. 'My' morphology is thus not DM-style, but more 'traditional' ('Minimalist Morphology', MM), as promoted in Kiparsky's work (following Wunderlich and Stiebels).

Morphology combines signs, i.e. units that have form and meaning (and usually a category label)

Complex words result from hierarchically combining morphemes and words, producing a minimal amount of structure and making no use of movement transformations.

(Note: 'non-native' morphology works differently, i.e., uses 'affix substitution'. See later.)

18 2. THEORETICAL PRELIMINARIES: 2.2 TWO PHONOLOGICAL STRUCTURES

I distinguish two phonological hierarchies (following Jorgen Rischel's *deep and surface phonology* and also informed by work by Aditi Lahiri and Frans Plank)

Deep phonology: **Phonotactic structure**
Preserves morphological structure, augmented with S/W labelling representing prominence.
Accounts for allomorphy [~ direct reference phonology]

Surface phonology: **Prosodic structure (domains: α, β, γ)**
Preserves word/phrasal accent, augmented with rhythm. Accounts for allophonic rules

This is not like SPE-style 'underlying/deep structure' and 'surface structure' in a derivational theory; Deep and surface phonology are 'co-phonologies' that do not stand in a derivational relation. Both levels are subject to their own constraints and rules.

19 WHY USE S/W LABELING? (RETRO?)

- The idea to represent prominence in terms of S/W labeling the morpho-syntactic structure goes back to Liberman (1975). Actually, Jorgen Rischel used '+' and '-' instead of S/W before Liberman's work.
- It was further developed in work by Heinz Giegerich during the 1980s who proposed 'metrical transformations' that would alter the structure and labelling under (eu)rhythmic pressures.
- This approach is continued in the 'metrically-interpreted syntactic structure' theory of Jean Roger Vergnaud and Anne Maria Luisa Zubizarreta (for phrasal accent).
- Arguably the **metrical grid** in the theory of Liberman and Prince captures the (eu)rhythmic structure of surface phonology. But then some abandoned the metrical structure (Prince) and others the grid (Kiparsky)
- For unknown reasons (to me) the original ideas were replaced by the Selkirkian prosodic hierarchy.
- In my approach, S/W structure is 'deep phonology', whereas Selkirk's prosodic hierarchy is 'surface phonology'. (In a sense, the 'classic' Selkirkian prosodic hierarchy is parceled out over two levels.)

20 PRIMARY ACCENT FIRST APPROACH TO WORD PROMINENCE

Consistent with distinguishing deep and surface phonology/levels, in all my previous work on word prominence, I have promoted an approach (PAF) that is Metrical Theory (MT) in 'reverse'.

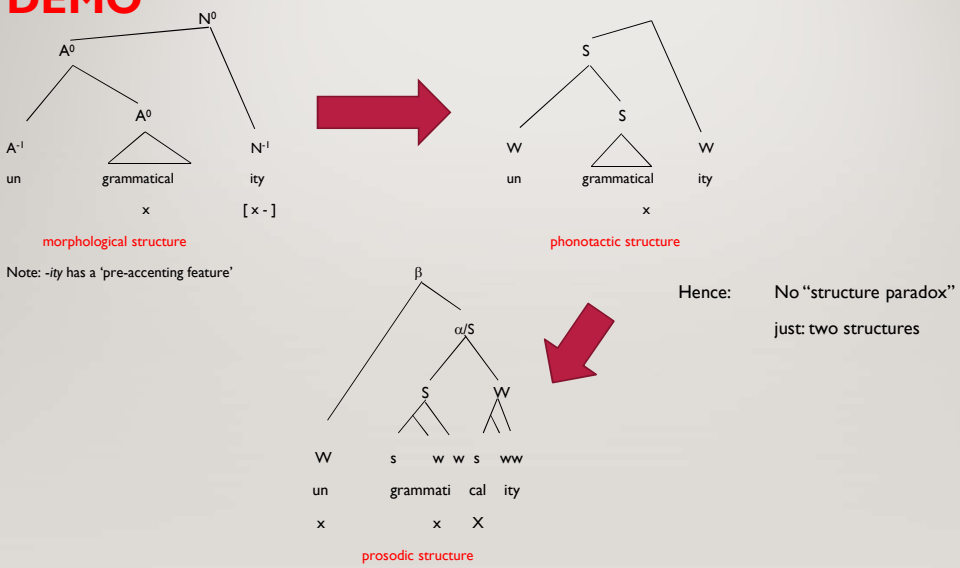
MT: First assign foot structure (=rhythm), then select one foot to represent 'primary accent/stress'.

