N-to-D movement within compounds and phrases

Referential compounding, *-s-* possessives and title expressions in Dutch

Marijke De Belder

Draft version, comments welcome
marijke.de.belder@uni-oldenburg.de

Abstract
In this article I argue that Dutch has referential compounds: compounds with a referential non-head, even though that claim seems to be a *contradictio in terminis*. They resemble the Dutch s-possessive in that their non-heads involve movement to a referential layer. However, unlike the possessive structures, the compounding structure contains head incorporation which results in word-hood. The referential compound is therefore an instance of the fourth logical possibility according to the two parameters defined by Borer (2011) for N-N combinations: compounding versus constructs and modificational structures versus referential ones. The article further discusses title expressions, such as *prince* Charles, which are argued to be referential construct states. Together with the syntactic structure of titles plus proper names, the referential compounds further contribute evidence to the idea that a ban on N-to-D movement for certain uniquely referring roots, such as *sun* and *Bronx* is extra-syntactic.

Keywords
compounding, construct state, N-to-D movement, possessives, Dutch

1. Introduction

In this article I present a previously unidentified type of compounding in Dutch, which, I argue, are compounds with referential non-heads as exemplified in (1) and I contrast this type of compounding with title expressions, as in (2), which I argue are referential construct states, and with *-s* possessives, as shown in (3):

(1) zon-s-hoogte
sun-S-height
'solar altitude'

(2) professor Bolleboos
professor Bolleboos
'professor Bolleboos'

(3) Marie-shypothese
Mary-S hypothesis
'Mary's hypothesis'

The referential compound is the fourth logical combination in Borer's (2011) classification, but it was assumed to be principally excluded and therefore non-existent. Borer (2011) argues, on the basis of Hebrew data, that N-N concatenations could either be constructs or compounds: constructs do not show syntactic head incorporation, whereas compounds do. Furthermore, the non-head could be either modificational or referential, a semantic distinction which is derived from a syntactic difference: modificational non-

heads occupy a classifier position, the referential ones include a determiner layer. Of the four possible combinations, three were argued to be realised in Hebrew: there is the modificational compound, the modificational construct and the referential construct. I aim to show that the fourth logical possibility occurs in Dutch. Compounds of the type *zonshoogte* 'solar altitude' are referential compounds and thus realise the fourth possible cell in table 1. I further argue that Dutch title expressions instantiate referential construct states. M-compounding patterns with Dutch nominal compounding.

Table 1: an inventory of possible N-N concatenations

	modificational non-head	**referential non-head**
construct	m-construct (Borer 2011)	r-construct (Borer 2011) & Dutch title expressions
compound	(m)-compounding (Borer 2011) & Dutch nominal compounding	r-compounding: *zonshoogte* 'solar altitude'

The -s possessive structure does not fall into the classification presented in table 1 as it is a possessive structure rather than a compounding or construct state structure, but it will become clear that it is relevant to compare referential compounding (r-compounding) to this syntactic structure.

The article is structured as follows. Section 2 contrasts the referential compound with other types of primary compounding in Dutch. Section 3 argues that the non-head of the referential compound indeed has referential properties. Section 4 shows that the referential compound is indeed a compound by contrasting the compound with the *-s-* possessive. Section 5 presents a referential construct state in Dutch. Section 6 discusses the consequences of these data for our understanding of N-to-D movement. Section 7 concludes.

2. Referential compounds versus other Dutch primary compounding types

De Belder (2017) and De Belder (to appear) identify two major Dutch primary compounding types: those of which the non-head is a bare root (De Belder 2017) and those for which the non-head merges with a classifier head (De Belder to appear). The non-head of the bare root compound can be associated with any category and never merges with an overt 'linking element' (examples taken from De Belder 2017:141):

(4) a. speur-hond
track-dog
'tracking dog'

b. drie-luik
three-panel
'triptych'

c. snel-trein
fast-train
'high-speed train'

d. achter-grond
back-ground
'background'

The non-head of the classifier type, in contrast, merges with a nominal classifying projection. It thus invariably gets a nominal interpretation. This type is perhaps the Dutch counterpart of what Borer (2011) identified as genuine compounding in Hebrew. The type is easily recognised by the presence of a so-called linking element *-s-* or *-en-* (NCM = nominal compound marker, i.e. the "linking element", examples taken from De Belder 2017:140):

(5) a. peer-en-boom
pear-NCM-tree
'pear tree'

b. varken-s-hok
pig-NCM-pen
'pig's pen'

In De Belder (to appear) it is argued that this 'linking element' realises the functional head of nominal classification. Furthermore, it is pointed out there that nominal classifying heads can in principle be realised by null morphemes as well and that there is actually dialectal evidence that this indeed happens in Dutch, probably also in the standard language. The reasoning is rather complex and lengthy, so I refrain from summarising it here as it would take us too far afield. According to the reasoning described there, the following compounds could be relevant examples (see De Belder to appear).

