The Dutch evidential NCI

A case of constructional attrition

Timothy Colleman and Dirk Noël
Ghent University and University of Hong Kong

Present-day Dutch has two entrenched "grammatical" hearsay evidentials: a construction with *zou* (originally the past tense form of the verb *zullen*, cognate with German *sollen*) and a construction with *schijnen* (literally, 'seem'). The closest English equivalent of both constructions is the "evidential nominative and infinitive" (NCI), which pairs an evidential meaning with the morphosyntactic pattern [Sʙᴊ *be* X*ed to* Inf]. This is a highly productive construction in English, the most typical instantiation of which is *be said to*. Present-day Dutch has an NCI construction as well, but the lexical possibilities of this construction are limited to a handful of cognition verbs, which – in their NCI use – encode deontic rather than evidential meanings. On the basis of historical corpus data, this paper shows that the Dutch equivalent of English *be said to*, i.e. *gezegd worden te*, looked ready at one time to become entrenched as a substantive hearsay construction as well. This paper traces its evolution and explores the questions of why the pattern disappeared and why Dutch, unlike English, did not develop a schematic evidential NCI construction.

Keywords: nominative-and-infinitive; evidential; diachronic construction grammar; constructional attrition; Dutch

1.   Introduction

The sentences in (1) and (2) are illustrations of a morphosyntactic configuration which at first glance seems to be shared by the grammars of

English and Dutch. It is sometimes referred to as the "nominative and infinitive", or the "NCI", short for a term used in Latin grammar: *nominativus cum infinitivo*.

(1)   Coca-Cola **is said to control** around 80 per cent of the French cola market and Pepsi is keen to strike back. (BNC A7T 354)

(2)   Ze werden populair bij atleten en wielrenners omdat meer rode bloed-lichaampjes **worden geacht** een 'zuurstofvoorsprong' **te geven**. (ConDiv *NRC*_VARIA01)
        'They [EPO hormones] became popular with athletes and cyclists because more red blood cells are thought to give an "oxygen advantage".'

The pattern consists of a subject, the passive form of a verb of a specifiable kind and an infinitive introduced by the infinitival particle specific to each language, i.e. *to* in English and *te* in Dutch.[1] The English pattern is a highly productive one: the verb can be any of those Givón (1990) has grouped as P-C-U verbs, i.e. utterance, cognition and perception verbs which take finite as well as infinitival complements. A search for the NCI in the 100-million-word *British National Corpus* (BNC) produced a list of no less than 112 verbs that can occur in the pattern (Noël 2008). Three more examples are given in (3) to (5). The first two, with *be supposed to* and *be thought to*, illustrate two other high-frequency NCI patterns (in addition to that with *be said to* presented in example (1)) and the third, with *be witnessed to*, a low-frequency one, which truly testifies to the productivity of the pattern.

(3)   "You **are supposed to be** suppressing dust, McAllister, not raising it and spreading it about," he said cheerfully to her back. (BNC HE 428)

(4)   If the subject is instructed to attend to one ear and ignore the sounds coming into the other ear, all the sounds in the attended ear will produce an enhanced N100 component in the ERP. […] This enhancement **is thought to be** associated with the extraction of additional information about the stimuli in the attended ear. (BNC A0T 921)

(5)   Although his yearly stipend only amounted to £6.00 or so in today's values, he **was** soon **witnessed to be** spending the equivalent of thousands of pounds on various good works for his village. (BNC G2Y 44)

The Dutch pattern, on the other hand, is restricted to less than a handful of verbs and can therefore be said to be less a "part of the grammar" than is the English pattern: only the cognition verbs *achten* 'consider', *veronderstellen* 'suppose' and *verwachten* 'expect' occur in the pattern with any frequency in the Dutch ConDiv corpus (Noël and Colleman 2009). Two additional examples are presented in (6) and (7).

(6)   Een abt **wordt verondersteld** in zijn abdij **te vertoeven**. (ConDiv *GVA*2)
        'An abbot is supposed to reside in his abbey.'

(7)   De Rijselse onderzoeksrechter Patrick Keil wil grondige tekst en uitleg bij de resultaten, zodat hij de correcte conclusies kan trekken uit de onderzoeken. De labresultaten **worden verwacht** klaar **te zijn** eind oktober. (ConDiv *LN_ALL*)
        'The "juge d'instruction" from Lille, Patrick Keil, wants the results to be thoroughly explained to him, so that he can draw the right conclusions from the tests. The lab results are expected to be ready by the end of October.'

That, at least, is the situation today; things used to be slightly different. Past diachronic contrastive research on the NCI in English and Dutch revealed opposite evolutions in these languages: whereas English saw a steady increase in the range of verbs that could enter the pattern ever since its introduction into the language, the Dutch pattern initially swelled as well but stopped expanding in the eighteenth century and then dwindled away, leaving only three remnants of what used to be a much more productive pattern: *geacht worden te* 'be considered/supposed to', *verondersteld worden te* 'be supposed to' and (less widely accepted) *verwacht worden te* 'be expected to' (Noël and Colleman 2009). This shows that a construction's movement along the schematicity cline need not be unidirectional: constructions can become more schematic, but they can also return to a less schematic state.

Interestingly, the Dutch cognate of one of the most typical representatives of the NCI pattern in English, *gezegd worden te* 'be said to', disappeared as well, in spite of the fact that at one time it was the most frequently realised NCI pattern in Dutch, with a frequency peak in the eighteenth century (Noël and Colleman 2009). The following is an eighteenth-century example:

(8)   Spreek, Virgilius, wat voor een zeilsteen trekt u herwaarts heen? wie is elders uw vyand? welke schrikdieren nestelen, onder de lauweren uwer kruin? en waarom hier gezogt? dat alleen, in 't Elizeesche

> lommer **wort gezegt** gevonden te moeten worden (W. van Swaanenburg, De herboore oudheit, of Europa in 't nieuw, 1724–1725)
> 'Speak, Virgil, what magnet pulls you here? Who is your enemy elsewhere? What fearful animals are nesting under your crown of laurels? And why look here for what is said to be found only in the Elisean darkness?'

The purpose of this paper is to document the use of this now extinct pattern in Dutch, to try to explain its demise, and ultimately to account for why a construction that is highly productive in English did not quite make it in Dutch. We do this with data from a corpus consisting of the first hundred volumes of the Dutch literary and cultural journal *De Gids* (1837–1936). By way of background, we first sketch the history of the Dutch NCI and explore its current functionality (section 2). We then do a frequency analysis of *gezegd worden te* in the *De Gids* corpus (section 3) and discuss the symbolic value the pattern had in the particular register of nineteenth-century Dutch represented in this corpus (section 4). Next we consider the question of why the pattern did not become an entrenched part of Dutch grammar in spite of its functional potential (section 5). We conclude with a note on the theoretical importance of the case study in terms of justification for the development of "(contrastive) diachronic construction grammar" as a discipline of linguistics, introducing the term "constructional attrition" for the kind of diachronic process observed here (section 6).

2.   The accusative and infinitive, and the nominative and infinitive

Since, from a strictly formal point of view, the NCI is a passive, and is often only conceived of as such, the Dutch historical linguistic literature only mentions the pattern in the margins of accounts of its active formal counterpart, the "accusative and infinitive", or "ACI", short for *accusativus cum infinitivo*. We first summarise these accounts in subsection 2.1, before zooming in on the history of the NCI and its present symbolic value in subsection 2.2.

2.1   The (genuine) ACI in Early Modern Dutch and afterwards

In (9) we list a number of sixteenth- and seventeenth-century examples of the "genuine" or "Latinate" ACI, which was quite common in formal registers in Early Modern Dutch in combination with all sorts of cognition verbs (*geloven*

'believe', *oordelen* 'judge', *erkennen* 'recognise', *denken* 'think', *menen* 'be of the opinion', etc.), utterance verbs (*zeggen* 'say', *heten* 'call', *verklaren* 'declare', *noemen* 'name', etc.) and verbs of (indirect) perception (*vernemen* 'learn', *merken* 'notice', etc.). The English glosses preserve the ACI structure of the Dutch examples, even in cases where this structure is not fully idiomatic in Present-day English, as in (9d) and (9f).

