The Dutch Anaphoric Possessive Construction

Frank Van Eynde University of Leuven

ABSTRACT

The Dutch anaphoric possessive construction (APC), as exemplified by *Tom zijn fiets* 'Tom his bike', shows a peculiar mix of regularity and idiosyncracy. The article provides a theory-neutral description of its properties and quantitative information about its use in two treebanks, one of spoken Dutch (CGN) and one of written Dutch (Lassy Small). It argues that the APC has a right branching structure and models it in the framework of Constructional Head-driven Phrase Structure Grammar. The latter's organization of constructions in terms of a fine-grained hierarchy of phrase types is shown to provide the means to capture both what the APC has in common with other possessive constructions and what is idiosyncratic of it.

Keywords: possessive constructions, anaphoric possessives, Dutch treebanks, Constructional Head-driven Phrase Structure Grammar, multiple inheritance

INTRODUCTION

Dutch has three semantically equivalent ways of expressing possession. Beside the PP[*van*] option and the genitive option, as exemplified in (1) and (2), there is the option of using a possessive determiner that is preceded by an NP, as in (3).

(1) Ik heb [de fiets van Tom] verkocht.I have the bike of Tom sold'I have sold Tom's bike.'

Journal of Language Modelling Vol 11, No 2 (2023), pp. 267-296

1

- (2) Ik heb [Toms fiets] verkocht.I have Tom.GEN bike sold'I have sold Tom's bike.'
- (3) Ik heb [Tom zijn fiets] verkocht.I have Tom his bike sold'I have sold Tom's bike.'

We use the term 'Anaphoric Possessive Construction' (APC) for the latter, since the determiner (*zijn* 'his') necessarily has the same referent as the preceding NP (*Tom*). The choice between the options is mainly determined by style and register. The genitive, for instance, is typical of written and slightly formal language, while the APC is typical of colloquial speech.

The APC displays a peculiar mixture of regularity and idiosyncracy. Its regularity is clear from the fact that the NP before the determiner and the nominal after it can take nearly any form. Its idiosyncracy is clear from the fact that it lacks a counterpart in other languages, including closely related ones. English, for instance, does not have it (**Tom his bike*), and the same holds for French (**Tom sa bicyclette*) and Italian (**Tom sua bici*).¹ This makes the APC an interesting test case for the larger issue of whether such constructions are amenable to formal analysis. We claim they are. To demonstrate it we adopt an approach that is inspired by the following quote: "To know what is idiomatic about a phrase one has to know what is nongeneral and to identify something as nongeneral one has to be able to identify the general ... The picture that emerges from the consideration of special constructions ... is of a grammar in which the particular and the general are knit together seamlessly" (Kay and Fillmore 1999, 30).

A framework that provides the tools for developing such a grammar is Constructional Head-driven Phrase Structure Grammar (HPSG), as pioneered in Sag 1997 and elaborated in more detail in Ginzburg and Sag 2000.² We will use it for a formal analysis of the APC in Section 4. To pave the way we first provide a theory-neutral description of

¹ A language which also uses the APC is German, see (31).

² Another framework that would serve the purpose well is Sign-Based Construction Grammar (Boas and Sag 2012). The analysis in this paper can be translated directly into SBCG terms.

	CGN	%	Lassy	%	Sum	%
Genitive	134	71.66	668	98.67	802	92.83
Anaphoric possessive	53	28.34	9	1.33	62	7.17
Sum	187	100.00	677	100.00	864	100.00

Table 1: Possessive NPs in the sample

2

the constuction in Section 2 and a discussion of its syntactic structure in Section 3.

For the purpose of exemplification we use a sample that consists of the treebank of the Spoken Dutch Corpus (CGN) (Oostdijk *et al.* 2002) and Lassy Small, a treebank of written Dutch (Van Noord *et al.* 2013). They contain approximately 1,000,000 words each and provide a syntactic analysis for every sentence that comprises both categorial and functional information. Every sentence is marked by an identifier.³ To extract relevant examples and quantitative data from the sample we use the XPath search mode of Gretel 3.0 (Augustinus *et al.* 2012). It facilitates amongst others the retrieval of all APC occurrences in the sample. As expected, it is far more common in spoken Dutch than in written Dutch: Of the 62 tokens, 53 (85.48%) are from CGN and only 9 (14.52%) from Lassy Small. Confirmation is provided by a comparison with the genitive. It is more common than the APC in both treebanks, but the difference is much larger in Lassy Small (98.67% *vs* 1.33%) than in CGN (71.66% *vs* 28.34%), see Table 1.

DESCRIPTION

The APC is described amongst others in Paardekooper 1984, 478–479, Haeseryn *et al.* 1997, 294–295, 822–823, and Broekhuis and Keizer 2012, 837–839.⁴ It consists of a possessive determiner that is preceded by an NP and followed by a bare nominal. The nominal can take any

³Identifiers with the prefix 'fn' stand for spoken data from the Netherlands, and identifiers with the prefix 'fv' stand for spoken data from the Dutch speaking part of Belgium. Identifiers with another prefix are taken from Lassy Small.

⁴In Broekhuis and Keizer 2012, 837, it is called the semi-genitival construction.

Frank Van Eynde

form but the possessive and the preceding NP are subject to a number of restrictions that will be presented in Section 2.1 and Section 2.2 respectively.

The possessive determiner

2.1

in

The possessive determiner is invariably of the third person. It can be singular masculine (*zijn* 'his'), singular feminine (*haar* 'her') or plural (*hun* 'their'). The former two often appear in the reduced form, i.e. z'n, 'r or d'r. Broekhuis and Keizer 2012, 837, claim that only the reduced forms can be used in the APC and that the plural *hun* 'their' is therefore not acceptable. This is not confirmed by the sample, where the full forms account for more than 60% of the tokens, including 6 for *hun*, see Table 2.⁵

Table 2: Anaphoric			Sing. masc.		Sing. fem.		Plural	
possessives	Full form	zijn	19	haar	13	hun	6	38
in the sample	Reduced form	z'n	20	'n, d'n	4			24
	Sum		39		17		6	62

Being anaphoric, the determiner shows number and gender agreement with the preceding NP. In (4), for instance, it is the singular masculine z'n 'his' that must be used, since the noun *Max* is singular and masculine.

 (4) dat komt gewoon op [Max z'n bankrekening] that comes simply on Max his bank.account
 'that simply goes to Max's bank account' [fnf007265_116]

That the agreement concerns natural gender is illustrated in (5).