(6) a. siroop-fles
syrup-bottle
'syrup bottle'

b. klei-grond
clay-soil
'clay soil'

c. wol-draad
wool-yarn
'wool-yarn'

As a result, the classifier head in Dutch can be realised by means of ∅, *-s-* [s] or *-en-* [ə(n)]. Compounds of this type with a null marker are at the surface of course indistinguishable from compounds with a bare root as their non-head.

De Belder (to appear) discusses the well-known fact that the non-head of the classifying (i.e. modificational) compounding type selects the classifier head and its exponent. One can thus expect a quite regular selection between the non-head's lexeme and its classifying exponent of choice (∅, -s- or -en-) (see Botha 1968 and see De Belder to appear for a more nuanced discussion and the examples below):

(7) a. kat-**en**-luik
cat-NCM-shutter
'cat flap'

b. kat-**en**-voer
cat-NCM-food
'cat food'

c. kat-**en**-staart
cat-NCM-tail
'cat tail'

d. kat-**en**-bak
cat-NCM-box
'cat litter box'

(8) a. ezel-**s**-dracht
donkey-NCM-pregnancy
'long pregnancy'

b. ezel-**s**-bruggetje
donkey-NCM-bridge.DIMINUTIVE
'mnemonic'

c. ezel-**s**-oor
donkey-NCM-ear
'dog-ear'

The present article, however, draws attention to the fact that sometimes non-heads occur with an *-s-*, even though they would typically be restricted to bare root compounding or select zero marking or *-en-* as a classifying compound. I argue that when these non-heads are bare roots or when they select their typical exponent of the classifying head (here zero or *-en-*), they are modificational compounds, if they select the unexpected *-s-*, they are referential compounds. This results in the following inventory for Dutch compounds:

Table 2: an inventory of Dutch primary compounding

	bare root compounding	**noun class marking compounding**	**referential compounding**
"linking element"?	none: the non-head is a bare root	∅, [s] or [ə(n)]	[s]
modificational or referential?	modificational		referential

These three types are illustrated in the following examples:

(9) a. kreeft-woord
lobster-word
'palindrome'
[bare root c.]

b. kreeft-en-soep
lobster-EN-soup
'lobster soup'
[noun class marking c.]

c. Kreeft-s-keerkring
lobster-S-tropic
'Tropic of Cancer'
[referential compound]

So, referential compounds can be recognised by the appearance of an *-s-* where it is not immediately expected. I now discuss this in some more detail. De Belder (to appear) argued that non-heads with the semantics of kinship names and proper names are probably always bare root compounds:

(10) a. moeder-melk
mother-milk
'breast milk'

b. moeder-taal
mother-language
'mother tongue'

c. vader-beeld
father-image
'conception of the father figure'

d. Pieter-baas
Peter-boss
'black Pete'

e. Pieter-man
Peter- man
ancient coin with the image of Saint-Peter/
name of a certain fish (*Trachinus draco*)

However, one does find instances of kinship names and proper names followed by an *-s-* in Dutch compounding:

(11) a. moeder-s-kind
mother-S-child
'child too dependent on the mother'

b. vader-s-zijde
father-S-side
'father's side of the family'

c. Pieter-s-zoon
Peter-S-son
(family name)

I will argue that these compounds are referential compounds.

Similarly, the Dutch roots *dag* 'day' and *jaar* 'year' typically do not select a 'linking element', either because they invariably occur in bare root compounding or because they are instances of noun class marking compounds which select a zero exponent of noun class marking (I am principally unable to tell):

(12) a. dag-deel
day-part
'part of the day'

b. jaar-beurs
year-fair
'trade fair'

c. jaar-balans
year-balance.sheet
'annual balance sheet'

d. jaar-getijde
year-tide
'season'

Yet, again, one does find instances of exactly these roots selecting an *-s-*:

(13) a. verjaardag-s-taart
birthday-S-pie
'birthday pie'

b. nieuwjaar-s-feest
new.year-S-party
'New Year's party'

c. zondag-s-kind
sunday-S-child
'child born on a Sunday and born for good luck'

Again, the claim is that they are referential compounds.

Then there are noun class marking compounds of which the non-head selects *-en-* to realise the noun class marking:

(14) a. zon-en-bank[1]

b. kreeft-en-soep

c. naam-en-lijst

d. maan-en-stelsel

[1] The official spelling would be *zonnebank*.

sun-EN-bench	lobster-EN-soup	name-EN-list	moon-EN-system
'tanning bed'	'lobster soup'	'list of names'	'moon system'

Yet, again, one does find instances of these roots selecting an -s-, which I claim to be instances of referential compounding:

(15) a.	zon-s-hoogte	b.	Kreeft-s-keerkring	c.	naam-s-wijziging	d.	maan-s-verduistering
	sun-S-height		Cancer-S-tropic		name-S-change		moon-S-eclipse
	'solar altitude'		'Cancer Tropic'		'name change'		'lunar eclipse'

In sum, there seems to be a type of compounding in Dutch, which is characterized by the occurrence of an -s- as its linking element. To be entirely clear, I do not aim to argue that *all* instances of Dutch compounds with an *-s-* are referential compounds. As can be deduced from table 2, the *-s-* can be a realisation of a nominal classifying head as well.[2] Note that it is thus probably so that there are compounds which select both the *-s-* as modificational compounds and as referential compounds.[3]

3. The referentiality of the non-head

The non-referentiality of the non-head is often taken to be a defining criteria of compounding (see e.g. Borer 2005a:84; Borer 2011), so claiming that it can be referential is quite controversial. Then, why would I think that, indeed, I have discovered instances of compounds with a referential non-head? In this section I present arguments in favour of the referentiality of the non-head, in the next section I present arguments in favour of the compounding status.