(9)   a.   Men kan immers niet **gelooven** al gantsch Israel soo puyr sot te zijn gheweest. (Coornhert, see WNT sub *geloven*)
'For one cannot believe the whole of Israel to have been so utterly mad.'

b.   Deeltgens die ik **oordeelde** bolletgens te zijn. (Van Leeuwenhoek, cited in Duinhoven 1991: 410)
'Particles which I judged to be tiny spheres.'

c.   Josephus **erkent** deze nederlage te zijn een wraecke. (Vondel, cited in Duinhoven 1991: 410)
'Joseph recognises this defeat to be a heavenly punishment.'

d.   Hier is een edelman, die men **zeidt** goede middel te hebben (Van Reigersbach, cited in Overdiep 1935: 410)
'Here is a nobleman, who one says to have good means.'

e.   Men **hiet** mij doe ter tijt geluckigh te wesen. (Florianus, cited in Fischer 1994: 113)
'They described me to be happy at the time.'

f.   In dus een zee van zorgen, **vernam** zij onder 't volk gezaaid te zijn zeker boek, getiteld… (Hooft, cited in Fischer 1994: 112).
'In such a sea of sorrow, she learnt to be disseminated among the people a certain book, entitled…'

The ACI combines the verb of a governing clause (also known as the matrix verb) with an accusative noun phrase (i.e. in the case of a pronominal NP an objective pronoun is used, as in (9e): *mij* 'me', rather than *ik* 'I') and an infinitival clause of which this NP is the implied subject. The NP and the infinitival clause form a unit with relation to the matrix verb whose meaning is comparable to that of a finite direct object clause. For instance in the case of (9b): *Ik oordeelde de deeltjes bolletjes te zijn* 'I judged the particles to be tiny spheres' ~ *Ik oordeelde dat de deeltjes bolletjes waren* 'I judged that the particles were tiny spheres'.

This genuine ACI should not be confused with a pattern that is sometimes also termed ACI and which is still common in Present-day Dutch in combination with a small group of verbs of (mainly) direct perception, viz. *zien* 'see', *horen* 'hear' and *voelen* 'feel', illustrated in (10).

(10) a.   Ik zag Jan/hem in bad zitten. (Haeseryn et al. 1997: 1011)
            'I saw John/him sit(ting) in the bathtub.'
      b.   Ik hoor Jan/hem piano spelen. (Haeseryn et al. 1997: 1011)
            'I heard John/him play(ing) the piano.'
      c.   Ik voel mijn eksteroog steken. (Haeseryn et al. 1997: 1013)
            'I can feel my corn itch(ing).'

The difference between the types is discussed at length in Klooster (1988), Duinhoven (1991, 1997), Fischer (1994) and Colleman (2007). In a nutshell, what the distinction amounts to is that in the case of the "non-genuine" ACI the accusative NP does not merely function as the implied subject of the infinitive but also (and mainly) as the direct object of the main clause. The infinitive can be thought of as a specification that is added to a sentence that already expresses a complete proposition without it. For instance, *Ik hoor de vogel zingen* 'I hear the bird sing' ~ *Ik hoor de vogel – al zingende* 'I hear the bird – singing' (cf. Fischer 1994: 103). In the case of the genuine ACI, on the other hand, the NP cannot on its own be understood as the notional object of the matrix verb: if the infinitive is left out, what remains is a defective proposition. In (9d), for instance, *een edelman* can hardly be the direct object of *zeggen* (*Men zegt een edelman 'One says a nobleman'). There is also a morphological difference between the two types: the genuine ACI is introduced by the infinitival particle *te*, the other ACI has kept the bare infinitive (see Duinhoven 1997; Colleman 2007).

In contrast to the non-genuine one, the genuine ACI is traditionally considered to be a syntactic borrowing from Latin (see, e.g., De Vooys 1947; WNT s.v. *te* I, B, 6; on the Latin ACI, see, e.g., Schoof 2003). Duinhoven (1991, 1997) has questioned this, arguing that the pattern is already attested in texts from before the Renaissance, and that these Middle Dutch attestations are too numerous for them to be attributable to the influence of Latin (a view not shared by Fischer 1994: 119–120, n. 6). Whatever the case may be, though the genuine ACI was not completely absent from Middle Dutch, it is certain that the pattern proliferated only in the formal written language of the (late) sixteenth and the seventeenth century: witness the many examples from literary and academic texts from this period cited in Duinhoven (1991), Fischer (1994), Overdiep (1935: 406–411), etc. Van Leuvensteijn emphasises that "the ACI hardly ever occurs in the written language of the common man, but mainly appears in literary texts of authors who have had a classical training" (1997: 331; our translation). It is also beyond question that the "popularity" of the construction was relatively short-lived. Fischer (1994) compares the history of the genuine ACI in English and in Dutch and shows that while the pattern became fashionable in both languages in the formal written language of

the Renaissance through the influence of Latin, it became a productive part of the grammar only in English and virtually disappeared again from Dutch after the seventeenth century. Klooster, as well, has stated that "the genuine ACI never got a foothold in the grammar of Dutch" (1988: 272; our translation), unlike in English.

In Present-day Dutch the genuine ACI is at best marginally present. According to Zajicek (1970), there is a small group of cognition verbs – namely *achten* 'consider', *rekenen* 'reckon', and to a lesser degree *oordelen* 'judge' and (*ver*)*onderstellen* 'suppose' – that still enters the pattern sporadically in formal registers (administrative and didactic texts). Relative clauses are said to be more tolerant of the pattern, also allowing ACIs with *menen* 'be of the opinion', for instance. The invented examples he provides, a selection of which is presented in (11), all sound fairly stilted, however, and such examples rarely occur in modern corpora.[2]

(11) a.   Het aantal zwakzinnigen in ons land **acht** men tussen de 150.000 en de 300.000 te liggen. (Zajicek 1970: 203)
'One considers the number of mentally ill in our country to be in-between 150,000 and 300,000.'

   b.   Men **oordeelde** dan ook deze dieren een afzonderlijke klasse te vormen. (Zajicek 1970: 203)
'One judged these animals to constitute a class of their own.'

   c.   Hij stapte in wat hij **meende** de bus naar Oegstgeest te zijn. (Zajicek 1970: 207)
'He stepped onto what he thought to be the bus to Oegstgeest.'

On the other hand, as we summarise in the next subsection, several authors do point out that the passive variant of the pattern has held its ground slightly better.

2.2   The NCI as passive counterpart of the ACI

Quite a few of the sixteenth- and seventeenth-century examples cited in the literature on the ACI contain a passive matrix verb, as in the examples in (12). Such sentences do not of course have an objective NP, but a subjective, or "nominative", one, which is both the subject of the passive main clause and the implied subject of the infinitival complement. The literature occasionally uses the Latin term *nominativus cum infinitivo* or NCI for this passive variant of the genuine ACI, and we have adopted this term throughout our paper (also see Noël 2008; Noël and Colleman 2009, 2010).

(12) a.   Daer over hy dan **geoordeelt wert** de doot weerdig te syn. (Adr. Valerius, 1626, cited in de Korne and Rinkel 1987: 79)

'For which he was judged to be worthy of death.'

    b.   Papier, dat **gesegt is** uyt den Hemel gevallen te sijn. (Van Leeuwenhoek, cited in Damsteegt 1981: 112)
          'Paper which is said to have fallen from Heaven.'

Both Zajicek (1970) and Duinhoven (1991) mention that in Present-day Dutch this passive pattern is slightly more usual than the ACI. Zajicek (1970: 204) prefers the passive matrix in (13a) to the active one in (13b), for instance, while Duinhoven (1991: 425) remarks that the passive in (14a) is perfectly acceptable, whereas the corresponding active sentence in (14b) is "barely possible" (our translation).

(13) a.   De nieuwe belasting **wordt gerekend** een bedrag van +/- 60 m. te zullen opbrengen.
          'The new tax is estimated to bring in around 60 million.'
    b.   *De nieuwe belasting **rekent** men (rekenen wij) een bedrag van +/- 60 m. te zullen opbrengen. (Zajicek 1970: 204)
          'One estimates (we estimate) the new tax to bring in around 60 million.'

(14) a.   Hij **wordt geacht** betrouwbaar te zijn.
          'He is supposed to be reliable.'
    b.   *Ik **acht** hem betrouwbaar te zijn. (Duinhoven 1991: 425)
          'I suppose him to be reliable.'