⁵ In written language it is more common to use the full form. The 9 occurrences in Lassy Small, for instance, all concern the full form. In spoken language both forms are used. CGN, for instance, contains 29 occurrences of the full form and 24 of the reduced form.

(5) van [dat meisje haar ex-lief] of that girl her ex-boyfriend'of that girl's ex-boyfriend' [fva400508_64]

The noun *meisje* 'girl' is grammatically neuter, but denotes a female person, and it is the latter that is relevant for the choice of the determiner. If the preceding NP is underspecified for number or gender, it is compatible with more than one. The demonstrative pronoun *die* 'that.one', for instance, is combined with all three in the sample, as shown in (6)–(8).

- (6) [die z'n idee] was dat that.one his idea was that 'that was his idea' [fva400459_141]
- (7) [die haar broer] is nog gekomen that.one her brother is still come
 'her brother has still come' [fvd900058_140]
- (8) hebben ze al [die hun meubels] d'ruit gegooid have they all those.ones their furniture out thrown 'they threw out all their furniture' [fva400466_99]

This anaphoric nature differentiates the APC from the English 'spossessive, in which the clitic indiscriminately combines with singular masculine NPs (*John's bike*), singular feminine NPs (*that girl's boyfriend*) and plural NPs (*the children's toys*).

The possessor NP

2.2

The possessor NP must be animate. This was already pointed out in Paardekooper 1984, 479, and is repeated in Haeseryn *et al.* 1997, 294–295, and Broekhuis and Keizer 2012, 838–839. It is a constraint that differentiates the APC from the genitive, as illustrated in (9) and (10).

Frank Van Eynde

- (9) a. ['s werelds eerste multinational] werd in 1602 the.GEN world.GEN first multinational was in 1602 ... opgericht ... founded
 'the world's first multinational was founded in 1602'
 - b. * [de wereld zijn eerste multinational] werd in 1602 ... the world his first multinational was in 1602 ... opgericht founded
- (10) a. het gevaar van [Iraks wapenarsenaal] the danger of Iraq.GEN weapon.arsenal
 'the danger of Iraq's weapon arsenal' [ws-u-e-a-0000000027.p.7.s.2]

[dpc-bal-001238-nl-sen.p.28.s.3]

b. * het gevaar van [Irak z'n/d'r wapenarsenaal] the danger of Iraq his/her weapon.arsenal

The animacy constraint is confirmed by the sample. Of the 62 tokens, 40 concern a proper noun that denotes a person, such as *Max*, or an animal, such as *Reynaert* (a fox). Proper nouns that denote a country or some other inanimate entity are not attested. 14 concern a pronoun with a human referent, such as *iemand* 'somebody' (3 tokens), *die* 'that.one' (10 tokens) or *wie* 'who'.⁶ Their [-HUMAN] counterparts, *iets* 'something', *dat* 'that' and *wat* 'what', are not attested in the APC. The 8 remaining ones concern the common nouns *mensen* 'people' (3 tokens), *kind* 'child', *man* 'man', *tante* 'aunt', *meisje* 'girl' and *advocaat* 'lawyer', all of which have a human referent.

A second constraint concerns the exclusion of "referential and reciprocal personal pronouns: **hij/hem z'n boek* 'he/him his book', **zij/haar d'r boek* 'she/her her book' and **elkaar z'n/hun boek* 'each.other his/their book" (Broekhuis and Keizer 2012, 839). This is another way in which the APC differs from the genitive, at least for the reciprocal pronouns, as illustrated in (11).

⁶ For *die* it is worth adding that it may have a non-human referent in its other uses.

- (11) a. dat mensen die op elkaar lijken van that people who each.other resemble of [elkaars paspoort] gebruik maken each.other.GEN passport use make 'that people who resemble each other use each other's passport' [ws-u-e-a-0000000240.p.30.s.4]
 - b. * dat mensen die op elkaar lijken van that people who each.other resemble of [elkaar hun paspoort] gebruik maken each.other their passport use make

Not mentioned in Broekhuis and Keizer 2012, but equally unfit for use in the APC, are the reflexive pronouns, as in **zich zijn kat* 'himself his cat'.

At the same time, there are no constraints on the internal structure of the possessor NP. The noun that heads the NP may be preceded by a dependent, such as the demonstrative *dat* 'that' in (5). If the preceding word is a possessive determiner of the third person, it is possible to embed one APC in another, as in (12).

(12) Ze heeft [[[mijn vader] z'n tante] d'r boeken] geërfd.she has my father his aunt her books inherited'She has inherited my father's aunt's books.'

The noun may also be followed by a dependent, such as the PP[*van*] in (13) and (14).

- (13) [[wie van jullie] z'n boek] is dit eigenlijk? who of you.PL his book is this really?'Who of you's book is this?'
- (14) ze werken liever onder [[iemand van ons] z'n they work rather under someone of us his hoede] surveillance
 'They'd rather work under the surveillance of someone of us.'

In this respect, the APC is less constrained than the genitive.

- (15) * [wiens van jullie boek] is dit eigenlijk? who.GEN of you.PL book is this really?
- (16) * ze werken liever onder [iemands van ons they work rather under someone.GEN of us hoede] surveillance

Notice, finally, that the possessor NP can be coordinated, as in (17).

(17) ... aansluitend bij [[Rianne en Rika] hun verhaal]
... related to Rianne and Rika their story
'... related to Rianne and Rika's story'
[fvf600083_106]

Summing up

Prenominal APCs are $[NP_i + Poss-Det_i + Nominal]$ -sequences, in which the possessive determiner is of the third person and anaphoric, in the sense of showing number and gender agreement with the preceding NP. The latter must be animate and cannot take the form of a personal, reciprocal or reflexive pronoun, but its internal structure is free.

STRUCTURE

To model the internal structure of the APC, one possibility is to treat the possessive determiner and the preceding NP as a phrasal determiner (DetP), as in (18).

(18) Ik heb [[Tom z'n] fiets] verkocht I have Tom his bike sold

2.3

3

This structure is adopted in the Dutch treebanks, and is suggested in Broekhuis and Keizer 2012, 837, which emphasizes the resemblance between the possessive determiner in *Tom z'n fiets* and the genitive affix in *Toms fiets*. In fact, this structure was already proposed in Paardekooper 1984, 478–479, albeit with the addition, put between

The Dutch Anaphoric Possessive Construction

	we	me	you	you	she/they	her
Full form	wij	mij	jij	jou	zij	haar
Reduced form	we	те	je	je	ze	d'r, 'r

Table 3: Personal pronouns with a reduced counterpart

parentheses, that z'n and d'r are the only unaccented words that can be the head of a phrase.⁷ This observation is worth a closer look.