Consider the following four arguments for the referentiality of the non-head. Firstly, the non-heads that occur in the referential compounds seem to belong to a specific group; they are highly reminiscent of the type of lexemes Longobardi (1994) identified as typically subject to N-to-D raising: proper names (*Pieter* 'Pete'), kinship names (*moeder* 'mother', *vader* 'vader'), names of days of the weeks (*zondag* 'Sunday') and holidays (*Nieuwjaar* 'New Year', *verjaardag* 'birthday'), and unique entities (*zon* 'sun', *Kreeft* 'Cancer (the constellation)', *maan* 'moon'). The first ingredient to become referential is thus in place: these lexemes are excellent candidates to raise to a D-layer syntactically where they can gain referential semantics.

Secondly, the referential compounding seems to imply universal uniqueness or uniqueness in the discourse context. For example, *Kreeftskeerkring* 'Tropic of Cancer' uniquely refers to the Cancer constellation. This contrasts with a modificational compound such as *kreeftensoep* 'lobster soup', which does not show the *-s-* and which does not imply unique reference to a lobster at all. In fact, it does not even imply the presence of lobster: one can easily to find a recipe for *kreeftensoep zonder kreeft* 'lobster soup without lobster' on Google. Similarly, *manenstelsel* 'moon system' does not refer to the Earth's unique moon, whereas *maansverduistering* 'lunar eclipse' does refer to the unique moon as we know it. For words such as *naam* 'name', *vader* 'father', *zondag* 'Sunday' and *verjaardag* 'birthday' the uniqueness is contextual rather than universal. This does not constitute a problem: it is known from research on definiteness that contextual uniqueness suffices for referentiality (Lyons 1999).

Thirdly, in the absence of uniqueness, the referential compound is excluded, as illustrated by the following minimal pair:

[2] I also do not exclude the possibility that there are yet other types of compounding in Dutch which have not been identified yet.

[3] The restricted inventory of Dutch 'linking elements' does not always allow one to distinguish between several types of compounding. Following the same reasoning, one may wonder how many different types of compounding English actually has, given that compounds in this language usually do not contain a linking element.

(16) zon-s-hoogte
sun-S-height
'solar altitude'

(17) * ster-s-hoogte
star-S-height

Fourthly, not only the uniqueness, but the referentiality itself is implied. Compare the following contrast. The following fully acceptable dialogue illustrates the classical, familiar non-referentiality of the non-head of a modificational compound (*hondenmand* 'dog bed'):

(18) A: We kopen een hondenmand.
B: Oh, heb jij een hond?
A: Nee, eigenlijk niet, we gaan de mand gebruiken voor onze kat.
B: Ja, je hebt gelijk, dan ligt ze wat ruimer.

'A: We are buying a dog bed.
B: Oh, do you have a dog?
A: No, actually not, we are going to use the bed for our cat.
B: Yes, I see, it will be a bit more spacious for her then.'

For the referential compound *verjaardagsfeest,* below, however, a dialogue parallel to the one in (16) is excluded -or at least very odd- due to the fact that the existence of the birthday is actually implied:

(19) #A: We organiseren een verjaardagsfeest.
B: Oh, is er een verjaardag?
A: Nee, eigenlijk niet, we organiseren het feest voor een huwelijk.
B: Ja, je hebt gelijk, dat is vast goedkoper.

'A: We are organising a birthday party.
B: Oh, is there a birthday?
A: No, actually not, we are organising the party for a wedding.
B: Yes, I see, it's probably cheaper.'

I conclude that the non-heads of these compounds have unique reference. The consequence is that the long-held believe that the non-head of a compound is by definition non-referential may be falsified.

4. Referential compounds versus the *-s-* possessive construction

The present section argues that the referential compound has true compounding status (and, as such, word-hood status) by comparing it to a phrase with similar properties: the Dutch *-s* possessive construction (see also De Belder 2009).[4]

The Dutch -s possessive resembles the English saxon genitive, but it is more restricted. In Dutch, the -s possessive construction is restricted to kinship nouns and proper names. Hence, the examples in (20) are excluded, whereas the ones in (21) are fine.

[4] I am using the term 'construction' in a descriptive, non-theoretical sense.