Otherwise the NCI is not given special attention in the existing Dutch literature: the pattern with the passive matrix verb is merely mentioned as an example of a syntactic environment which was (and to a certain extent still is) more tolerant of the ACI for structural reasons, often in the same breath as another facilitating context, the relative clause.[3]

    For English, Noël (2001, 2008) has developed an alternative hypothesis that the NCI is more than simply the passive counterpart of the ACI and that the NCI is better conceived of as a separate "construction" in the theoretical sense of "construction grammar" approaches (Kay and Fillmore 1999; Goldberg 1995; etc.), or rather as a set of constructions, i.e. as a set of form-meaning pairings which have their own place in the grammar of the language, with specific properties that cannot be reduced to the combined semantics of the ACI and the general passive construction. One type of evidence for this discrete status is the fact that though both the ACI and the NCI are well-established patterns in English, some verbs are much more frequently used in the NCI pattern than in the ACI pattern, while not otherwise displaying any particular preference for the passive (Noël and Colleman 2010). Good

examples of this are *suppose* and *think*, which both occur much more often in sentences of type (a) in (15) and (16) than in sentences of type (b), and *say*, which in Present-day English is incompatible with the ACI (17) (for frequency details, see Noël 2001).

(15) a. By the time the meeting was over I was in such a state of excitement and fright that I forgot all about the clothing parcel I **was supposed** to collect. (BNC B0U 2508)

  b. Traditional readings of the Shipman's Tale, however, **suppose** the husband to be a, or even the, target figure, and not without reason. (BNC HXS 820)

(16) a. Other northern Malawians arrested between February and May **are thought** to be among those set free. (BNC A03 59)

  b. I had always **thought** him to be egotistical and attention-seeking. (BNC AE0 2777)

(17) a. The 680 pastors that serve the region **are said** to preach to well-attended churches on Sundays. (BNC AAB 132)

  b. *One **says** the pastors to preach to well-attended churches on Sundays.

The NCI patterns with these verbs have special semantics. In sentences like (15a) *be supposed to* has a deontic meaning similar to that of the modals *should* and *ought to* and the pattern is consequently often analysed as a "semi-modal" (see, e.g., Depraetere and Reed 2006: 272; on the development of this "deontic NCI", see Noël and van der Auwera 2009). By contrast, *be thought to* and *be said to* in (16a) and (17a), respectively, have an "evidential" function, which we will say more about below. Suffice it to say here that these NCI patterns make a specific contribution to the semantics of the sentence which is shared neither by the ACI nor by the general passive construction and this special functionality explains their higher frequency.

  The Dutch NCI has a functionality comparable to the English pattern, but, as already indicated, it is much less often recruited to perform these functions, in terms of both token and type frequency. In a fragment of about 12.5 million words of the newspaper component of the ConDiv corpus (Grondelaers et al. 2000) used for the study reported on in Noël and Colleman (2009), only three NCI patterns were found: *geacht worden te* 'be considered/supposed to' occurred 44 times, *verondersteld worden te* 'be supposed to', ten times, and there was one example of *verwacht worden te* 'be expected to'; other corpora have thrown up more examples of the pattern, however, the example in (20a) originating from the 38-million-word INL

corpus).[4] The majority of these attested NCI patterns have a deontic function, similar to that of *be supposed/expected to* in English. Examples are given in (18) to (20), in which the infinitival complement expresses some sort of obligation resting on the referent of the implied subject. As the (b) sentences make clear, the conversion to the ACI pattern always produces an ungrammatical sentence, which supports the analysis that the (a) sentences are not merely passive versions of active sentences.

(18) a.   Chefs van afdelingen **worden geacht** excessief of nutteloos surfen te voorkomen. (ConDiv *NRC* Handelsblad)
      'Heads of departments are supposed to prevent excessive or useless surfing.'
      b.   *We **achten** / De directie **acht** chefs van afdelingen excessief of nutteloos surfen te voorkomen.
      'We suppose / The management supposes heads of departments to prevent excessive or useless surfing.'

(19) a.   Een abt **wordt verondersteld** in zijn abdij te vertoeven. (ConDiv Gazet van Antwerpen)
      'An abbot is supposed to reside in his abbey.'
      b.   *De orde **veronderstelt** een abt in zijn abdij te vertoeven.
      'The order supposes an abbot to reside in his abbey.'

(20) a.   Het steekt bij ons dat we **worden verwacht** om dingen te doen waarvoor we gewoon geen tijd hebben. Vooral de individuele aandacht voor kinderen die moeite hebben de lessen bij te houden, gaat al snel verloren. (INL 38MWC *Meppeler Courant*)
      'It galls us that we are expected to do things which we simply have not got the time for. Giving individual attention to kids who are experiencing difficulties in keeping up is something which is easily discarded.'
      b.   *Ze **verwachten** ons dingen te doen waarvoor we gewoon geen tijd hebben.
      'They expect us to do things which we simply haven't got the time for.'

All three attested patterns can also be used evidentially, however, as shown for NCI *verwachten* and NCI *achten* in the earlier examples (2) and (7), respectively, and for NCI *veronderstellen* in the following Internet example.

(21)   De belangen van personages blijken uit hun handelingen en die roepen bij het publiek bepaalde affectieve disposities op jegens hen: positief (met empathie als gevolg) of negatief (onverschilligheid of

anti-empathie als reactie). Ook het uiterlijk van personages **wordt verondersteld** een rol te spelen in die dispositie, hetgeen nadrukkelijk bevestigd is in experimenteel onderzoek. (www.ethesis.net/dramafilms/dramafilms_inhoud.htm)

'The interests of characters are obvious from their actions, which evoke certain affective dispositions towards them in the audience: positive ones (resulting in empathy) or negative ones (indifference or anti-empathy as a reaction). Also the outward appearance of characters is assumed to play a part in this disposition, which has been confirmed by experimental research.'

Though the NCI cannot therefore be considered to be a productive pattern in Dutch, a few lexical NCI islands stayed behind in the grammar, so to speak. They are not matched by corresponding ACI patterns and have a specific semantic value. Earlier corpus research has confirmed, however, that more types of the pattern used to be quite common at one time, that with the highest frequency being *gezegd worden te* (Noël and Colleman 2009, 2010). Examples (8) and (12b) above suggest that this lexical pattern could be used evidentially. We now examine a snapshot of its history, first looking at its frequency in a corpus of nineteenth- and early twentieth-century Dutch and the evolution of its frequency during the period covered, and then consider its functionality.

3.   The incidence of *gezegd worden te* in nineteenth-century Dutch

The decision to focus on the nineteenth (and early twentieth) century, rather than on the eighteenth century, which saw the highest token frequency of the Dutch NCI and of *gezegd worden te* in particular (Noël and Colleman 2009, 2010), was very much determined by the availability of a large corpus of formal written Dutch covering this period. This is a corpus that contains the full text of the first hundred volumes, from 1837 till 1926, of the literary/cultural journal *De Gids* ('The Guide'), totalling approximately 73 million words. Using the *WordSmith* concordancing tool (Scott 2004), we queried the corpus for the term *gezegd* followed by the term *te* within a span of ten words, which, after manual filtering, produced 488 attestations of NCI *zeggen*. A few randomly chosen examples are presented in (22).

(22) a.   De krachtige regels toch, waarmede BYRON de Tiberstad aanspreekt, kunnen niet **gezegd worden** alleen Rome te verheerlijken. (*De Gids* 1844)

'The forceful Lines with which Byron addresses the city on the Tiber cannot be said to only extol Rome.'

b.   Ook de Nederrhijnsche of Duitsche school **wordt gezegd** eene dochter der Byzantijnsche te zijn geweest. (*De Gids* 1845)
'Also the Lower Rhine or German school is said to have been a daughter of the Byzantine one.'

c.   Laten wij nooit vergeten, dat de oorlog, gevoerd tegen Cromwell in de twee jaren van 1652 tot 1654, **gezegd wordt** meer gekost te hebben dan de tachtigjarige oorlog tegen Spanje. (*De Gids* 1852)
'Let us never forget that the war waged against Cromwell in the two years from 1652 to 1654 is said to have cost more than the Eighty Years' War against Spain.'

d.   Zoo vereeren de Alfoeren van de Minahasa, bepaaldelijk die van den stam der Tombulu, eene godin Lumimuut, die **gezegd wordt** uit de aarde te zijn voortgekomen. (*De Gids* 1891)
'For example, the Alfores of the Minahasa, specifically those of the Tombulu tribe, worship a goddess called Lumimuut, who is said to have sprung from the earth.'