For a start, notice that many of the Dutch personal pronouns have both a full form and a reduced form. Some examples are given in Table 3. Typical of the full forms is that they have a full vowel or a diphtong as their nucleus, while the reduced forms have the schwa. This phonological distinction corresponds to differences in syntactic potential. One concerns the fact that full forms can take dependents, while their reduced counterparts cannot (Van Eynde 1999). Notice, for instance, that both can be used as the subject in (19), but that only the full form can be used if the pronoun is modified by the adverb *alleen* 'alone', as in (20).

- (19) Wij/we hebben dat bericht nog niet gekregen.we have that message still not received'We have not yet received that message.'
- (20) [Alleen wij/*we] hebben dat bericht nog niet gekregen.
 alone we have that message still not received
 'We alone have not yet received that message.'

Similarly, while both forms can be used as the subject in (21), only the full form can be used if we add a relative clause, as in (22).

- (21) Zij/ze krijgen een bonus.they receive a bonus'They receive a bonus.'
- (22) [Zij/*ze die het verdienen] krijgen een bonus. they who it deserve receive a bonus'Those who deserve it receive a bonus.'

 $^{^{7}}$ This is a translation of the Dutch original: "*Z'n* en *d'r* zijn de enigste onbeklemtoonde woorden die kern van een patroon kunnen zijn" (Paardekooper 1984, 479).

Frank Van Eynde

A related difference concerns the fact that full forms can be conjoined, while their reduced counterparts cannot.

- (23) [Wij en zij] denken daar hetzelfde over. we and they think there the same about 'We and they think the same about that.'
- (24) * [We and ze] denken daar hetzelfde over. we and they think there the same about

Similar observations apply to the possessives. While both forms can be used in the determiner position of *bericht* 'message' in (25), only the full form can be used if the possessive is modified by *alleen*, as in (26).⁸

- (25) Zijn/z'n bericht is al verstuurd.his message is already sent'His message has already been sent.'
- (26) [Alleen zijn/*z'n] bericht is verstuurd. Het onze niet.
 only his message is sent. the our.NOM not
 'Only his message was sent. Ours was not.'

This is confirmed by the coordination test:

- (27) Als je [zijn en haar] getuigenis vergelijkt, dan ...if you his and her testimony compare, then ...'If you compare his and her testimony, then ...'
- (28) * Als je [z'n en d'r] getuigenis vergelijkt, dan ...
 * if you his and her testimony compare, then ...

(26) and (28) pose a problem for the DetP-analysis of the APC, since they make the reduced possessives doubly exceptional. Beside the stipulation that they are "the only unaccented words that can be the head of a phrase", we also need the stipulation that this exceptional behavior is limited to their use in the APC, since their incompatibility with

⁸ The reduced form can be used in (26) if *alleen* modifies the entire NP, as in [*alleen* [*zijn/z'n bericht*]], but in that case *alleen* 'only' is not a dependent of the possessive. Similarly, in *al z'n berichten* 'all his messages', the quantifying *al* is not a dependent of the possessive, but of the NP *z'n berichten*. This is clear from the fact that it is the messages that are quantified over, rather than him.

adverbial modifiers and their non-conjoinability is as expected. Neither of these stipulations is needed if the possessive is taken to form a unit with the bare nominal that follows it, as in (29).

(29) ik heb [Tom [z'n fiets]] verkocht I have Tom his bike sold

In this structure, the possessive is part of an NP in which it takes its usual specifier position. Confirming evidence is provided by the fact that this NP can be conjoined, as in *Tom z'n schoenen en z'n laarzen* 'Tom his shoes and his boots'.⁹

Within the rightmost NP in (29), the head is not the possessive determiner but the common noun that follows it. Evidence is provided by the fact that it shares the number and gender values of the NP. The NP *z*'n *zussen* 'his sisters', for instance, is not singular and masculine, like *z*'n 'his', but plural and feminine, like *zussen* 'sisters'.

For the APC as a whole, we assume that the rightmost NP is the head, since the APC shares its number and gender with that NP. Notice, for instance, that the finite verb in (30) shows number agreement with the plural *z'n fietsen* 'his bikes', rather than with the singular *Tom*.

 (30) [Tom [z'n fietsen]] zijn/*is gestolen. Tom his bike.PL are/*is stolen
 'Tom's bikes are stolen.'

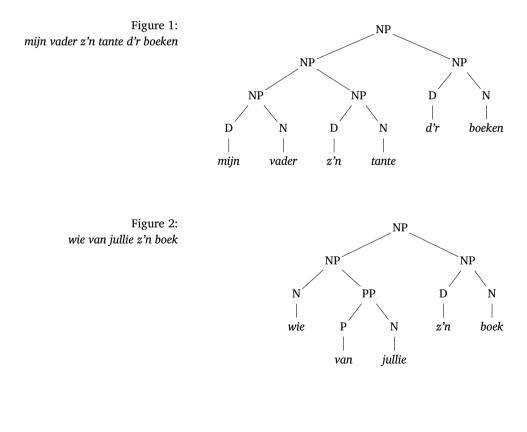
Confirming evidence is provided by the the German equivalent of the APC, exemplified in (31).

(31) Kennst du [dem Hans [seine Mutter]]? know you the.DAT Hans his.ACC mother'Do you know Hans' mother?'

⁹ An anonymous reviewer points out that it is also possible to conjoin the combination of the possessive and the preceding NP, as in *Tom z'n en Marie d'r kinderen* 'Tom his and Mary her children'. This combination sounds awkward to the native speakers I consulted, but for those who consider it well-formed, it can be described as an instance of Right Node Raising, comparable to *vier grote en twee kleine kinderen* 'four tall and two small children'.

As the glosses show, the APC in (31) contains a dative NP followed by an accusative NP. The former's case is fixed: The possessor NP is invariably dative in the APC. The latter's case, by contrast, is determined by the context in which the NP as a whole appears. In this case, this is accusative, since the NP is the direct object of the verb *kennst* 'know'. The fact that it is the rightmost NP that bears accusative case, hence, indicates that it is the head. Beside the syntactic arguments for treating the rightmost NP as the head of the APC, there is the obvious semantic argument that the APC denotes something of the kind that is denoted by the rightmost NP. *Tom z'n fiets*, for instance, denotes a bike, rather than a person named *Tom*.