PAF:

First: determine word accent (aka primary accent) (phonotactic structure)

Second: assign rhythm, respecting word accent (prosodic structure)

21 DEMO



Note: *-ity* has a 'pre-accenting feature'

Hence: No "structure paradox"

just: two structures

22 LINEARIZATION

I assume that the morphological AND the phonotactic structure are non-linearized; they are 'mobiles'.

Linearization is the first step toward the prosodic structure feeding syllabification.

I will show later that linearization involves a notion of Head Adjunction (*hold that thought*).

23 WORD ACCENT SYSTEMS

Phonologically predictable accent: dependent on *domain size* (bounded/unbounded), *domain edge* (left/right), *syllable weight, with diacritic accents for exceptions* (following linearization)

Lexical accent (lexical accent system): diacritic accents (accented and unaccented morphemes)

Resolution strategies (especially noticeable in the lexical accent systems) (Bogomolets 2020)

Linear (phonological): first/last, last/first, last/last, first/first (following linearization in lexical accent systems, maybe even in prosodic structure if purely based on phonological weight).

hierarchical (morphological): directly dependent on S/W labelled morphological structure

Focus of this talk: the role of S/W labelling in hierarchical accent resolution, with specific reference to synthetic compound (in Dutch)

24 5 EMPIRICAL SUPPORT FOR THE AFFIXATION ANALYSIS

I will suggest that the suffixation analysis must be the correct one for at least some SCs, based on *phonological evidence* that involves *word accent*, which, as far as I know, has never been considered. As we will see, there are some specific data in Dutch (which do not have analogues in English, but possibly in other languages), that can only be explained if we adopt a suffixation analysis and hierarchical resolution for what appear to be synthetic compounds.

25 THE DUTCH FACTS

Consider the example *bierbrouwerij* 'beer brewery'. There are two possible (binary) analyses:

(3) a. [[bier brouw] erIJ] 'the suffixation analysis'
b. [[bIEr] [[brouw] erij]] 'the compound analysis'

As usual, the compound analysis in (3b) is possible for the reading: 'a brewery that has something to do with beer' (thus likely one where they brew beer).

26 ANOTHER READING

However, in this case both structures are possible, but they correspond to different meanings and accentuation.

Structure (3a) means 'the activity of brewing beer':

(4) a. BierbrouwerIJ is een winstgevende activiteit

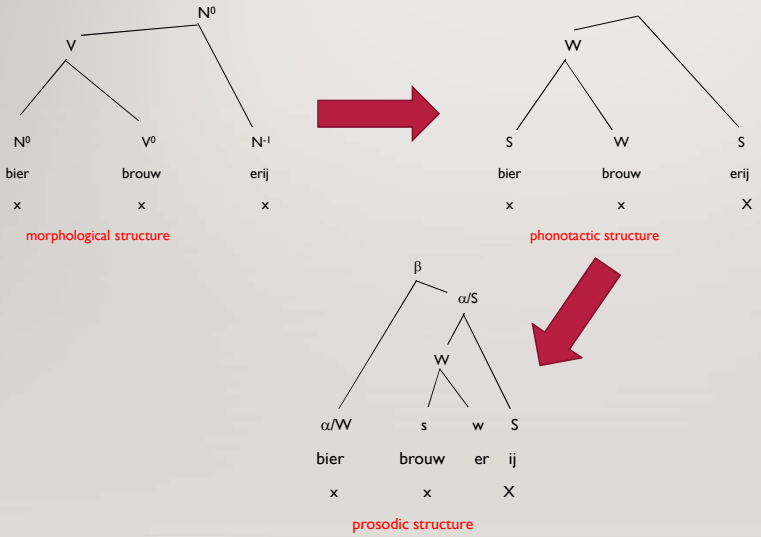
'brewing beer is a profitable activity/business'

Structure (3b) means 'a brewery as an entity':