The sheer number of hits suggests that NCI *zeggen* must have been a fairly well-established pattern in nineteenth-century (and early twentieth-century) Dutch, at least in the formal written register we find in *De Gids*. Such a conclusion needs to be nuanced, however, because if we sort these hits chronologically, we can observe a very clear evolution in the frequency of the pattern. Table 1 presents the counts for four 25-year periods and these show an unmistakable downward trend. In the first period *gezegd worden te* occurs 13.86 times per million words, while in the last period this normalised frequency has come down to 1.22 instances per million words. Notice that the total frequency of the form *gezegd* stays more or less the same, so the drop in the NCI pattern cannot be attributed to a drop in the occurrence of the verb generally (there is a very significant difference between the four periods in the observed number of NCI instances compared to all other instances of *gezegd*: χ^2 (df = 3) = 258, p < .0001).

Table 1. Absolute and normalised frequencies of NCI *zeggen* and the verb form *gezegd* in the *De Gids* corpus

[INSERT TABLE 1 ABOUT HERE]

NCI *zeggen* is therefore shown to have been relatively common still in the first half of the period covered by the corpus, though the fall in its frequency had clearly already started then. It plummets after that, to (almost) completely evaporate in the course of the twentieth century. Searches in corpora of Present-day Dutch have produced no hits, including in the fragment of the ConDiv corpus referred to above, and in the 38-million-word INL corpus (without the juridical component).[5]

Very relevant now for the constructional status of NCI *zeggen* is the question of the extent to which the pattern is still "supported" by the corresponding ACI pattern. Are there still patterns to be found in *De Gids* of the kind [NP *zegt* [NP *te* Inf]]? We examined this for the first 25-year period (in which the NCI is most strongly present) with a query for all forms of the verb *zeggen* followed by *te* within a span of ten words. The decision to restrict this to the first period is a practical one: a query like this produces a huge number of hits that need to be manually filtered. The results are summarised in table 2. ACI *zeggen* is not completely absent in the investigated period, but with only ten attestations (= 0.68 hits per million words) it can definitely be considered to be an extremely rare pattern. Two examples are presented in (23) and (24). The first of these, it can be added, is also the only example in which the subject of the infinitive (*deze bijzonderheid omtrent de H*) is placed in its "normal" position between the matrix verb and the infinitive. In the other nine examples this subject is to be found in matrix clause-initial position, usually taking the form of a relative pronoun, as in (24).

(23)   REITZ, Belga Graeciss. pag. 129, **zegt** deze bijzonderheid omtrent de H den Brabanderen bijzonder eigen te zijn, en dat zij daarin door hunne taalverwanten, de Zeeuwen en de Goudenaars, gevolgd worden. (*De Gids* 1838)
'REITZ, Belga Graeciss. p. 129, reports (lit. 'says') this peculiarity regarding [h] to be typical of the inhabitants of Brabant and that they resemble the inhabitants of Zeeland and Gouda in this.'

(24)   Chanteloup is een lustslot, dat men **zegt** haar behoord te hebben. (*De Gids* 1846)
'Chanteloup is a summer mansion that one says to have belonged to her.'

Table 2. Absolute and normalised frequency of ACI and NCI *zeggen*

[INSERT TABLE 2 ABOUT HERE]

The frequency difference with the passive pattern is evident: in the first 25-year period of the *De Gids* corpus NCI *zeggen* occurs more than 20 times more often than ACI *zeggen*. This means that the use of *zeggen* in this period is very similar to that of *achten* and *voorstellen* in Present-day Dutch and to that of *suppose*, *say*, *think*, etc. in Present-day English: its NCI use far exceeds its ACI use. In the next section we show that like these other NCI patterns *gezegd worden te* was more than just a passive.

4.   The evidential function of *gezegd worden te*

To start the argument, let's consider a first example, from a historical contribution to *De Gids*.

(25)   Die onderhandelingen met Frankrijk hier na te gaan, ligt geheel buiten ons plan; alleen een paar opmerkingen mogen hier nog plaats vinden. *De Witt, die bij de aannadering der Franschen gezegd wordt geheel radeloos te zijn geweest en de eerste aanleiding te hebben gegeven tot eene bezending naar Lodewijk XIV*, schrijft den 18^{den} Junij aan zijnen broeder: …' (*De Gids* 1851; italics ours)
'It is beyond our scope to examine these negotiations with France; there is only space here for a few remarks. De Witt, who is said to have been completely desperate at the approach of the French and to have taken the initiative for sending a delegation to Louis XIV, writes to his brother on 18 June, saying: …'

If we want to convert the fragment in italics to current Dutch usage while retaining the verb *zeggen*, we could use a passive construction with an extraposed *dat*-clause (the equivalent of an English *that*-clause), as in (26a), or a sentence with the impersonal subject pronoun *men* ('one') and an object *dat*-clause, as in (26b).

(26) a.   Er wordt gezegd dat De Witt bij de nadering van de Fransen volkomen radeloos was.
'It is said that De Witt was completely desperate at the approach of the French.'
  b.   Men zegt dat De Witt bij de nadering van de Fransen volkomen radeloos was.
'They say that De Witt was completely desperate at the approach of the French.'

Other modern translations are provided in (27) to (29), respectively containing the forms *schijnen* (literally, 'seem'), *zou* (originally the past tense form of the

verb *zullen*, cognate with German *sollen*) and *naar verluidt* (literally, 'as announced').

(27)   De Witt schijnt bij de nadering van de Fransen volkomen radeloos te zijn geweest.

(28)   De Witt zou bij de nadering van de Fransen volkomen radeloos zijn geweest.

(29)   De Witt was bij de nadering van de Fransen naar verluidt volkomen radeloos.

The sentences in (27) to (29) are all illustrations of evidential constructions, i.e. constructions that give an indication of the source of the information presented in a sentence. More specifically, we are dealing with "hearsay" markers here: speakers report information they have been told, something they have not observed or come up with themselves, but which unspecified sources hold to be the case. The fact that *gezegd worden te* in (25) can be paraphrased by these evidential constructions is evidence that this pattern, too, is an instantiation of an evidential construction. In other words, the pattern could at one time function similarly to *be said to* illustrated in (1) above, or in (31) below, which needs to be contrasted with (30).

(30)   The statement of the Propaganda Minister that Hitler had a "sixth sense" for seeing what remained hidden to ordinary mortals **was** sarcastically **said** by a young secretary to explain his choice of Italy as an ally. (BNC ADD 778)

(31)   INDEPENDENTS GET BREAK FROM BRAVO
        by David Goymour
        BRAVO, a new on-screen booking system which puts British hotels and tourist attractions on travel agents' counters, has entered its launch phase. It **is said to** offer independent hotels the kind of exposure which hotels in big groups can derive from international booking systems – owned, typically, by the big airlines. (BNC A0C 456)

In (30) the literal meaning of the matrix verb *say* forms part of the explicitly stated meaning of the sentence; its meaning is "foregrounded", or is "on-stage" (cf. Langacker 1987). The sentence refers to a situation where a certain individual has said something. The explicit reference to a speaker in a prepositional phrase (*by a young secretary*) and the presence of a manner adverbial (*sarcastically*) highlight the fact that the sentence reports on a

specific communicative event. The NCI pattern is used here simply for discourse-pragmatic, information-structural reasons, i.e. to topicalise the Propaganda Minister's statement by placing it in sentence-initial position. In other words, it basically instantiates the general passive construction, or somewhat more accurately, it combines an ACI construction with the passive construction.

In (31), on the other hand, we have a typical example of the evidential use of *be said to*. Here there is no reference as such to a specific spatio-temporally locatable utterance. The sentence containing the pattern does not report on a particular communicative event. Instead, the NCI is a means to indicate that the information that is presented can be attributed to an external source. Its effect is two-sided and somewhat paradoxical: on the one hand, the suggestion of an external source underscores the reliability of the information (it is not something the writer has made up); on the other hand, the source indication removes the writer's responsibility for the accuracy of the information. In other words, the pattern makes a text appear more objective, but at the same time offers the writer a backdoor escape, as it were, when it comes to guaranteeing the truthfulness of the information offered. Not surprisingly, it is especially typical of journalistic texts (Noël 2008), a genre in which subjectivity generally needs to be avoided and in which the accurateness of the presented information needs to be guaranteed to the maximum feasible extent. This kind of use of the pattern, or the realisation of what in previous work we have called "the evidential NCI construction", occurs much more frequently than the "plain passive" NCI illustrated in (30) (Noël 2008).

The Dutch example in (32), from a review of a book by a physics professor, is comparable to the English example in (30) in that the NCI is used to report on a specific statement by a specific individual. The writer of the review even provides the page number where the paraphrased assertion is to be found and stresses from whom it originates in an agent phrase. The infinitival complement expresses the content of the assertion.