Pulling the strings together, we assume that the APC consists of two NPs and that the rightmost one is the head. Figures 1, 2 and 3 show how this applies to APCs in which the possessor NP has a more complex internal structure, as in (12), (13) and (17).



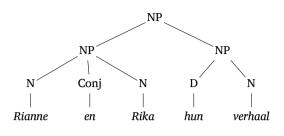


Figure 3: *Rianne en Rika hun verhaal*

4

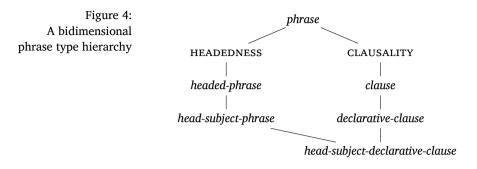
ANALYSIS

Having described the main properties of the APC in theory-neutral terms we now turn to a formal analysis. The aim is to show, first, that also constructions with a high degree of idiosyncracy, such as the APC, are amenable to formal analysis, and second, that this requires a framework that allows the attribution of properties to phrase-size combinations; in other words, that it requires a constructional, rather than a purely lexicalist, approach.

The framework we employ is that of Constructional HPSG, as pioneered in Sag 1997 and developed more fully in Ginzburg and Sag 2000. A key property of that framework is the classification of phrases in terms of a bidimensional phrase type hierarchy.

In the first dimension, called HEADEDNESS, phrases are classified in terms of syntactic dependency. The basic distinction is that between headed and non-headed phrases. *Kim smiled*, for instance, is a headed phrase in which the verb is the head and the noun its subject. By contrast, *Kim and Mary* is a non-headed phrase, consisting of two conjuncts and the conjunction *and*. Characteristic of the HEADED-NESS classification is its cross-categorial nature. Heads, subjects and conjuncts, for instance, can belong to any syntactic category. In early HPSG this was the only dimension of classification for phrases (Pollard and Sag 1994). Modeling their properties was reduced to the interaction of a small number of highly abstract cross-categorial phrase types with a very large number of detailed category specific lexical types. When applied to phrases with idiosyncratic properties, this radically lexicalist stance turned out to have its limitations.

Taking a cue from the organization of the lexicon in terms of a bidimensional hierarchy of lexical types in Pollard and Sag 1987,



191–218, the phrase type hierarchy was enriched with a second dimension, called CLAUSALITY. The basic distinction in that dimension is between clauses and non-clauses. In contrast to the distinctions in the HEADEDNESS dimension, these are not cross-categorial. Instead, they capture generalizations about specific syntactic categories and/or semantic types, differentiating, for instance, between declarative, interrogative, imperative and exclamative clauses (Ginzburg and Sag 2000, 38–42).

Given that the HEADEDNESS and CLAUSALITY dimensions are orthogonal, in the sense that they make mutually independent distinctions, it is possible to define phrase types that combine properties from a type in the HEADEDNESS dimension, on the one hand, and properties from a type in the CLAUSALITY dimension, on the other hand. *Kim smiled*, for instance, is subsumed by a type, called *head-subjectdeclarative-clause*, that is a subtype of *head-subject-phrase*, on the one hand, and of *declarative-clause*, on the other hand, see Figure 4. The purpose of this hierarchy is to provide the means to capture generalizations at various levels of specificity, ranging from the very general, such as the properties that all headed phrases share, to the very specific, such as the idiosyncratic properties of an inverted exclamative clause like *Am I tired*!

While Ginzburg and Sag (2000) mainly focus on clauses, Van Eynde (2018) shows how the approach can be extended to noun phrases, yielding a framework in terms of which it is possible to model both ordinary nominals, such as *red boxes*, and idiosyncratic ones, such as the Big Mess Construction (*so big a man*) and the Binominal Noun Phrase Construction (*her nitwit of a husband*). It is this framework that we will adopt for an analysis of the APC. More specifically, we assume that there is a type, called *anaphoric-possessive*, that inherits the constraints of one of the subtypes of *headed-phrase*, on the one hand, and one of the subtypes of *non-clause*, on the other hand.

The section subsequently focusses on the HEADEDNESS dimension (Section 4.1), the CLAUSALITY dimension (Section 4.2) and their combination (Section 4.3). The resulting hierarchy is then used to model the anaphoric possessive construction (Section 4.4). In a final step, we compare the resulting analysis with a lexicalist alternative (Section 4.5). Throughout, we use the Typed Feature Structure notation that has been employed in HPSG since Pollard and Sag 1987. A recent comprehensive survey is provided in Müller *et al.* 2021.

The HEADEDNESS dimension

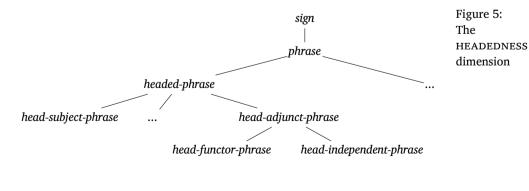
4.1

Building on the conclusion in Section 3 that the APC is an [NP + NP]-sequence in which the rightmost NP is the head, we start with a look at the hierarchy of headed phrases in Figure 5.

The properties of the various types in the hierarchy are represented by features. Which features are relevant for which types is spelled out in terms of feature declarations, as in (32).

(32) sign : [PHON list(phone) SYNSEM synsem
]
phrase : [DTRS list(sign)]
headed-phr : [HEAD-DTR sign]

Every sign, whether lexical or phrasal, has a PHON(OLOGY) feature whose value is a list of phonemes, and a SYN(TAX-)SEM(ANTICS)

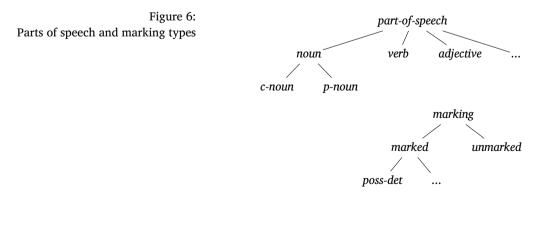


feature whose value is an object of type *synsem*, which stands for the various syntactic and semantic properties of a sign. In contrast to lexical signs, phrasal signs also have a D(AUGH)T(E)RS feature whose value is a list of signs. Moreover, headed phrases have a HEAD-D(AUGH)T(E)R feature whose value is a sign. The phrase *red box*, for instance, has two daughters, of which the head daughter is *box*. Non-headed phrases, such as *Kim and Mary*, lack this feature.