(20) a. * gisteren-s arrangementen
yesterday-S arrangements

b. * vrouw-s theorie
woman-S theory

c. * een vrouw-s theorie
a woman-S theory

d. * de vrouw-s theorie
the woman-S theory

(21) a. Marie-s hypothese
Mary-S hypothesis
'Mary's hypothesis'

b. papa-s auto
daddy-S car
'daddy's car'

When one studies the data in somewhat more detail, the generalisation is that the Dutch -s possessive construction is restricted to those nouns which can undergo N-to-D raising in argument position. Kinship nouns and proper names are of course roots that have the appropriate semantics to undergo N-to-D-raising, hence they can occur without an overt determiner in argument position, suggesting that the root itself has raised to D (Longobardi 1994):

(22) Ik ontmoette Marie in Parijs.
I met Mary in Paris.
'I met Mary in Paris.'

b. Ik zag papa.
I saw daddy
'I saw daddy'.

If the noun cannot undergo N-to-D-raising in argument position, the *-s* possessive construction is excluded. The contrast between the examples (23) and (24) shows that the noun (i.e. the root in the structurally nominal position) *zon* 'sun' requires an overt determiner in Dutch in argument position. Example (25) shows that the noun cannot occur as the possessor in the *-s* possessive construction. The proper names of rivers illustrate the same fact. They cannot occur without an overt determiner, and they do not occur as the possessor in the *-s* possessive either (see section 6 for an account):

(23) Ik zie de zon.
I see the sun
'I see the sun.'

(24) # Ik zie zon.[5]
I see sun

(25) * zon-s zachte warmte
sun-S gentle warmth

(26) Ik zie de Seine.
I see the Seine
'I see the Seine'

(27) * Ik zie Seine.
I see Seine

[5] A mass reading is possible, as in 'I see some sunlight.'

(28) * Seine-s flikkerende spiegeling
Seine-s flickering reflection

The -s possessive construction and the referential compound show resemblances. Their non-heads are referential and they are both marked by an *-s*. Yet, they also show significant empirical differences, showing that we are dealing with two distinct structures.

Consider the following four criteria that distinguish between the two structures. Firstly, consider the restriction that only non-heads that can undergo N-to-D movement when occurring in argument position can occur as the possessor in *-s* possessive constructions. A parallel restriction does not hold for referential compounds. The referential compound allows all roots as a non-head that have a unique - albeit universal or contextual- reference. The illicit *-s* possessive construction in (25) thus contrasts with the licit referential compound in (29):

(29) zon-s-verduistering
sun-s-eclipse
'solar eclipse'

I postpone an account for this contrast till section 6. For now it is important to note that the contrast exists as a criterion to distinguish between the two structures.

Secondly, the compounds qualify for word-hood in the sense that they can be lexicalised: they are stored in the native's speaker memory and in Dutch dictionaries. In that sense, a native speaker can distinguish between stored compounds, perhaps even with an idiomatic meaning, as shown in (30), and newly formed compounds, i.e. neologisms, which are to be interpreted literally, as illustrated in the examples in (31) and (32):

(30) Ze is een zondag-s-kind.
she is a Sunday-s-child
'She is born on a Sunday and thus for good luck.'

(31) # Ze is een maandagskind.
she is a Monday-s-child
(The speaker expresses that there is a salient connection in the discourse between the child and Mondays.)

(32) Ze is een zondag-s-oma.
she is a Sunday-s-grandmother
(The speaker expresses that there is a salient connection in the discourse between the grandmother and Sundays, for example, because this grandmother only visits the family on Sundays.)

I do not fully exclude a creative, humorous use of the neologism in (31), analogue to the idiomatic reading in (30) (i.e. a child that systematically fails to grasp good luck). However, the mere fact that it would be considered humorous illustrates the point that it is not lexicalised. The *-s* possessive construction does not qualify for word-hood in that sense: it is simply a freely generated syntactic constituent. Syntactic constituents are never experienced as 'neologisms':

(33) moeders auto/fiets/jurk/laptop/…
mother's car/bicycle/dress/laptop/…

Thirdly, the compounds qualify for word-hood phonologically: they receive compound stress, i.e. main stress falls on the non-head. The *-s* possessive construction, in contrast, receives the stress of a syntactic constituent:

(34) a. 'moeder-s-kind
mother-S-child
'child too dependent on the mother'

b. 'Kreeft-s-keerkring
Cancer-S-tropic
'Cancer Tropic'

(35) a. moeder-s 'auto
mother-S car
'mother's car'

b. Marie-s 'feestje
Mary-S party
'Mary's party'

Fourthly, the compound qualifies for word-hood in the sense that it cannot be interrupted by other words: an intervening adjective is excluded. The *-s* possessive construction, in contrast, allows for intervening adjectives:

(36) moeders mooie autootje
mother's pretty car.DIMINUTIVE
'mother's pretty little car'

(37) Ze was mijn grootmoeder van vader-s-zijde.
she was my grandmother of father-S-side
'She was the grandmother of my father's side of the family.'