(32)   Zoodanig passen toch is juist het aanwenden der analyse op de hypothese, welke zoo te regt door den Hoogleeraar COP (bl. 25) **gezegd wordt**, het eenige geldige middel te zijn, om over mogelijkheid en waarschijnlijkheid te kunnen oordeelen. (*De Gids* 1846)
'Yet very fitting is the application of the analysis to the hypothesis, which is rightly said by Professor Cop (p. 25) to be the only valid means to decide on possibility and probability.'

As is the case for *be said to* in English, however, most instances of NCI *zeggen* in the *De Gids* corpus occur in sentences that do not refer to a specific, spatio-temporally locatable utterance act. Some examples are (33) to (36):

(33)   Zelf een grondig beoefenaar der wetenschap, beminnaar aller schoone kunst, verzamelde hij alle dichters en geleerden aan zijn glansrijk hof, die in zijn eigen of in naburige landen zich beroemd hadden gemaakt. De spraakkunst van Panini of de zoogenaamde "groote spraakkunst", de *Mahabhashya*, **wordt gezegd** door hem in Kasjmir te zijn ingevoerd. (*De Gids* 1867)
'A very thorough practitioner of science himself, and a lover of fine arts, he gathered all poets and men of learning at his magnificent court that had become famous in their own or in neighbouring countries. Panini's grammar, or the so-called "extensive grammar", the *Mahabhashya*, is said to have been introduced by him in Kashmir.'

(34)   Evenwel beginnen reeds de zoovele vreemde musici die hier concerten of opera's geven, den smaak te wijzigen, en weldra zal het ook op dit gebied zijn tout comme chez nous, vooral bij het heerschen en het steeds meer door- en voordringen van het Duitsche element. Niet minder toch dan 200,000 Duitschers **worden gezegd** in New-York gevestigd te zijn. (*De Gids* 1874)
'Yet the many foreign musicians who perform concerts and operas here are already starting to change musical tastes and soon it will be tout comme chez nous in this area as well, especially with respect to the dominance and pervasiveness of the German element. No fewer than 200,000 Germans are said to be living in New York.'

(35)   En de brieven? Ergens in de Amsterdamsche Universiteitsbibliotheek kan men een verzegelden trommel vinden, die **gezegd wordt** de kostbare documenten te bevatten. (*De Gids* 1899)
'And the letters? Somewhere in the Amsterdam University Library one can find a sealed box which is said to contain the valuable documents.'

(36)   Engeland wordt in Mexico vertegenwoordigd door Sir Lionel Carden, een diplomaat die **gezegd wordt** Latijnsch-Amerika te kennen op zijn duim. (*De Gids* 1913)
'England is represented in Mexico by Sir Lionel Carden, a diplomat who is said to know Latin America like the inside of his pocket.'

In all these instances NCI *zeggen* signals that the author has a source for the presented information, and that consequently the responsibility for its accuracy is not (solely) his/hers. If we remove *gezegd worden te* from the last example, for instance, and change the infinitive clause to a finite clause, as in (36') below, we end up with a sentence that conveys exactly the same information, but without the distancing evidential qualification.

(36')   Engeland wordt in Mexico vertegenwoordigd door Sir Lionel Carden, een diplomaat die Latijnsch-Amerika kent op zijn duim.
        'England is represented in Mexico by Sir Lionel Carden, a diplomat who knows Latin America like the inside of his pocket.'

In other words, in examples (33) to (36) the proposition conveyed is expressed by the infinitival clause and the passive matrix merely serves as a qualification of that proposition. These NCIs therefore conform to one of the semantic recognition criteria for evidentials listed by Anderson (1986: 274), and endorsed by De Haan (1999: 75): "Evidentials are not themselves the main part of the clause."

Interestingly, one of the examples of *gezegd worden te* from *De Gids* shows very nicely that for nineteenth-century Dutch speakers there was a functional similarity between this pattern and evidential *zou*. Example (37) is from a book review in which a few inadequacies in language and style are pointed out. The reviewer explicitly presents NCI *zeggen* as an alternative to *zou*.

(37)   Op bl. 63 van het Iste Stuk lezen wij: *Omtrent de list, door den Koning KAREL zullende zijn in het werk gesteld*, enz., voor: *omtrent de list, welke door Koning KAREL ZOU zijn in het werk gesteld*, of *wordt gezegd te zijn in het werk gesteld*, enz. (De Gids 1841)
       'On p. 63 of Part I we can read: *Regarding the ruse that will have been set in motion by King Charles* etc., instead of: *regarding the ruse that [literally:] would have been set in motion by King Charles*, or *is said to have been set in motion* etc.'

The boundary between the plain passive NCI and the evidential NCI construction is a fuzzy one, however, and attestations of the NCI pattern cannot always be categorised unambiguously as either the one or the other. Rather, there is a continuum from most propositional / least qualificational to least propositional / most qualificational along which individual attestations need to be positioned. In a sense the basic meaning of the verb *zeggen*, or *say* for that matter, already has an evidential component: though a sentence like (38) below reports on the fact that somebody, i.e. Jan, has said something, it can also be interpreted to mean that Jan is the source of the proposition "Ajax

is still the best Dutch football club", in which case the reliability of this information can be dependent on Jan's reputation as a football buff. This distinction corresponds to the one made between "qualificational" and "non-qualificational" uses of cognition and utterance verbs in Aijmer (1972: 39) and Nuyts (2001: 109–122), or the one made in Verhagen (2005: chapter 3) between "objective" complement clause constructions – in which the matrix clause describes an event (of saying, thinking, etc.) – and "intersubjective" ones – in which the primary function of the matrix clause is to invite "an addressee to identify with a particular perspective on an object of conceptualisation that is itself represented in the embedded clause" (2005: 79).

(38)   Jan heeft gisteren gezegd dat Ajax nog steeds de beste club van
         Nederland is.
         'John said yesterday that Ajax Amsterdam is still the best club of the
         Netherlands.'

In passive sentences with *zeggen* the speaker is left unmentioned, which can easily lead to a hearsay effect. This is also the case in patterns other than the NCI, as in (39), a corpus example that has the same structure as the invented example (26a) above.

(39)   Op straat worden [de artiesten] nagekeken door de politie. Je struikelt
         daar als het ware over de agenten. Er wordt **gezegd** dat Oezbekistan
         de hoogste politie-dichtheid per vierkante kilometer in centraal Azië
         heeft. (ConDiv *De Telegraaf*)
         'In the street the artists are checked by policemen. You keep falling
         over policemen there. It is said that Uzbekistan has the highest density
         of police per square kilometer in central Asia.'

As Cornelis (1997: 226) has put it, the passive of verbs like *zeggen* can function as "a means to abstract away from a particular sayer and his or her words".[6] This is also what happens in (39), in which a passive matrix clause attributes the proposition presented in the *dat*-clause to an unspecified source.

Compared to this construction with a passive matrix and extraposed subject clause, the NCI offers an important advantage, however. To explain which, let's return to example (25) from *De Gids* offered at the start of section 4, a simplified version of which is presented in (40a), and compare this with the paraphrase that was offered in (26a), repeated here as (40b).

(40) a. De Witt **wordt gezegd** bij de nadering van de Fransen geheel
         radeloos te zijn geweest.
         'De Witt is said to have been completely desperate at the approach
         of the French.'

b.   Er **wordt gezegd** dat De Witt bij de nadering van de Fransen geheel radeloos was.
'It is said that De Witt was completely desperate at the approach of the French.'

While *De Witt* in (40b) functions as the subject of the *dat*-clause, it is the subject of the matrix clause in (40a). As a result, De Witt is more unequivocally presented as the "topic" or the "theme" of the utterance in the sentence with the NCI than in the other one; i.e., the NCI allows speakers to start their sentence in the most natural of possible ways with an NP whose referent is what the sentence is about. It is this combination of properties that gave NCI *zeggen* an edge, especially over the ACI pattern: it makes it possible not only to demote the information source but also to promote the topical referent from the qualified statement to subject position in the matrix clause.

Naturally, this apparently favourable situation raises the question of why the Dutch NCI practically disappeared.