For our purpose, it is mainly the values of the SYNSEM feature that matter. They comprise among others the CATEGORY feature, whose value captures most of the syntactic properties of signs. Technically, this value is of type *category* and is declared to have the features in (33).

(33)	category :	HEAD	part-of-speech	
		MARKING	marking	
		SUBJ	list(synsem)	
		COMPS	list(synsem)	

The HEAD value is a part of speech, such as *verb* or *noun*. A partial inventory is given in Figure 6. Each of these may be declared to have other features. Verbs, for instance, have a VFORM feature, differentiating amongst others between finite and nonfinite forms, while nouns have a CASE feature. The finer-grained distinction between common nouns, on the one hand, and proper nouns and pronouns, on the other hand, is motivated by the fact that the former have morpho-syntactic NUMBER and GENDER features, while the latter lack these.



```
(34) verb : [VFORM vform]
noun : [CASE case]
c-noun : [NUMBER number]
GENDER gender
```

The main reason for differentiating the HEAD feature from other syntactic features is that its value captures precisely those properties which a phrase shares with its head daughter. The head daughter of a finite VP, for instance, is a finite verb, and the head daughter of an accusative NP is an accusative noun. Technically, this sharing is captured in terms of the implicational constraint in (35), quoted from Sag 1997, 439.¹⁰

(35) Head Feature Principle (HFP)

 $headed-phrase \Rightarrow \begin{bmatrix} synsem | ... | HEAD & 1 part-of-speech \\ HEAD-DTR | synsem | ... | HEAD & 1 \end{bmatrix}$

The HFP is the HPSG equivalent of the central principle of x-bar syntax that phrases are projections of lexical categories, albeit with the important qualification that the HFP applies to surface structures, not to the abstract underlying structures of Transformational Grammar and its descendants.

The other syntactic features capture properties that are not shared between a phrase and its head daughter. In X-bar syntax this includes the bar level. In HPSG it concerns the properties that are captured by the valence features (SUBJ and COMPS) and the MARKING feature. For modeling the APC it is mainly the latter that matters. One of its functions is to register the degree of saturation of a projection, more specifically its degree of functional saturation. For nominal projections it differentiates those which are partially saturated, such as *box* and *red box*, from those which are fully saturated, such as *this box* and *she*. The relevance of this distinction is clear amongst others from the fact that the former can be combined with other prenominal dependents, as in *small red box* and *that red box*, while the latter cannot: **small this box* and **the she*. Formally, the partially saturated signs have the value *unmarked* and the functionally saturated ones the value *marked*.

¹⁰ The recurrence of the boxed integer stands for token-identity, i.e. sharing. The feature paths are abbreviated, as indicated by the dots.

They both have a range of more specific subtypes, but the inventory in Figure 6 only contains the one we need for the APC, i.e. *poss-det*. It is assigned to the possessive determiners and to the NPs which contain them. This sharing between the determiner and the NP is not modeled by the HFP, since it is not the determiner, but rather the nominal, that is treated as the head of an NP in HPSG (Van Eynde 2021). Instead, the determiner is treated as adjoined to a nominal projection and the sharing is modeled in terms of the Marking Principle. Formally this is an implicational constraint on phrases of type *head-adjunct-phrase*. It is spelled out in (36).

(36) Marking Principle

head-adjunct-phrase \Rightarrow

 SYNSEM | ... | MARKING \square marking

 DTRS $\langle [SYNSEM | ... | MARKING \square], 2 \rangle$

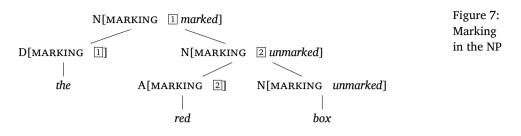
 HEAD-DTR 2 sign

What this says is that signs of type *head-adjunct-phrase* share the MARK-ING value of their non-head daughter (1).

An example is provided in Figure 7. The adjective *red* has a MARK-ING value of type *unmarked*, and shares this with the phrase *red box* (2), while the article *the* has the value *marked* and shares this with the NP (\square).¹¹

While the MARKING value registers the degree of saturation, something more is needed to prevent the formation of ill-formed combinations, such as **small this box* and **the she*. For that purpose we use the SELECT feature. Its value specifies the properties which a sign imposes on its head sister. Prenominal adjectives and determiners, for

¹¹ An anonymous reviewer points out that in *red box* the sharing might as well be with the noun, since it has the same MARKING value as the adjective. This is indeed true for the example in Figure 7, but it is not true in cases where finergrained distinctions are needed. One of them concerns the contrast between the Dutch nominals *een zwart paard* 'a black horse' and *het zwarte paard* 'the black.DCL horse'. In both cases the noun is unmarked, but while the addition of an adjective in the base form yields a bare nominal, the addition of a declined adjective yields a nominal that must be preceded by a definite determiner. This can be modeled if one assigns different subtypes of *unmarked* to the adjectives, depending on whether they are in the base form or declined, and if this more specific value is shared with the nominal, see Van Eynde 2006, 170–178.



instance, select an unmarked nominal and are, hence, incompatible with a nominal that contains a determiner. Technically, SELECT is assigned to objects of type *part-of-speech* and its value is either a bundle of syntactic and semantic properties or *none*.¹²

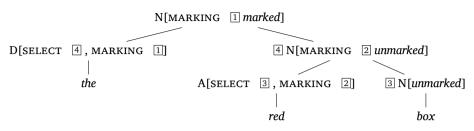
(37) part-of-speech : $\begin{bmatrix} \text{SELECT} & \text{synsem} \lor \text{none} \end{bmatrix}$

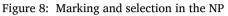
Since *part-of-speech* is the value of the HEAD feature, it is subsumed by the HFP, so that the SELECT value is shared between a phrase and its head daughter. For instance, if the adjective *large* selects an unmarked nominal, then so does the AP *very large*. To model the sharing that the selection involves, we employ a constraint on signs of type *head-functor-phrase*, which is a subtype of *head-adjunct-phrase*. It is spelled out in (38).