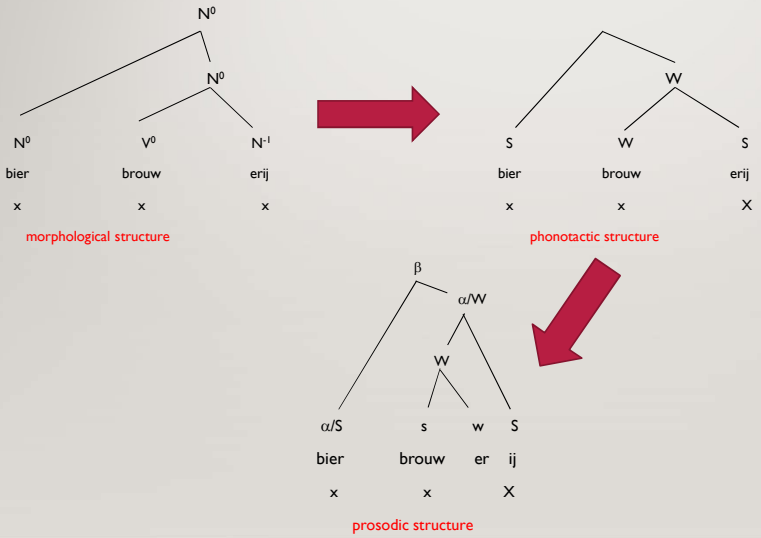
b. Dat gebouw aan de overkant is een bIErbrouwerij

'that building across the street is a beer brewery'

27 *BIERBROUWERIJ* ('ACTIVITY' READING)



28 *BIERBROUWERIJ* ('ENTITY' READING)



29 THE COMPOUND ANALYSIS DOES NOT ALWAYS WORK

The existence of two accentuations, corresponding to two different readings, indicates that there are two structures, one being an affixation structure, the other a compound structure.

Word accent in Dutch is thus dependent on the S/W labelling of morphological structure, which means that resolution is hierarchical.

Can we find other cases in Dutch or other languages where difference in accentuation (or other phonological effects) correspond to SCs and 'root' compounds? I think so…

30 SEMANTIC AMBIGUITY

One can argue that the English word *beer brewery* is simply semantically ambiguous between an 'entity' and an 'action' reading, which also correspond to two different structures, but since *-ery* is not accented you can't 'hear' the difference.

The issue of ambiguity has recently been addressed:

"The basic ambiguity between a derivational and compositional analysis of SCs can be resolved by assuming dualism and superposition of suffixing and compounding, similar to Albert Einstein's assumption of duality and superposition of waves and particles (photons) for light." (Elisa Mattiello and Wolfgang U. Dressler 2022).

[**Superposition** = the ability of a quantum system to be in multiple states at the same time until it is measured.]

"We investigate compounds headed by suffix-based deverbal nouns and propose that they are ambiguous between true synthetic compounds, which include verbal structure, and root compounds." (Iordăchioaia, G. & Alexiadou, A. 2022)

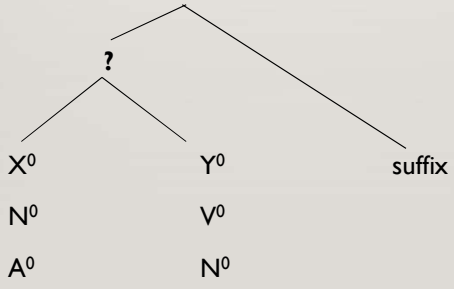
31 SEMANTIC DIFFERENCES BETWEEN 'ROOT' COMPOUNDS AND SYNTHETIC COMPOUNDS

While my Dutch data suggest semantic differences between root compounds and synthetic compounds, more needs to be said about what semantics can tell us about the difference between 'root' compounds and synthetic compounds … Albeit not here…

Note: Hagit Borer addresses the related issue of semantic differences between Argument Structure Nominals [like *the driving of the truck*]) and the event structure and properties of Synthetic Compounds [like *truck driving*] traditionally assumed to derive from the incorporation, into a derived nominal, of the internal argument of the verb. A lot of work on these issues was done by Teun Hoekstra and his students during the 1990s.

32 EXTENSION OF THE SUFFIXATION ANALYSIS TO DENOMINAL SYNTHETIC COMPOUNDS

Given that the affixation structure is supported for deverbal synthetic compounds, (*bier brouwerij / beer brewery*), it can also be adopted for the deverbal type but also the denominal type (*breedgeschouderd / broad shouldered*):



33 6 60 MILLION $ QUESTION: WHAT IS '?'