5.   Why did the Dutch NCI fade away?

Though earlier research revealed the NCI to have had a frequency peak in the eighteenth century and to already be on the decline in the period covered by the *De Gids* corpus (Noël and Colleman 2009, 2010), data from this corpus confirm that the pattern was still more diverse in the nineteenth century than is currently the case. Some of the examples found with verbs other than *zeggen* are the following ones with *geloven* 'believe', *beweren* 'claim' and *verhalen* 'narrate'.

(41)   … een geestelijke, die […] met den hoogsten eerbied al de reliquiën toonde, die van de heilige zaken overig waren, of vromelijk **geloofd werden** overig te zijn, van nagels van het kruis af tot melaatsche overblijfsels toe. (*De Gids* 1839)
'… a clergyman who reverentially showed all relics remaining of the saint, or which were piously believed to be remnants, from nails of the cross to leprous remains.'

(42)   Aangenaam wordt ons oog verrast door groene stippen in de verte; het zijn tamariskenstruiken, waaruit des morgens eenig vocht loopt, eene soort van gom, welke **beweerd wordt** de manna te zijn waarmede de Israëlieten in de woestijn gevoed werden. (*De Gids* 1869)
'Our eyes are pleasantly drawn to green dots in the distance; they are tamarisk shrubs, which emit some fluid in the morning, a kind of

gum, which is claimed to be the manna with which the Israelites were fed in the desert.'

(43)   Vooral op het terrein van belang en hartstocht kunnen soms coalities worden gesmeed, wier aanrakingspunt in een gemeenschappelijken afkeer tegen eene derde grootheid moet gezocht worden. Eene coalitie, als ons **verhaald wordt** getroffen te zijn tusschen Herodes en Pilatus. (*De Gids* 1862)
'Coalitions can especially be formed at the levels of self-interest and passion, whose point of contact needs to be sought in a common aversion to a third party. A coalition of the kind reported to us to have been founded between Herod and Pilate.'

Such examples are scarce, however. A search in five mid-nineteenth-century volumes of *De Gids* (1850–1854) for the past participle forms of the 31 P-C-U verbs listed in note 4 followed by the term *te* within a span of ten words produced NCI cases only for the verbs listed in table 3.

Table 3. Frequency of the NCI in five volumes of *De Gids* (1850–1854; 2.99 million words)

[INSERT TABLE 3 ABOUT HERE]

Both the fact that only 11 of the 31 NCI patterns searched for were found to be present in the sample looked at and that the frequencies of the patterns found were low are clear indications that the NCI was not a well-established pattern in nineteenth-century Dutch. Only the frequency of *geacht worden te*, a pattern that has survived to this day, matches the frequency of *gezegd worden te*; the nine other patterns have a much lower frequency. In other words, its eighteenth-century frequency peak is not likely to have been high enough to firmly entrench the pattern in the grammar of Dutch. The dwindling frequency of NCI *zeggen* in the course of the period covered by the *De Gids* corpus also confirms this, as does the merely residual presence of NCI patterns in Present-day Dutch. The question, therefore, is why in Dutch, unlike in English, the NCI did not become a firmly entrenched part of the grammar.

This question perhaps needs to be split into two separate questions. Firstly, why was *gezegd worden te,* which appeared to have become entrenched as an evidential construction, not joined by a "critical mass" of NCI patterns needed to ensure the entrenchment of a schematic evidential NCI in Dutch? And secondly, why did *gezegd worden te* itself disappear, in spite of

having survived for some time without the "support" of an ACI counterpart? An additional question, of course, is how come three patterns, *geacht worden te*, *verondersteld worden te* and to a lesser extent *verwacht worden te*, now form part of the grammar of Dutch, supported neither by a schematic NCI, nor by an ACI construction.

We can only offer speculative answers, but it is very likely that the disappearance of the Dutch ACI was a major factor preventing the creation of a schematic evidential NCI in Dutch. In English both the ACI and the plain passive NCI always remained productive patterns, whose presence formally supported the evidential NCI construction, for, though we have argued the latter to be semantically independent from both the ACI and the plain passive NCI, the fact remains that the existence of these morphosyntactic configurations made the evidential NCI quite "normal" grammar. In Dutch the ACI and the plain passive NCI had already become fairly "unusual" grammar in the nineteenth century (Noël and Colleman 2010) and this very likely stood in the way of the evidential NCI becoming more schematic (i.e. with the disappearance of the ACI and the plain passive NCI, the "source" for new substantive evidential patterns dried up, so to speak, before the above-mentioned critical mass needed to ensure the entrenchment of a schematic evidential was attained). *Gezegd worden te* held out a little longer than most other Dutch NCI patterns, but its frequency was already dwindling in the nineteenth century, after having enjoyed a sort of heyday as an evidential construction in the eighteenth century.

Another factor that probably contributed to nipping the development of a Dutch schematic evidential construction in the bud is the existence in Dutch of competing evidential constructions which not only convey the same hearsay meaning as NCI *zeggen* but also share the information-structural advantage offered by the NCI of presenting the topic of the qualified proposition in subject position. The strongest competition probably came from the constructions with *schijnen* and *zou* illustrated in (27) and (28) above.[7] These constructions were already available in the nineteenth century; indeed, they were already more frequent than evidential *gezegd worden te*. This is supported by a spot-check we carried out of one mid-nineteenth-century volume of *De Gids*, viz. the 1850 volume. NCI *zeggen* occurs 15 times in this volume, 6 instances of which are very clear hearsay markers. By contrast, 34 examples were found of evidential *zou*, 2 of which we present here as (44) and (45), and 31 examples of evidential *schijnen*, illustrated in (46) and (47).

(44)   Wij hebben in de laatste tijden veel gelezen en gehoord over het regt van souvereiniteit, dat het huis van Oranje over ons land **zoude** toekomen. (*De Gids* 1850)

'Lately, we have read and heard a great deal about the right to sovereignty of our country, which is said to fall to the House of Orange.'

(45)   Zijne hoofdbewijzen zijn, dat er enkele dieren voorkomen, die pas in den lateren tijd, na de schepping van den mensch, ontstaan kunnen zijn. Zoo bij voorbeeld **zou** de oinopota cellaris een diertje zijn, welks pop alleen in wijn en bier leeft; een ander schepseltje **zou** alleen in chocolade gevonden worden, en eene soort van mot **zou** slechts gevolde en bereide wol, maar nimmer de nog ruwe stof aangrijpen. (*De Gids* 1850)
'His main evidence is that there are some animals which can only have come into existence in later times, after the creation of man. The oinopota cellaris, for instance, is said to be a little animal whose pupa only lives in wine and beer, another creature is said to be found only in chocolate, and a particular kind of moth is said to only harm fulled and processed wool, but never the raw material.'

(46)   Dit was welligt in Mozes' tijd nog het geval niet, ofschoon het ondenkbaar is, dat de vallei in den regentijd niet van water zou zijn voorzien geweest. Overigens **schijnt** deze beek niet de eenige oorzaak der vruchtbaarheid van den Wady Feiran te zijn, maar de eigenaardige gesteldheid van den bodem daartoe veel bij te dragen. (*De Gids* 1850)
'This was probably not yet the case in Moses' time, although it is inconceivable that the valley would have been lacking in water during the rain season. Besides, this stream is said not to have been the only source of the fertility of the Wadi Feiran, but the peculiar condition of the soil (is said) to have contributed a lot to it.'

(47)   … eer men op het denkbeeld kwam, dat men daartoe geen geheelen kogel, maar ook slechts een kogelsegment behoefde: iets dat eerst in de 11de eeuw **schijnt** ontdekt te zijn. (*De Gids* 1850)
'… before one realised that one did not need a whole bullet for that, but that a segment would suffice; something which is said to have only been discovered in the 11th century.'

The English evidential NCI, on the other hand, did not meet with similar competition for its functional niche, i.e. competition from constructions which share its evidential and information-structural potential.[8]

Finally, *geacht worden te* and *verondersteld worden te* have probably survived to this day because, as we have shown in previous work (Noël and Colleman 2009), they are mainly used as deontic constructions. They can also

be used evidentially, but the patterns probably owe their presence as evidential constructions in Present-day Dutch to their deontic sister constructions. Of the 44 tokens of NCI *achten* found in the fragment of the ConDiv corpus referred to above, 33 are clearly deontic in meaning (see (18a) above for a good example) and only 7 are clearly evidential (see (2) above for an example), with the remaining 4 cases allowing both interpretations. NCI *veronderstellen* occurred 10 times in the same corpus fragment, and 9 of those examples are deontics of the kind illustrated in (19a) above; the only remaining example is again ambiguous between both readings. Like their English constructional equivalent *be supposed to*, deontic *geacht worden te* and *verondersteld worden te* in all probability fill a functional niche which is not optimally served by other deontic constructions (see, e.g., Collins 2009: 81 on how *be supposed to* complements *should* and *ought to*). The English construction could even be a factor in the retention of the two Dutch patterns.[9] Similarly, we would hypothesise both evidential and deontic *verwacht worden te* to be fairly recent additions to the language resulting from the influence of English *be expected to*. All this must remain a matter for future research, however.