(38) Head-Functor Phrase

The SELECT value of the non-head daughter is required to match the SYNSEM value of the head daughter (1). It interacts with the MARK-ING value in a way that is effective to prevent overgeneration, as illustrated in Figure 8. Adjectives and determiners both select an unmarked nominal sister, but since the former has a MARKING value of type *unmarked* (or one of its subtypes), while the latter has a MARKING

¹²Earlier versions of HPSG made a distinction between selection by members of the substantive parts of speech (noun, verb, adjective, preposition) and selection by members of functional parts of speech (determiner, complementizer). The former was modeled by MOD(IFIED), the latter by SPEC(IFIED) (Pollard and Sag 1994). This distinction is neutralized in the functor treatment.





value of type *marked*, it is possible to stack adjectives but not determiners. Exploiting the potential of this combination of marking and selection, Allegranza (2006) provides a detailed analysis of Italian NPs and Van Eynde (2006) of Dutch NPs.

Not all adjuncts select their head sister. Loose apposition, for instance, as exemplified by *Berlin, the current capital of Germany*, concerns a juxtaposition of two NPs in which neither NP selects the other (Van Eynde and Kim 2016). To model this we add another subtype of *head-adjunct-phrase*, called *head-independent-phrase*, in which the SE-LECT value of the non-head daughter is *none*.

(39) Head-Independent Phrase

4.2

 $\begin{array}{c} head-independent-phrase \implies \left[DTRS \left< \left[SYNSEM | \dots | HEAD | SELECT none \right], 1 \right> \right] \\ HEAD-DTR 1 \end{array} \right]$

This type of phrase is also used in Van Eynde 2018 for the analysis of a number of nominals with idiosyncratic properties, such as the English Big Mess Construction (*so big a mess*) and the Binominal Noun Phrase Construction (*her nitwit of a husband*). It will play a role in our analysis of the APC as well (Section 4.4).

The CLAUSALITY dimension

Orthogonal to the dimension of HEADEDNESS is the dimension of CLAUSALITY. The types that populate this dimension tend to include constraints on semantic types, see Ginzburg and Sag 2000. To model the semantic properties of signs HPSG employs the CONTENT feature. It is part of the objects of type *synsem*, along with the CATEGORY feature. Its value is an object of type *semantic-object* and these come in

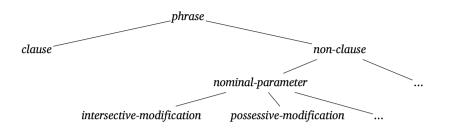


Figure 9: The CLAUSALITY dimension

a variety of subtypes, one of which is *scope-object*. This type is used to model the semantic properties of nouns, adjectives and determiners. Technically, it consists of an index and a set of facts that jointly restrict the denotation of the index.

(40) scope-object : [INDEX index RESTR set(fact)]

The CONTENT value of the noun *bike*, for instance, consists of an index i and the restriction that i is a bike. HPSG indices are comparable, but not identical, to Predicate Logic variables. One of the differences concerns the fact that indices are declared to have agreement features, as spelled out in (41).¹³

(41)	index :	PERSON	person	
		NUMBER	number	
		GENDER	gender	

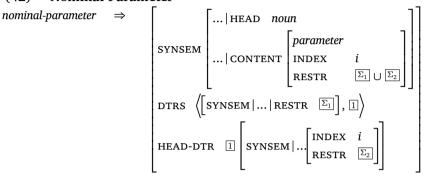
Co-indexed nominals share the values of these features, thus modeling amongst others the agreement between an anaphoric pronoun and its antecedent: *I wash myself/*yourself/*ourselves*. Scope-objects come in two subtypes, depending on whether or not they contain a quantifier, such as *every* or *no*. Those which are not explicitly quantified are called *parameter*.

Making use of the semantic types and their associated features Van Eynde 2018 presents a type hierarchy for nominal phrases that is partially reproduced in Figure 9. The type *nominal-parameter* subsumes

¹³ The GENDER feature in the index concerns natural gender. As such, it contrasts with the GENDER feature in the HEAD value of common nouns, which concerns grammatical gender.

nominals that are not explicitly quantified. Technically, its characteristic properties are spelled out in the following implicational constraint.¹⁴

(42) Nominal Parameter



In plain English, phrases of type *nominal-parameter* share the index (*i*) of their head daughter (1) and the set of restrictions on their denotation is the union of the restrictions that hold for the daughters.

At the next level we differentiate a number of subtypes, depending on the semantic relation between the daughters. The most common subtype is that of intersective modification, as exemplified by *red box*, which denotes entities which are boxes and which are red. To model this we use the constraint in (43), quoted from Van Eynde 2018, 14, where it is called restrictive modification.

(43) Intersective Modification

intersective-modification $\Rightarrow \begin{bmatrix} DTRS \langle [SYNSEM | ... | INDEX i], 1 \rangle \\ HEAD-DTR 1 [SYNSEM | ... | INDEX i] \end{bmatrix}$

What this adds to (42) is that the head daughter also shares its index (i) with its non-head sister. As such, it contrasts with combinations in which the non-head daughter has another index than the head daughter. This is the case in NPs which contain a possessive. To model it we introduce the type *possessive-modification* and define it as in (44).

¹⁴Boxed Greek characters stand for sets of objects. Recurrence stands for token-identity, i.e. sharing.

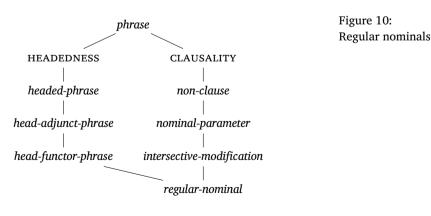
(44) **Possessive Modification**

$$possessive-modification \Rightarrow \begin{bmatrix} synsem | ... | restr & \left\{ \begin{bmatrix} poss-rel \\ Possessor & j \\ Possessed & i \end{bmatrix} \right\} \\ DTRS & \left\langle \begin{bmatrix} synsem | ... | INDEX & j \end{bmatrix}, \boxed{1} \right\rangle \\ HEAD-DTR & \boxed{1} \begin{bmatrix} synsem | ... | INDEX & i \end{bmatrix} \end{bmatrix}$$

This type subsumes all instances of NP-internal possessives, including postnominal PP[*van*]s, genitives and NPs with a possessive determiner. It deliberately leaves the MARKING value of the possessive undefined, since it may as well be in a determiner position (being *marked*) as in a position where it intermingles with other adjuncts (being *unmarked*). Notice, for instance, that the Dutch and English possessive determiners are marked, while their Italian counterparts are not, as illustrated by their compatibility with a preceding determiner, as in *la mia tavola* 'the my table', and their admissibility in postnominal position, as in *tesoro mio* 'treasure my'.