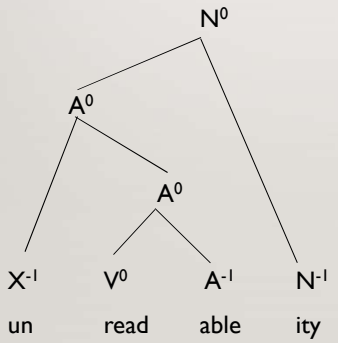
Several linguists have worried about the nature of the constructions that form the basis for SCs (according to the suffixation hypothesis).

- If it's a phrase, it is subject to various restrictions that prevent the availability of 'full phrases' (cf. Botha's approach).
- If it is a compound (i.e., a word, X^0), one issue is the unavailability of NV compounds (at least in English and Dutch).
- More importantly: if it is a compound why are the parts 'separable' (in the morphology and in the syntax); see below.

34 UNDERSPECIFATION TO THE RESCUE

- My approach aligns with the approach of Ackema and Neeleman, making use of the notion underspecification: '**?**' is neither a word nor a phrase. (Peter Ackema and Ad Neeleman: "complex lexical items can be underspecified in various ways: one type of underspecification concerns their locus of realization (in syntax or morphology)"
- Formalization: For morphologically structured words, I adopt an X-bar type notation for node labels (using superscripts), originally proposed in Hoekstra, Moortgat & van der Hulst (1980). Word as 'zero level' units (most linguists use that), and affixes as 'minus one': X^{-1}. Affixes are heads of complex words that (potentially determine the category of the word that they derive). Phrases are 'higher than 0', minimally level 2. (There are similar looking proposals, which are nevertheless different.)

35 AN EXAMPLE: *UNREADABILITY*



X = no categorial features (non-head)

36 UNDERSPECIFICATION TO THE RESCUE

The proposal (given this notational system) is that **?**-constructs are un(der)specified for a level.

Informally, I will refer to the underspecified constructs as ambivalent constructs (ACs).

(Mark Baker, p.c. called my appeal to an X-bar notation 'a little retro'.)

37 7 AMBIVALENT CONSTRUCTIONS ARE 'SEPARABLE'
7.1 EVIDENCE FROM DENOMINAL SCs

- Why not say that '?' has a compound structure, taking compounds to be words (X^0)?
- There is an interesting difference between denominal SCs in English and Dutch:
- The affix forming denominal SCs of the *broad shouldered* type is discontinuous and it 'wraps' around the N-part:

breed GE schouder D

38 DOES THIS SUGGEST A COMPOUND ANALYSIS?

I will assume that the suffixation analysis is the correct one for *breedgeschouderd* and that the discontinuous affix *ge-X-d* is attached to the unit [breed schouder]. However, as we can see, the prefixal part occurs in front of *schouder*, rather than in front of the whole unit (**gebreedschouderd*). Interestingly, this fact has been used to argue in favor of the compound analysis in (5b):

(5) a. [[breed schouder] ge-Xd] suffixation analysis
b. [[breed] [ge- schouder -d]] compound analysis

39 PHONOTACTIC STRUCTURE TO THE RESCUE

By adhering to the suffixation analysis, I will be committed to the proposal that ge-X-d 'wraps around' the word *schouder.*

Interestingly this indicates that [breed schouder] is not a compound word. Compare:

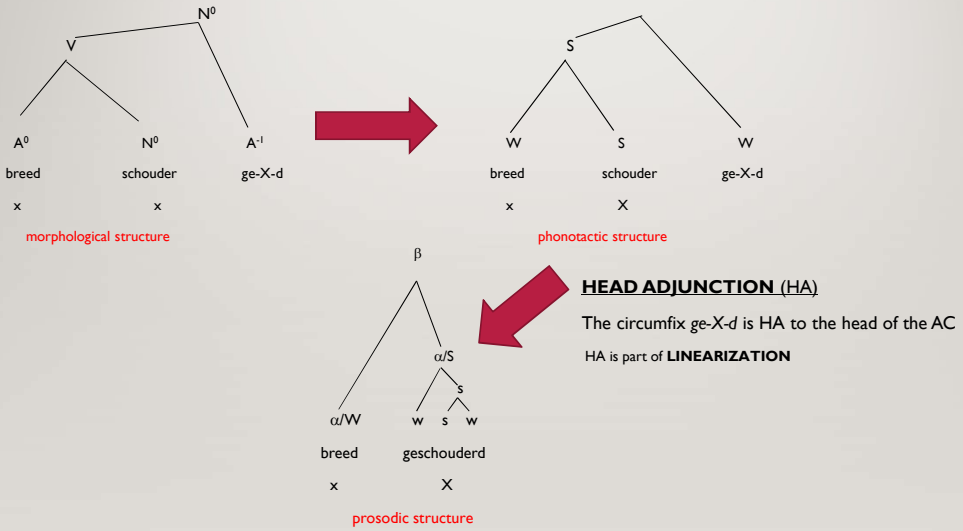
(6) (ge (hoofdtooi) d) *hoofdtooi* is a root compound 'head ornament'