6.   Conclusion

In Early Modern Dutch the genuine accusative and infinitive and its formal passive, the nominative and infinitive, regularly occurred with all kinds of verbs of cognition, perception and communication in the literary and academic texts of classically trained authors. The same thing happened in Early Modern English, but while in this language these patterns became thoroughly entrenched, they only occupy a marginal position in the grammar of Present-day Dutch, where the ACI is well-nigh absent and the NCI reduced to a small handful of patterns which mainly (though not exclusively) instantiate deontic constructions: *geacht worden te*, *verondersteld worden te* and *verwacht worden te*. The Dutch cognate of English *be said to*, *gezegd worden te*, at one time regularly occurred as a hearsay evidentiality marker, the way *be said to* is still used today, but the pattern never became sufficiently entrenched to survive next to constructions filling the same functional niche.

On the theoretical level, the case study presented here underscores the need for a diachronic construction grammar, or a contrastive diachronic construction grammar even, to form a framework of investigation, in order to more fully understand the current composition of the "constructicon" (cf. Goldberg 2003) of a language. Traditional grammaticalisation theory, which concentrates on how lexical material turns into (partially) substantive grammatical constructions (cf. Noël 2007), can only offer part of the

explanation of what set of constructions end up being used on a regular basis by the speakers and writers of a language, so that they can be said to be part of the grammar. The pool of constructional resources of a language, even the subset which is traditionally termed "grammatical", is far greater than the totality of those resulting from grammaticalisation proper. Grammatical constructions can be borrowings, for instance, rather than grammaticalisations, the evidential NCI being a case in point (Noël and Colleman 2009). Such borrowings can then develop into highly productive schematic constructions, like the English evidential NCI, or they can fail to become entrenched and become less productive again, like the Dutch evidential NCI. We could term the former development "constructionalisation" or "schematisation" (cf. Noël 2007), the latter "constructional attrition". Contrastive diachronic construction grammar can account for such opposite developments in different languages with reference to differences in the number of constructions which are competing for the same position in their semantically organised constructicons.

Acknowledgements

Notes

1. Dutch, like English, has an analytic passive, which is formed by combining the past participle of the main verb with a passive auxiliary (viz. *worden* 'become' in imperfect tenses and *zijn* 'be' in perfect tenses; see Cornelis 1997 on the difference between the two). Past participles are formed by circumfixing *ge- … -d/-t* to the verbal root: e.g. *achten* 'consider' > *ge-acht*, *zeggen* 'say' > *ge-zeg-d*. The *ge-* prefix is absent in the participle form of prefixed verbs, e.g. *veronderstellen* 'suppose' > *veronderstel-d*.

2. For instance, a query in the full INL 38-million-word corpus (minus the legal component) for all occurrences of any form of the verb *achten* 'consider' followed by the infinitival particle *te* with a maximum of ten intervening words only threw up a single ACI instance, viz. (i) below (next to many passive, i.e. NCI, instances, see the next subsection). The example in

question is from the fairly high-brow political periodical *Liberaal Reveil*. We refer to Kruyt and Dutilh (1997) for detailed information on the INL corpora.

> (i) Onder aanroeping van het subsidiariteitsprincipe **achten** de meeste Nederlanders het voor de hand te liggen dat elke lidstaat zijn eigen infrastructuur moet financieren. (INL 38MWC *Liberaal Reveil*)
> 'Referring to the principle of subsidiarity, most Dutchmen consider it to go without saying that each member state should finance its own infrastructure.'

3. Both Duinhoven (1991: 410–411, 423–424) and Fischer (1994: 114–118) relate this greater tolerance to the fact that in passives as well as in relative clauses the implied subject of the infinitival clause is not in its basic position between the matrix verb and the infinitival clause, as a consequence of which the construction is less blatantly at odds with the "normal" rules of Dutch grammar.

4. This test started out from a set of 31 P-C-U verbs which were quite frequently used with ACI/NCI syntax in Early Modern Dutch according to Fischer (1994) and the other historical studies cited above and/or which are claimed to be still compatible with ACI/NCI syntax (in facilitating syntactic environments) in Present-day Dutch in Zajicek (1970) or Duinhoven (1991). We searched for present-day NCI instances via queries for the past participle forms of the test verbs followed by the infinitival particle *te* within a distance of maximally ten intervening words. Except for the three verbs mentioned in the main text, none of the test verbs was encountered with NCI syntax a single time. The 28 remaining verbs are listed and glossed here:

> *aantonen* 'demonstrate', *begrijpen* 'understand', *beoordelen* 'judge', *berekenen* 'calculate', *beschouwen* 'consider', *beseffen* 'realise', *bevinden* 'find', *beweren* 'claim', *bewijzen* 'prove', *denken* 'think', *erkennen* 'recognise', *geloven* 'believe', *hopen* 'hope', *houden* 'hold', *menen* 'be of the opinion', *onderstellen* 'suppose', *oordelen* 'judge', *rekenen* 'reckon', *schatten* 'estimate', *verklaren* 'declare', *vermoeden* 'suspect', *vernemen* 'learn', *verstaan* 'understand', *vertellen* 'tell', *vinden* 'find', *wanen* 'imagine', *weten* 'know', *zeggen* 'say'.

5. Google web searches do produce a number of hits, but these are usually from translations of English texts (like example (i) below; notice the URL) or from academic texts where the synchronic influence of English is probably substantial (example (ii)). On this basis the NCI with *zeggen* can hardly be said to be a common pattern in Present-day Dutch.

> (i) Binnen die bol zijn er roosters en krachtenvelden waardoor de levering van het oordeel **wordt gezegd te** passeren.
> (http://www.shangrala.org/M_Dutch%20Translations/Wie_GroteGoddelijkeBestuurder .html; accessed on 01.08.2008)
> 'Inside this sphere there are grids and powerfields through which the delivery of judgement is said to pass.'

(ii) Dat is dan ook waaruit Wittgensteins bijdrage aan de filosofie (voor zover hier nog sprake kan zijn van een positieve bijdrage) vaak **wordt gezegd te** bestaan.
(http://dare.ubvu.vu.nl/bitstream/1871/11536/1/bax.pdf; accessed on 01.08.2008)
'This is exactly what Wittgenstein's contribution to philosophy (to the extent that we can speak of a positive contribution here) is often said to consist of.'

6. Cornelis makes this point in the context of a theory of percolation, building on work by Spooren and Jaspers (Spooren 1989; Spooren and Jaspers 1990). In a given piece of (argumentative) discourse, the information presented in an embedded clause under a matrix verb of utterance or cognition may or may not be associated with ("percolate upwards to") the speaker's/author's standpoint. The use of a passive matrix verb typically signals that the information should not percolate to the author's own individual standpoint. There may be various reasons why the author should want to give such a signal: s/he does not want to identify with the "sayer" because s/he has a different opinion, s/he wants to present the embedded information as generally accepted (i.e. not just the words of a particular "sayer"), etc. See Cornelis (1997: chapter 6) for further elaboration.

7. The full set of competitors is larger, as other hearsay evidential constructions exist in Dutch, such as *naar verluidt*, mentioned above and illustrated in (29), *volgens X* 'according to X' and *naar men zegt* 'they say' (somewhat more literally, 'as they say'), but these have an adverbial nature. Similar evidential adverbials exist in English.

8. *Schijnen* literally translates as 'seem', but though *seem* can be used evidentially in English, it is not a hearsay evidential.