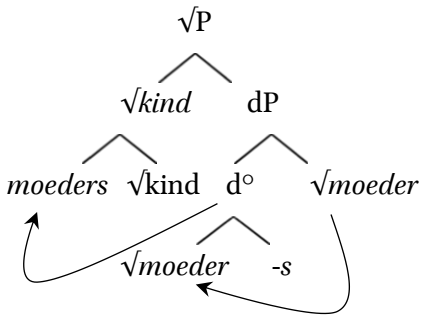
(38) * Ze was mijn grootmoeder van vader-s arme zijde.
she was my grandmother of father-S poor side

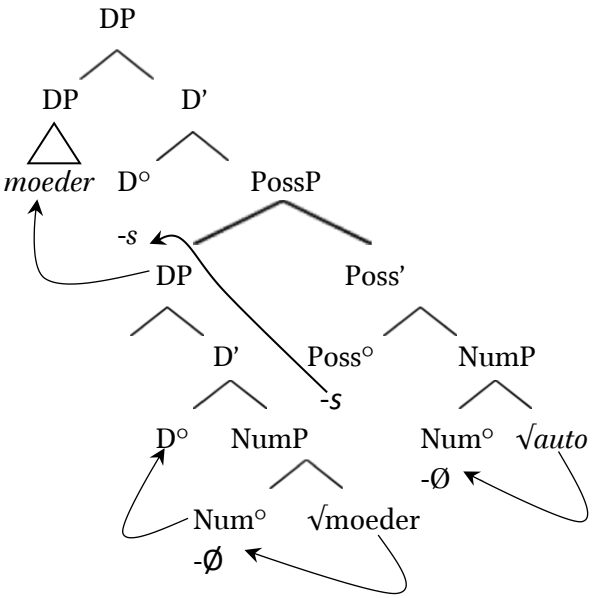
I conclude that the referential compound differs from the *-s* possessive constructon. The referential compound qualifies for word-hood and is thus truly a compound. Syntactically, this implies that it is derived through head incorporation, which counts as the syntactic movement that defines word-hood for compounds (Mithun 1984, Harley 2009). I propose that the non-head incorporates into a functional head called little d°, which is characterised by nominality and uniqueness: [n, unique]. It surfaces as *-s-*. The compound's non-head incorporates into this little d° head. The non-head plus *-s-* subsequently incorporates into the head of the compound.

The *-s* possessive construction, in contrast would not undergo head incorporation. The non-head rather moves to Spec,DP and the *-s* occupies the D° position. It moved there from its base position, which could either be analysed as Spec,nP or Spec,PossP. (See Abney 1987 on the Spec,DP position for the possessor and see Radford 2000, Alexiadou, Haegeman and Stavrou 2007:568 for the base position in Spec,nP for possessors. Corver 1990 proposes the *-s* ends in the D° head.)

(39) a. moeder-s-kind
mother-s-child
'child too dependent on the mother'

b. moeder-s auto
mother-s car
'mother's car'





5. The referential construct state

In this section I argue that referential construct states also occur in Dutch: they are title expressions (see also author 2009).

Title expressions are illustrated in (40)-(41). They always consist of one or more bare common nouns followed by a proper name.

(40) graaf Dracula
count Dracula
'count Dracula'

(41) professor doctor Einstein
professor doctor Einstein
'professor doctor Einstein'

The order cannot be changed, in the sense that the bare common noun needs to precede the proper name:

(42) * Dracula graaf
Dracula count

The construction needs to contain at least one proper name:

(43) * koningin paleis
queen palace

The whole functions semantically as a proper name as can be deduced from the following two tests. Firstly, Dutch definite NPs allow for a generic reading, as shown in (44), except when they are proper names, as in (45). As can be seen in (46), title expressions pattern with proper names in this respect.

(44) **De hond** is een trouw dier.
the dog is a loyal animal
✓ generic reading: 'Every dog is a loyal animal.'

(45) **Albert** is een flamboyante man.
Albert is a flamboyant man
* generic reading: 'Every Albert is a flamboyant man.'

(46) **Prins Albert** is een flamboyante man.
Prince Albert is a flamboyant man
* generic reading: 'Every prince Albert is a flamboyant man.'

Secondly, Dutch definite NPs allow for a *de dicto* and a *de re* reading, as in (47), proper names only allow for a *de re* reading, as in (48). Again, the title expression patterns with the proper names as it only allows for a *de re* reading, as illustrated in (49).

(47) Marie wil met **haar buurman** trouwen.
Marie wants with her neighbour marry
✓ 'Marie has the wish to marry the man living next to her, whoever that may be.' (de dicto)
✓ 'Marie wants to marry a specific man and this man is her neighbour.' (de re)

(48) Marie wil met **William** trouwen.
Marie wants with William marry
* 'Marie has the wish to marry no matter which man, as long as he is called William.' (de dicto)
✓ 'Marie wants to marry a specific man and this man is called William.' (de re)

(49) Marie wil met **prins William** trouwen.
Marie wants with prince William marry
* 'Marie has the wish to marry any man, as long as this man is called prince William.' (de dicto)
✓ 'Marie wants to marry a specific man and this man is prince William.' (de re)

I conclude that title expression are DPs which consist of one or more bare common nouns followed by a proper name and the whole functions semantically as a proper name.