I proposes that the discontinuous affix is properly attached in the *prosodic structure,* through head adjunction, which is part of linearization:

(7) ((breed)(geschouderd))

Assumption: A discontinuous affix wraps around the head X^0 that is a sister of the affix. For Head Adjunction to apply properly it is crucial that the AC does not have a zero level specification.

40 *BREEDGESCHOUDERD*



41 RULES FOR S/W LABELLING

Note that the AC [breed schouder] gets labeled W-S, which mirrors the W-S labelling in Noun Phrase that contain an adjective:

(8) hij heeft [brede schOUders]

'he had broad shoulders'

By the same reasoning the S/W labelling in ACs that form the basis of deverbal SCs is S-W, again as in verb phrases:

(9) Hij [drinkt bIEr] dat hij [bIER drinkt]

'he drinks beer' 'that he drinks beer'

That said, in specific cases, the S/W labelling can deviate from these patterns in the context of specific suffixes, which suggest that ACs are actually unspecified for prominence. Time permitting, I can talk about this.

42 7 AMBIVALENT CONSTRUCTIONS ARE SEPARABLE
7.1 EVIDENCE FROM SEPARABLE VERBS ('SAMENKOPPELINGEN')

The idea that the lexicon contains complex units that are neither words nor phrases, but can end up as being either, is strikingly confirmed by a class of complex verbs, called *separable verbs* (SVs) which has attracting A LOT of attention by both (mainly Dutch and German) morphologist and syntacticians, as well as Canadian linguists (Heather Newell).

The most widespread (and productive) type contains of a so-called particle (often homophonous with prepositions, but not always) and a verb.

Grammaticalization at work:

Historically, separable verbs are re-analyses of constructions that contain a verb and an adposition. The lexical status of SVs is consistent with many of them having a non-compositional meaning (with the particle having lost it prepositional meaning).

A further historical development has been that SVs would be come inseparable, with the particle becoming a prefix.

43 A TYPICAL SEPARABLE VERB

Op bellen 'to phone, to call up'

Root clause

Ik *bel* hem morgen *Op* (verbal part moves to 'second position'; *verb second*)

* Ik opbel hem morgen

Embedded clause (OV order assumed to be the 'basic order')

Dat ik hem morgen *Opbel*

dat ik hem *Op* wil *bellen* ~ dat ik hem wil *Opbellen* (either verb alone or part-verb combination raises; (*verb raising*); the latter case shows that *opbellen* can behave like a word in the syntax;).

44 SEPARABILITY IS ALSO APPARENT FROM PARTICIPLE AND INFINITIVAL FORMS

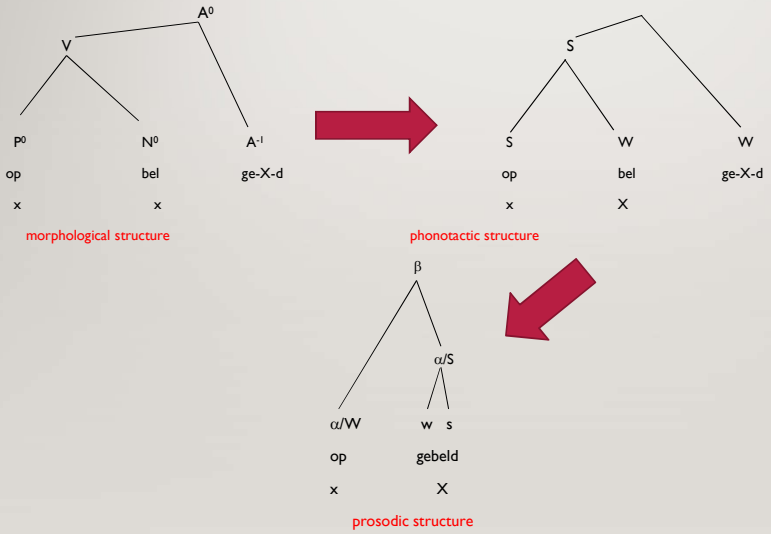
Hij heeft mij *opGEbelD* (uses a discontinuous suffix *g-X-d*)