9. An anonymous reviewer raises the question whether it might not be the case that *geacht worden te* and *verondersteld worden te* are undergoing the same process of liquidation as *gezegd worden te*, but at a slower pace. The corpus data do not permit a conclusive answer to this question. As a spot-check, we investigated the occurrence of NCI *geacht worden te* – the most frequent of the two substantive patterns – in two five-year subperiods of the *De Gids* corpus with a fifty-year interval, viz. 1850–1854 (comprising 2.99 million words of text) and 1900–1904 (comprising 3.97 million words of text). We found 43 NCI instances in the former subperiod (= 14.38 instances per million words) versus 38 NCI instances in the latter (= 8.06 instances per million words), which, indeed, suggests a downward trend. However, the overall frequency of the participle form *geacht* dwindles as well (from 193 to 120 instances), so that, in contrast to what we found for *zeggen* 'say', there is no statistically significant decrease in the proportion of NCI to other uses of the *achten* 'suppose' verb. In any event, if NCI *achten* and *veronderstellen* are in the process of disappearing from the language, too, this process is far from completed, as these patterns are still relatively frequently encountered in Present-day Dutch (cf. section 2).

References

Aijmer, Karin. 1972. *Some Aspects of Psychological Predicates in English.* Stockholm: Almqvist and Wiksell.

Anderson, Lloyd B. 1986. Evidentials, paths of change, and mental maps: Typologically regular asymmetries. In: Wallace Chafe and Johanna Nichols (eds.). *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, NJ: Ablex, 273–312.

Colleman, Timothy. 2007. *Ik zie hem morgen (te) komen*. Een "echte" accusativus-cum-infinitivo met zien. *Taal and Tongval* 59, 3–29.

Colleman, Timothy, and Dirk Noël. 2009. *Gezegd worden + te*-infinitief: een verdwenen evidentiële constructie. *Tijdschrift voor Nederlandse Taal- en Letterkunde* 125, 385–403.

Collins, Peter. 2009. *Modals and Quasi-modals in English.* Amsterdam: Rodopi.

Cornelis, Louise. 1997. *Passive and Perspective*. Amsterdam: Rodopi.

Damsteegt, B.C. 1981. *Van Spiegel tot Leeuwenhoek*. Leiden: Repro RU.

De Haan, Ferdinand. 1999. Evidentiality in Dutch. *Proceedings of the 25th Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society, 74–85.

de Korne, Annet, and Tineke Rinkel. 1987. *Cursus zestiende- en zeventiende-eeuws Nederlands*. Groningen: Wolters.

De Vooys, C.G.N. 1947. *Nederlandse Spraakkunst*. Groningen: Wolters.

Depraetere, Ilse, and Susan Reed. 2006. Mood and modality in English. In: Bas Aarts and April McMahon (eds.). *The Handbook of English Linguistics*. Oxford: Blackwell, 269–290.

Duinhoven, Anton M. 1991. *Dat siet men wit ende reine wesen*. A.c.i.-constructies in het Nederlands. *Nieuwe Taalgids* 84, 409–430.

Duinhoven, Anton M. 1997. *Middelnederlandse Syntaxis, Synchroon en Diachroon. Deel 2: De Werkwoordgroep*. Groningen: Martinus Nijhoff.

Fischer, Olga. 1994. The fortunes of the Latin-type accusative and infinitive construction in Dutch and English compared. In: Toril Swan, Endre Mørck and Olaf Jansen Westvik (eds.). *Language Change and Language Structure: Older Germanic Languages in a Comparative Perspective*. Berlin and New York: Mouton de Gruyter, 91–133.

Givón, Talmy. 1990. *Syntax: A Functional-typological Introduction*, vol. 2. Amsterdam and Philadelphia: John Benjamins.

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Grondelaers, Stefan, Katrien Deygers, Hilde van Aken, Vicky Van den Heede and Dirk Speelman. 2000. Het ConDiv-corpus geschreven Nederlands. *Nederlandse Taalkunde* 5, 356–363.

Haeseryn, Walter, Kirsten Romijn, Guido Geerts, Jacques de Rooij and M.C. van den Toorn (eds.). 1997. *Algemene Nederlandse Spraakkunst*. Second edition. Groningen and Deurne: Martinus Nijhoff and Wolters Plantyn.

Kay, Paul, and Charles J. Fillmore. 1999. Grammatical constructions and linguistic generalizations: The What's X doing Y? construction. *Language* 75, 1–33.

Klooster, W.G. 1988. Historische en systematische verklaringen. *Tijdschrift voor Nederlandse Taal- en Letterkunde* 104, 269–277.

Kruyt, J.G. and M.F.W. Dutilh. 1997. A 38 Million Words Dutch Text Corpus and its Users. *Lexikos* 7, 229–244.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar, vol. 1: Theoretical Prerequisites*. Stanford: Stanford University Press.

Noël, Dirk. 2001. The passive matrices of English infinitival complement clauses. Evidentials on the road to auxiliarihood. *Studies in Language* 25, 255–296.

Noël, Dirk. 2007. Diachronic construction grammar and grammaticalization theory. *Functions of Language* 14, 177–202.

Noël, Dirk. 2008. The nominative and infinitive in Late Modern English: A diachronic constructionist approach. *Journal of English Linguistics* 36, 314–340.

Noël, Dirk, and Timothy Colleman. 2009. The nominative and infinitive in English and Dutch: An exercise in contrastive diachronic construction grammar. *Languages in Contrast* 9, 144–181.

Noël, Dirk, and Timothy Colleman. 2010. Believe-type raising-to-object and raising-to-subject verbs in English and Dutch: A contrastive investigation in diachronic construction grammar. *International Journal of Corpus Linguistics* 15, 157–182.

Noël, Dirk, and Johan van der Auwera. 2009. Revisiting *be supposed to* from a diachronic constructionist perspective. *English Studies* 90, 599–623.

Nuyts, Jan. 2001. *Epistemic Modality, Language and Conceptualization: A Cognitive-pragmatic Perspective*. Amsterdam and Philadelphia: John Benjamins.

Overdiep, G.S. 1935. Zeventiende-eeuwsche Syntaxis (Derde Stuk). Groningen: Wolters.

Schoof, Susanne. 2003. Impersonal and personal passivization of Latin infinitive constructions: A scrutiny of the structures called AcI. In: Jong-Bok Kim and Stephen Wechsler (eds.). *Proceedings of the Ninth International Conference on Head-driven Phrase Structure Grammar*. Stanford: CSLI Publications, 293–312.

Scott, Michael. 2004. *WordSmith Tools* version 4. Oxford: Oxford University Press.

Spooren, Wilbert. 1989. Some aspects of the form and interpretation of global contrastive coherence relations. Ph.D dissertation, University of Nijmegen.

Spooren, Wilbert, and Joep Jaspers. 1990. Tekstoperaties en tekstperspectieven. *Gramma* 14, 195–220.

Van Leuvensteijn, J.A. 1997. Vroegnieuwnederlands (circa 1550–1650). In: M.C. van den Toorn et al., *Geschiedenis van de Nederlandse Taal*. Amsterdam: Amsterdam University Press, 227–272.

Verhagen, Arie. 2005. *Constructions of Intersubjectivity: Discourse, Syntax, and Cognition*. Oxford: Oxford University Press.

WNT = Matthias De Vries, Lambert Te Winkel et al. (eds.). 1882–1998. *Woordenboek der Nederlandsche Taal* [Dictionary of the Dutch Language]. 's Gravenhage: Martinus Nijhoff.

Zajicek, Jacques. 1970. Réflexions sur l'accusativus cum infinitivo. *Nieuwe Taalgids* 63, 198–208.

Table 4. Absolute and normalised frequencies of NCI *zeggen* and the verb form *gezegd* in the *De Gids* corpus

	size subcorpus (number of words)	*gezegd worden te*		*gezegd*	
		absolute frequency	frequency/ million words	absolute frequency	frequency/ million words
Period I (1837–1861)	14,788,104	205	13.86	3,637	245.94
Period II (1862–1886)	22,132,530	197	8.90	4,797	216.74
Period III (1887–1911)	20,336,520	66	3.25	4,563	224.37
Period IV (1912–1936)	16,348,275	20	1.22	3,733	228.34
Total	73,605,429	488	6.63	16,703	226.91

Table 5. Absolute and normalised frequency of ACI and NCI *zeggen*

	ACI *zeggen*		NCI *zeggen*	
	absolute frequency	frequency/ million words	absolute frequency	frequency/ million words
Period I (1837–1861)	10	0.68	205	13.86

Table 6. Frequency of the NCI in five volumes of *De Gids* (1850–1854; 2.99 million words)

P-C-U verb	NCI frequency
achten 'consider'	43
bevinden 'find'	6
bewijzen 'prove'	7
denken 'think'	2
erkennen 'recognise'	1
houden 'hold'	1
oordelen 'judge'	2
rekenen 'reckon'	12
verklaren 'declare'	4
veronderstellen 'suppose'	5
zeggen 'say'	45