One could think that the bare common noun needs to belong to a specific semantics field as it commonly refers to nobility, clergy, military ranks or professions, as shown in the following examples:

(50) prins Charles
prince Charles
'prince Charles'

(51) priester Damiaan
priest Damian
'father Damian'

(52) kapitein Von Trapp
captain Von Trapp
'captain Von Trapp'

(53) professor Curie
professor Curie
'professor Curie'

This could suggest that nouns referring to nobility, clergy, military ranks or professions form a closed class in the lexicon, which, for example, share a feature [+ title] or [+unique reference] or lexical selection properties. However, this cannot be the case, given that the title expression can be applied productively. Indeed, any common noun is licit as the bare common noun in the title expression:

(54) boekenkast Billy
book case Billy
'book case Billy'

Even nonce formations can occur as the bare common noun in the title expression:

(55) naakstrandgemeente Bredene
nude.beach.town Bredene
'Bredene, the town that has a nude beach'

Note that the 'proper name' does not need to be stored as a proper name either, the title expression is interpretable, regardless which roots realise the syntactic positions (cf. Borer 2005a):

(56) bedsofa Vimle
bed.couch Vimle
'bed couch Vimle'

This shows that the title expression does not depend on a specific property of the lexical items involved. Structural properties license them. This raises the questions which syntactic structure is

realised by the title expression and why the whole functions as a proper name. I propose that the bare common noun is licensed by N°-to-D° movement in a construct state (see also Author 2009). The title interpretation is a result of this syntactic structure.

In Hebrew, a construct state is a DP which consists of a bare, unstressed head noun which is immediately followed by a genitival phrase that is not overtly case marked. A variety of semantic relations can hold between them. (Cf. Borer 1984, Ritter 1991, Siloni 1997)

(57) beyt ha-'is [Hebrew examples are taken from Siloni 1997:21-26]
house the-man
'the man's house'

Title expressions and the Hebrew construct state share structural similarities. Firstly, in both constructions prepositions cannot intervene between the head and the complement:

(58) beyt (*sel) ha-'is [Hebrew]
house of the-man

(59) professor (*van) Einstein [Dutch]
professor of Einstein

Secondly, in both constructions an initial determiner is illicit. This is immediately clear for Hebrew:

(60) (*ha)-beyt ha-'is [Hebrew]
the-house the-man

For the Dutch title expressions, we first need a context in which a construction with a proper name would tolerate a determiner in the first place. Such examples exist: definite determiners may merge with proper names referring to males in Belgian Dutch.

(61) Ik heb de Larousse gezien. [Belgian Dutch]
I have the Larousse seen
'I have seen Larousse.' (*Larousse* is a family name.)

(62) Ik heb de Jan gezien. [Belgian Dutch]
I have the John seen
'I have seen John.'

Yet, in the title expression such an initial determiner is excluded, on a par with the Hebrew title expression in (63).

(63)* Ik heb de professor Larousse gezien. [Belgian Dutch]
I have the professor Larousse seen

Thirdly, in both constructions the whole DP inherits the referential properties of the second part. This is shown in examples (64)-(65) for Hebrew: the (in)definiteness of the second part of the construct state determines the (in)definiteness of the whole construct state:

(64) ben ha-melex [Hebrew]
son DEFINITE-king (example taken from Borer 1984: 45)
'the prince'

(65) ben melex [Hebrew]
son king (example taken from Borer 1984: 45)
'a prince'

For Dutch, we have seen that the entire title expression inherits the referential properties of the second part: the whole expressions functions semantically as a proper name, as they resist a generic reading and a *de dicto* reading.

Fourthly, both structures are recursive:

(66) gag beyt ha-'is [Hebrew]
roof house the-man
'the roof of the house of the man'

(67) *ingenieur doctor doctor professor ererector* [Dutch]
engineer doctor doctor professor honorary.president
associatievoorzitter baron Oosterlinck
association.president baron Oosterlinck
'engineer doctor doctor professor honorary president association president baron Oosterlinck'

Fifthly, for both constructions the presence of the complement is obligatory, as can be seen in the following examples:

(68) beyt ha-'is [Hebrew]
housethe-man
'the man's house'

(69) *beyt [Hebrew]
house

(70) Ik heb paus Benedictus gezien. [Dutch]
I have pope Benedict seen
'I have seen pope Benedict.'

(71) * Ik heb paus gezien. [Dutch]
I have pope seen

Sixthly, in both constructions, the head noun is de-stressed.

(72) * BAYIT ha-'is [Hebrew]
house the-man
(*bayit* is the stressed form of the unstressed *beyt*)

(73) * koning'in Fabiola [Dutch]
queen Fabiola

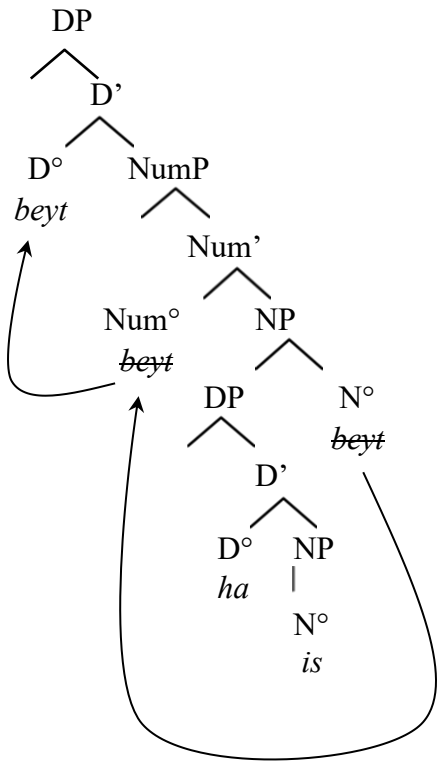
(74) koningin Ma'thilde [Dutch]
queen Mathilde
'queen Mathilde'