'he has called me'

Hij heeft me gevraagd hem *op TE bellen* (*te* 'to' comes in between particle and verb)

'he asked me to call him'

45 *OPGEBELD*



46 SEPARABLE VERBS AS THE BASIS IN SYNTHETIC COMPOUNDS

When SVs are provided with a suffix, the resulting structure is a de facto SC:

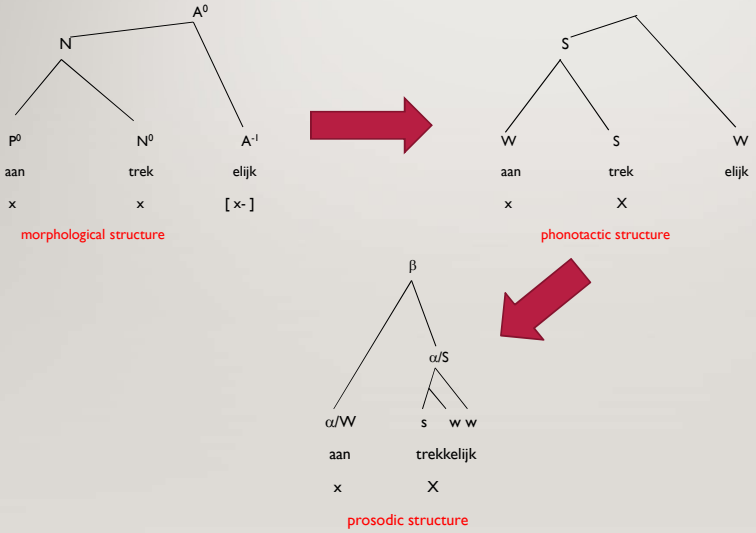
(8) aan trEk elijk 'attractive'

aan nEm elijk 'acceptable'

op mErk elijk 'remarkable'

The suffix *-elijk* is pre-accenting, which cases the S/W labelling of the AC to be W-S

47 *AANTREKKELIJK*



48 POTENTIAL EVIDENCE FOR *-ELIJK* BEING PERIPHERAL IN SYNTHETIC COMPOUNDS

Adjectives such as [[on][wens-elijk]] 'undesirable' display a rhythmic alternation, albeit optional:

een ónwenselijke rede ~ onwénselijke rede 'a undesirable speech'

deze rede is onwénselijk 'this speech is undesirable'

However, this rhythmic effect does not occur in SCs with *-elijk*:

een aantrékkelijke meid	* een áantrekkelijke meid	'an attractive girl' speech'
een toepásselijke grap	* een tóepasselijke grap	'an appropriate joke'
een uitdrúkkelijke wens	* een úitdrukkelijke wens	'an explicit wish'

The explanation could be that in *onwenselijk*, *-elijk* being at a inner 'cycle' cannot exercise its pre-accenting power on the word as a whole, whereas in SCs it can. This support the claim that in SCs *-elijk* is the outer layer, which thus support the affixation hypothesis.