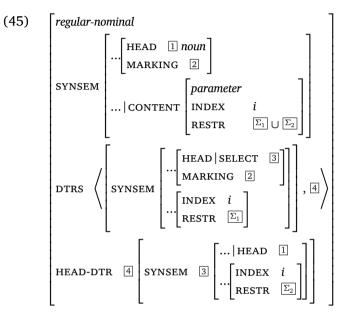
Multiple inheritance

Since the HEADEDNESS dimension is orthogonal to the CLAUSAL-ITY dimension, it is possible to define types that inherit properties from types of either dimension. This is known as multiple inheritance. Exploiting this possibility, Van Eynde 2018, 15 defines a type,



4.3

called *regular-nominal*, that is a subtype of *head-functor-phrase*, on the one hand, and *intersective-modification*, on the other hand, as made explicit in Figure 10. The properties of this type are spelled out in (45).¹⁵



All of the constraints in (45) are inherited from the supertypes of *regular-nominal*. First, being a headed phrase, the HEAD value of the phrase is identical to that of its head daughter (1) (Head Feature Principle). Second, since the non-head daughter is an adjunct, the MARKING value of the phrase is identical to that of its non-head daughter (2) (Marking Principle). Third, since the non-head daughter is a functor, the latter's SELECT value is identical to the SYNSEM value of the head daughter (3). Fourth, being a nominal parameter, its HEAD value is of type *noun*, its CONTENT value is of type *parameter*, its index is shared with the head daughter (*i*), and its RESTR(ICTION) set is the union of the RESTR values of the daughters. Fifth, being an instance of intersective modification, the phrase also shares its index with the non-head daughter (*i*). An

¹⁵ This is a slightly simplified version of that in Van Eynde 2018, 15. The full version also contains the valence features (SUBJ and COMPS).

instance of this phrase type is the combination of *red* with *box* in Figure 8.

In the same way, one can define a type that inherits the properties of *head-functor-phrase* and *possessive-modification*, and that subsumes among others the combination of a possessive determiner or genitive noun with its nominal head sister, as in *his bike* and *Tom's bike*.

Typical of highly regular combinations is that their properties are all inherited from their supertypes. Less regular ones, by contrast, show a mixture of inherited (general) properties and inherent (idiosyncratic) properties. The APC is one of those.

The anaphoric possessive construction

4.4

As demonstrated in Section 3, the APC concerns a juxtaposition of two NPs of which the rightmost one is the head daughter. To model its properties we add a type to the hierarchy of phrases, called *anaphoric*-*possessive*, which is a subtype of *head-independent-phrase*, on the one hand, and *possessive-modification*, on the other hand, as spelled out in Figure 11.

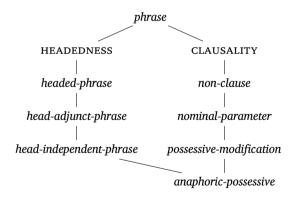


Figure 11: Anaphoric possessive construction

We treat it as a subtype of *head-independent-phrase*, rather than of *head-functor-phrase*, since there is no selection involved. It does not make much sense, for instance, to treat the name *Tom* as a noun that selects *z'n fiets* 'his bike' in *Tom z'n fiets*, since names are fully saturated NPs by themselves.¹⁶ Moreover, we treat *anaphoric-possessive* as a subtype of *possessive-modification*, rather than of *intersective-modification*, since the index of the possessive NP and its head sister are different. Beside the inherited properties, there are some properties that are specific for anaphoric possessives. They are spelled out in (46).

(46) Anaphoric Possessive Construction

The head daughter (1) is required to contain a possessive determiner, as made explicit by its MARKING value (*poss-det*). This blocks the combination with NPs in which the possessor is expressed by a genitive or a PP[*van*], as in **Tom Leo's fiets* and **Tom de fiets van Leo*.¹⁷ Given the presence of a possessive determiner, the RESTR value of the head daughter contains a *poss(essive)-rel(ation)*.

The non-head daughter is required to share the index of the possessor denoting element in its head sister (*j*). The requirement that that index must be of the third person excludes combinations with first and second person forms, as in **ik mijn fiets* 'I my bike' and **jij jouw huis* 'you your house'. The restriction that it must have an animate referent excludes combinations, such as **de wereld zijn eerste multinational* 'the world his first multinational'. Coincidentally, this restriction also

¹⁶Likewise, it does not make much sense to treat z'n fiets 'his bike' as an NP that selects *Tom*, since it is a fully saturated NP by itself.

¹⁷ The constraint could also be captured in another way, invoking Principle B of the Binding theory, according to which referential (i.e. non-anaphoric) NPs must be free (Sag *et al.* 2003, 207). *Tom Leo's fiets* and *Tom de fiets van Leo* are then excluded, since the index of *Leo* must be distinct from that of *Tom*. A problem for this alternative is that it does not exclude **Tom_i de fiets van zichzelf_i/hemzelf_i*.

blocks the combination with personal, reflexive and reciprocal pronouns, as in **hem z'n huis* 'him his house', **zich haar fiets* 'herself her bike' and **elkaar hun boek* 'each other their book'. This is due to the fact that these pronouns are assigned the empty set as their RESTR value in HPSG, reflecting their absence of descriptive content (Pollard and Sag 1994, 250). As such, they do not match the conditions on the non-head daughter in (46), whose RESTR value is required to be non-empty.

A lexicalist analysis

The proposed analysis is constructional in the sense that the defining characteristics of the APC are captured by an implicational constraint on a phrasal type, see (46). It is not impossible to develop a lexicalist alternative. Taking a cue from a reviewer's comment, one could adopt the specifier treatment of determiners (rather than the functor treatment that we adopt) and treat the possessor NP as the specifier of the possessive determiner, yielding a DetP, that is in its turn the specifier of the nominal. In that analysis, the third person possessives can be claimed to select an NP as their specifier, requiring that NP to be animate and sharing its index. In that way, the constraints on phrases of type *anaphoric-possessive* are made part of the lexical entries of the possessive determiners that are used in the APC.

This is, in essence, a variant of the analysis that was described in the opening paragraph of Section 3, and the reasons for dismissing it there also apply to this variant: It violates the constraint that reduced forms of pronouns and determiners cannot take dependents, and it requires special measures to deal with the coordination in *Tom z'n schoenen en z'n laarzen* 'Tom his shoes and his boots'. Besides, it necessitates the postulation of separate lexical entries for possessive determiners that are used in the APC (requiring a specifier) and possessive determiners in other contexts (not requiring a specifier), which is unfortunate since their other properties are the same.