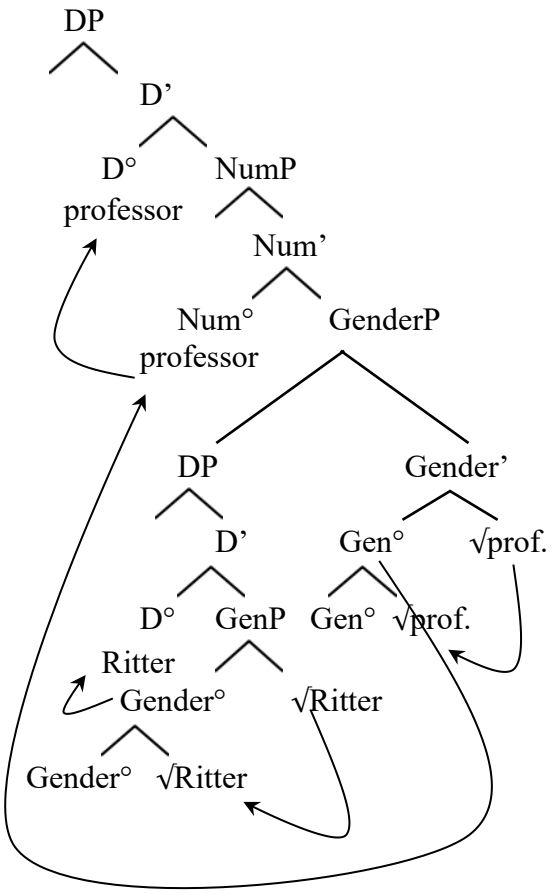
Note that this stress pattern clearly sets apart the title expression from a compound, as compounding stress would entail stress on the left-hand part in Dutch.

Given the fact that the second part of the construction functions as a proper name, I assume it undergoes N-to-D raising (Longobardi 1994) and as such acquires the referential properties of the proper name. Given the empirical similarities of the title expression to the construct state in Ritter (1991), I analyse it further on a par with Ritter's (1991) analysis for the Hebrew construct state in 2008:

(75) beyt ha-'is [Hebrew, Ritter 1991]
house the-man
'the man's house'

(76) professor Ritter [Dutch]
professor Ritter
'professor Ritter'





As can be concluded from the structures, the Dutch title expression is analysed as a construct state, but one in which the second part undergoes N-to-D movement, blocking the insertion of a determiner. In other words, it is a referential construct state.

In the construct state, the head noun moves through the nominal inflectional domain and does allows for the possibility of nominal inflection of the head noun in the form of number marking and diminutives:

(77) professor**en** Chomsky en Kayne [Dutch]
professors Chomsky and Kayne
'professors Chomsky and Kayne'

(78) prinses-je Elizabeth [Dutch]
princess-DIMINUTIVE Elizabeth
'little princess Elizabeth'

(79) prinses-je-s Elizabeth en Amalia [Dutch]
princess-DIMINUTIVE-PLURAL Elizabeth and Amalia
'the little princesses Elizabeth and Amalia'

Note that the agreement in number between the title and the proper name in the construction clearly distinguishes these constructions from compounds. In sum, Dutch has referential construct states alongside referential compounds.

6. When is N-to-D movement licit?

In section 4 I noticed a contrast between the referential compound and the *-s* possessive. It involved the issue that some roots which qualify for unique reference do not undergo N-to-D raising when in argument position. This goes hand in hand with the fact that they cannot occur as the non-head of the s-possessive either:

(80) # Ik zie zon.
I see sun

(81) * zons zachte warmte
sun's gentle warmth

Yet, they can be the non-head of the referential compound:

(82) zon-s-verduistering
sun-S-eclipse
'solar eclipse'

Why are some roots unable to undergo N-to-D raising and what causes the opposition between the referential compound and the *-s* possessive? In Borer (2005:84-85) the issue is discussed that certain proper names, such as *the Bronx* or *the Pacific Ocean* cannot occur without an article. After some discussion, Borer concludes that 'for reasons we can only speculate on' certain roots are banned from being proper names. She assumes that the reason is to be situated outside of syntax proper.

The present data indeed point in this direction. First, let us assume that for some roots the combination of the definite article plus the root is actually stored at Encyclopedia as the proper way to refer to the entity. For example, *de zon* is the conventional, lexicalised Dutch way to refer to the sun, whereas Z*on* is not. If *zon* 'sun' then would move to D, Encyclopedia would not be able to assign a reference to the construction. More generally, it follows that there is a rather superficial, extra-syntactic ban on the N-to-D movement. Syntax itself does not prohibit the movement. Interestingly, this makes the empirical predictions that if one can alleviate the extra-syntactic limitation, the N-to-D movement should be licit again.