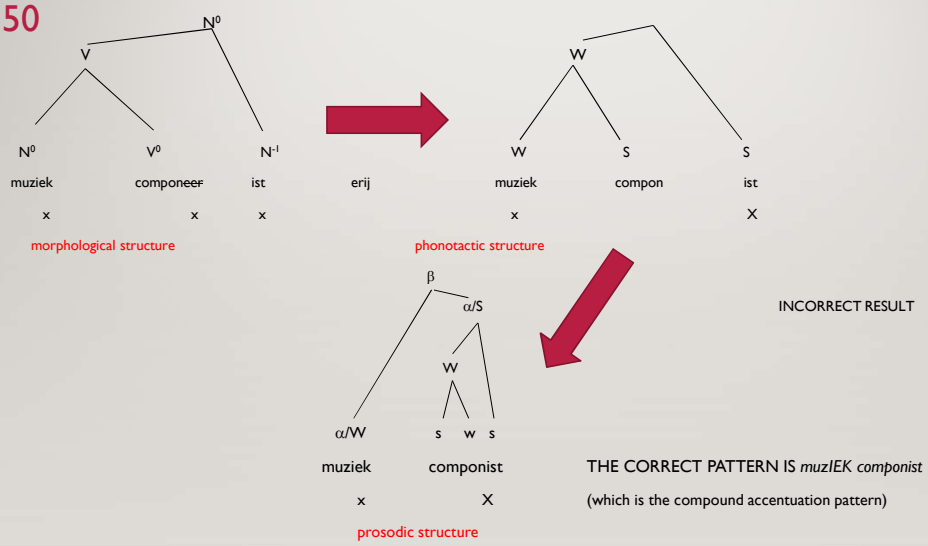
49

8. DOES NON-NATIVE SUFFIXATION PROOF ME WRONG?

I conclude with an apparent problem for my proposal. There are *cohering* ('non-native) suffixes that are stress-bearing and with those we can form SCs. Consider the example *muziek componist* 'music composer; someone who composes music'. Since the suffix *-ist* is accented, and added last, one would expect the word accent to be on the suffix. However, that is not the case. Whether this form is analyses as a regular NN compound or as a SC, the word accent is on the second syllable of *muziek* (which it where the word accent falls on the word when use in isolation).

MUZIEKCOMPONIST

50



51 NON-NATIVE SUFFIXATION IS NOT ADDITIVE: *AFFIX SUBSTITUTION*

'The only recursive rules of morphology, to my knowledge, are a matter of layering. You have a unit, you add something to it, you get a new unit, you add something to it, you get a new unit, and so on. And that is a very primitive sort of recursiveness'.

(Chomsky, 1982: 96-97)

However, some morphologists have argued for the operation of *affix substitution* (and against truncation).

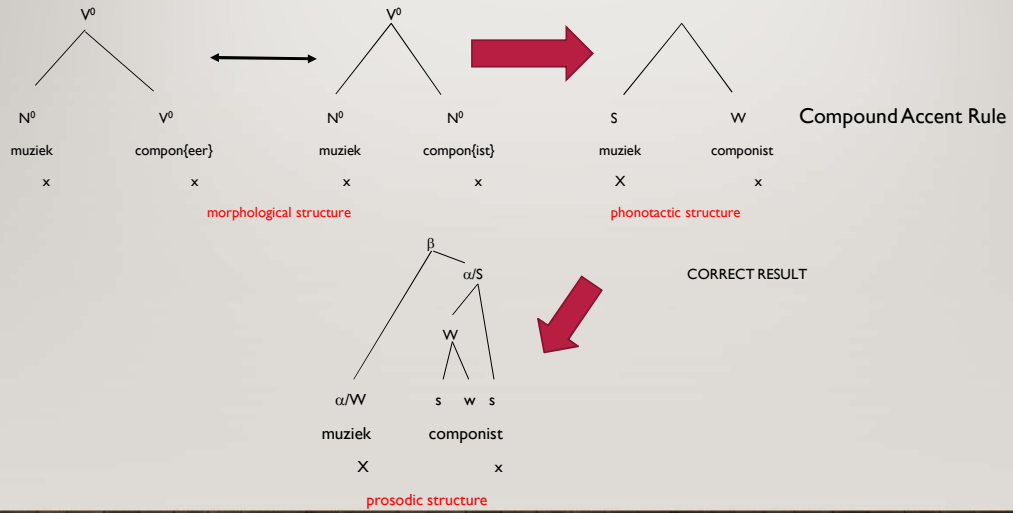
(Marle, van. 1985. *On the paradigmatic dimension of morphological creativity.* Dordrecht: Foris Publications.)

(Comparable data in English: *rúbbish remòval, hólocaust denìal, próject renewal)*

Affix substitution also occurs in native morphology:

veroveren 'to concur' heroveren 'to re-concur'

52 *MUZIEKCOMPONIST*



53

9 CONCLUSIONS AND PROSPECTS

- The *-erij* data show that synthetic compounds exist and that not all of them can be reduced to so-called 'root' compounds.
- Much work remains to be done, including on:
  - Why no de-adjectival SCs? (**horsefondness*, **verylateness*)
  - A cross-linguistic angle; do all languages have SCs? A typology of possible synthetic compounds.
  - Whether prefixes can form SCs (*enlargecage); in Dutch there is one prefix *her-* that can attach to SVs, which effectively forms a SC (*herinrichten 're-decorate'*).
  - Do all suffixes lend themselves to forming SCs? (Depends on their semantics; *-ig* vs. *-erig*).
  - Are there (specific uses of) suffixes that only occur in SCs? (Yes; type: *driewieler*).
  - Finding more cases where phonology discriminates between SCs and root compounds.

54

THANK YOU!