CONCLUSION

The Dutch APC is an example of a construction which has both regular and idiosyncratic properties. To pave the way for a formal analysis we 4.5

5

first described its main syntactic and semantic properties in a theoryneutral way (Section 2) and argued that it consists of two NPs, of which the rightmost one is the head (Section 3). For the analysis, we employed the framework of Constructional HPSG, as defined in Sag 1997 and Ginzburg and Sag 2000. More specifically, we used the bidimensional phrase type hierarchy for nominals developed in Van Eynde 2018 and extended it with a type for the anaphoric possessive construction which, on the one hand, shares a number of properties with its supertypes and, on the other hand, has some properties which are unique to the APC (Section 4). As anticipated in the introduction, the resulting grammar fragment is one "in which the particular and the general are knit together seamlessly" (Kay and Fillmore 1999, 30).

GLOSSES

- ACC accusative
- CMP complementizer
- DAT dative
- DCL declined
- DIM diminutive
- GEN genitive
- PL plural

ACKNOWLEDGEMENTS

This text is based on ideas that I first presented at the 11th International Conference on Construction Grammar, held at the University of Antwerp in the summer of 2021. It was part of the workshop on *Constructional Approaches in Formal Grammar*, organized by Manfred Sailer and Nurit Melnik, who subsequently took care of the editing of this volume for the *Journal of Language Modelling*. They have done an excellent job, both as organizers and as editors. Thanks are also due to the participants of the workshop, to the three anonymous reviewers who gave detailed and very useful feedback on the first version of the text, and to Jan Odijk, who provided me with valuable comments on a later version. If this final version still has any shortcomings, they are entirely my responsibility.

REFERENCES

Valerio ALLEGRANZA (2006), The signs of determination. Constraint-based modelling across languages, Peter Lang.

Liesbeth AUGUSTINUS, Vincent VANDEGHINSTE, and Frank VAN EYNDE (2012), Example-based treebank querying, in *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC)*, pp. 3161–3167, Istanbul.

Hans BOAS and Ivan A. SAG (2012), *Sign-based construction grammar*, CSLI Publications, Stanford.

Hans BROEKHUIS and Evelien KEIZER (2012), Syntax of Dutch. Nouns and Noun Phrases. Volume I, Amsterdam University Press, Amsterdam.

Jonathan GINZBURG and Ivan A. SAG (2000), *Interrogative investigations*, CSLI Publications, Stanford.

Walter HAESERYN, Kirsten ROMIJN, Guido GEERTS, Jaap DE ROOIJ, and Maarten C. VAN DEN TOORN (1997), *Algemene Nederlandse spraakkunst*, Nijhoff and Wolters Plantyn.

Paul KAY and Charles J. FILLMORE (1999), Grammatical constructions and linguistic generalizations: The *what's x doing y*? construction, *Language*, 75:1–33.

Stefan MÜLLER, Anne ABEILLÉ, Robert D. BORSLEY, and Jean-Pierre KOENIG, editors (2021), *Head-driven Phrase Structure Grammar. The handbook*, Language Science Press, Berlin.

Nelleke OOSTDIJK, Wim GOEDERTIER, Frank VAN EYNDE, Louis BOVES, Jean-Pierre MARTENS, Michael MOORTGAT, and Harald BAAYEN (2002), Experiences from the Spoken Dutch Corpus Project, in Manuel GONZALEZ RODRIGUEZ and Carmen Paz SAUREZ ARAUJO, editors, *Proceedings of LREC 3*, pp. 340–347, Paris.

Piet C. PAARDEKOOPER (1984), Beknopte ABN-syntaksis. zevende druk, sterk uitgebreid, Uitgave in eigen beheer, Eindhoven.

Carl POLLARD and Ivan A. SAG (1987), *Information-based syntax and semantics*, CSLI Publications, Stanford.

Carl POLLARD and Ivan A. SAG (1994), *Head-driven Phrase Structure Grammar*, CSLI Publications and University of Chicago Press, Stanford/Chicago.

Ivan A. SAG (1997), English relative clause constructions, *Journal of Linguistics*, 33:431–484.

Ivan A. SAG, Thomas WASOW, and Emily M. BENDER (2003), *Syntactic theory*. *A formal introduction. Second edition*, CSLI Publications, Stanford.

Frank Van Eynde

Frank VAN EYNDE (1999), Major and minor pronouns in Dutch, in Gosse BOUMA, Erhard HINRICHS, Geert-Jan KRUIJFF, and Richard OEHRLE, editors, *Constraints and Resources in Natural Language Syntax and Semantics*, pp. 137–151, CSLI Publications, Stanford.

Frank VAN EYNDE (2006), NP-internal agreement and the structure of the noun phrase, *Journal of Linguistics*, 42:139–186.

Frank VAN EYNDE (2018), Regularity and idiosyncracy in the formation of nominals, *Journal of Linguistics*, 54:823–858.

Frank VAN EYNDE (2021), Nominal structures, in Stefan MÜLLER, Anne ABBEILLÉ, Robert D. BORSLEY, and Jean-Pierre KOENIG, editors, *Head-Driven Phrase Structure Grammar. The Handbook*, pp. 275–313, Language Science Press, Berlin.

Frank VAN EYNDE and Jong-Bok KIM (2016), Loose apposition. A construction-based analysis, *Functions of Language*, 23:17–39.

Gertjan VAN NOORD, Gosse BOUMA, Frank VAN EYNDE, Daniël DE KOK, Jelmer VAN DER LINDE, Ineke SCHUURMAN, Erik TJONG KIM SANG, and Vincent VANDEGHINSTE (2013), Large scale syntactic annotation of written Dutch: Lassy, in P. SPYNS and J. ODIJK, editors, *Essential Speech and Language Technology for Dutch*, pp. 147–164, Springer.

Frank Van Eynde © 0000-0002-6172-442X frank.vaneynde@kuleuven.be

University of Leuven

Frank Van Eynde (2023), The Dutch Anaphoric Possessive Construction, Journal of Language Modelling, 11(2):267–296 (***) https://dx.doi.org/10.15398/jlm.v11i2.337

This work is licensed under the *Creative Commons Attribution 4.0 Public License*. (a) http://creativecommons.org/licenses/by/4.0/