Now consider the syntax of title expressions. They do not undergo N-to-D raising in argument position:

(83) * Ik feliciteer professor.
I congratulate professor

(84) Ik feliciteer de professor.
I congratulate the professor
'I congratulate the professor.'

The ban on the N-to-D movement for such roots is rather clear: they lack the typical unique reference which Encyclopedia requires to interpret the structure and there is no idiom stored to interpret the structure either. Note also the default encyclopedic interpretation for N-to-D movement would fail as well: *professor* should not be interpreted as a proper name, the sentence is not about a person whose proper name is Professor.[6] In sum, the structure is uninterpretable.

However, in title expressions the title itself is subject to N-to-D movement, allowing the entire title expression to pattern with proper names both syntactically and semantically (see section 5). Crucially, when combined with a proper name, there is no ban on moving to D for a title. After all, why should there be such a ban? A title expression has the same semantics as a proper name and it is thus fully interpretable: it has unique reference. These observations illustrate that syntax has no general ban for certain roots to move to D, the ban is interpretational.

Consider further the fact that many of the referential compounds under discussion are stored, lexicalised words. In other words, encyclopedia has an interpretation stored that matches their structure:

(85) zonsverduistering ↔ "SOLAR ECLIPSE"

The fact that the root *zon* 'sun' would require the determiner in argument position to refer to the star so familiar to us is simply irrelevant for the interpretation of the compound at Encyclopedia. Note that even if the compound had not been stored, the reasoning still holds. The non-head of a compound is a quite unique syntactic position which is arguably subject to its own interpretational rule at Encyclopedia.

[6] Unless of course, the person is indeed called Professor, as in *Money heist*.

More generally, it had been noted before that there is no ban on incorporating 'proper names that require an article' in compounds, as shown in (86) (example taken from Borer 2005a:84, who cites Peter Ackema for suggesting it):

(86) Bronx-lover

I conclude that it follows that the set of roots that can occur in r-compounds is a superset of the set that can occur in the -s-possessive: roots with unique reference that would otherwise require a definite article in argument position may occur in the compounds, but not in the -s possessive. More generally, the idea that the ban on N-to-D movement for roots such as *sun* and *Bronx* is extra-syntactic seems to be on the right track.

7. Conclusion

In this article I argued that Dutch has referential compounds: compounds with a referential non-head, even though that claim seems to be a *contradictio in terminis*. They resemble the Dutch s-possessive in that their non-heads involve movement to a referential layer. However, unlike the possessive structures, the compounding structure contains head incorporation which results in word-hood. The referential compound is therefore an instance of the fourth logical possibility according to the two parameters defined by Borer (2011) for N-N combinations: compounding versus constructs and modificational structures versus referential ones. Together with the syntactic structure of titles plus proper names, the referential compound further contributes evidence to the idea that a ban on N-to-D movement for certain uniquely referring roots, such as *sun* and *Bronx* is extra-syntactic.

References

Author (2009) To be completed.

Abney, Steven Paul (1987) *The English noun phrase in its sentential aspect.* Doctoral dissertation, MIT.

Alexiadou, Artemis, Liliane Haegeman and Melita Stavrou (2007) Noun phrase in the generative perspective. Studies in Generative Grammar Volume 71. Berlin & New York: De Gruyter Mouton.

Borer, Hagit (2005) *In name only. Structuring Sense Volume I.* Oxford: OUP.

Borer, Hagit (2011) Compounds: the view from Hebrew. In: Lieber, R. & P. Štekauer (eds.) *The Oxford handbook of compounding.* Oxford: OUP.

Botha, Rudolf P. (1968) The function of the lexicon in transformational generative grammar. *Janua Linguarum. Series Maior, Band 38.* The Hague: Mouton.

Corver, Norbert (1990) *The syntax of left branch constructions.* Doct. dissertation, Tilburg University.

De Belder, Marijke (2009) *On the syntax of titles.* Talk presented at ConSOLE, Nova Gorica, January 16-18, 2009.

De Belder, Marijke (2017) The root and nothing but the root: primary compounds in Dutch. *Syntax* 20(2). 138-169.

De Belder, Marijke (to appear) The syntax-semantics interface in Distributed Morphology. To appear in: A. Alexiadou, R. Kramer, A. Marantz & I. Oltra-Massuet (eds). *The Cambridge Handbook of Distributed Morphology*. Cambridge: CUP. (accessible on Lingbuzz as lingbuzz/006470)

Harley, Heidi (2009) Compounding in DM. In Lieber, R. & P. Štekauer (eds.) *The Oxford handbook of compounding*, 129-144. Oxford: OUP.

Longobardi, Giuseppe (1994) Reference and proper names: a theory of N-movement in syntax and logical form. *Linguistic Inquiry* 25(4). 609-665.

Lyons, Christopher (1999) *Definiteness*. Cambridge: Cambridge University Press.

Mithun, Marianne (1984) The Evolution of Noun Incorporation. *Language* 60(4). 847-894.