<table>
<thead>
<tr>
<th>Title</th>
<th>Same formal pattern, different contact situation, different propagation: evidential vs. deontic NCI constructions in Dutch (contrasted with English)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Noel, D; Colleman, T</td>
</tr>
<tr>
<td>Citation</td>
<td>The 2014 Workshop on Grammatical Hybridization and Social Conditions, Leipzig, Germany, 16-18 October 2014.</td>
</tr>
<tr>
<td>Issued Date</td>
<td>2014</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/10722/206867">http://hdl.handle.net/10722/206867</a></td>
</tr>
<tr>
<td>Rights</td>
<td>This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.</td>
</tr>
</tbody>
</table>
Same formal pattern, different contact situation, different propagation

Evidential vs. deontic NCI constructions in Dutch (contrasted with English)

Dirk Noël & Timothy Colcman

This paper is about the Dutch equivalent of the pattern illustrated by the English sentence *John is believed to be in Leipzig*. This is a specific type of what is sometimes called “subject-to-subject” raising. Another name for it is “nominative and infinitive”, or NCI in short, the abbreviation of the Latin term “nominativus cum infinitivo”. 
Believe-type “raising to subject”, or NCI, patterns

1. In this book authorities are said to be limited also by the kinds of reasons on which they may or may not rely in making decisions and issuing directives, and by the kind of reasons their decisions can pre-empt.

2. AMERICAN ring doughnuts from The Delicious Donut Co are made from a flour which is said to give them a light, fluffy, and non-greasy consistency.

3. A market can be said to be a place where buyers and sellers meet to make an exchange of goods (or services).

We have argued in previous work that NCI patterns can instantiate no less than five different constructions, i.e. they can have five different functions, both in English and in Dutch (Noël & Colleman 2010).

We call the NCI in (1) a “plain passive” (even though no active “equivalent”, i.e. an ACI, with say is possible in Present-day English) because the NCI is used here for the same information packaging reasons that motivate the use of the passive generally (compare: The authors of this book say that authorities…). The meaning of the matrix verb (said) is very much part of the propositional meaning of the sentence, which refers to a spatiotemporally locatable utterance act: a statement made “in this book”. In this example the meaning of are said to is very much “on-stage” (cf. Langacker 1987): if you remove it, the sentence becomes nonsensical, or at the very least conveys an altogether different meaning (?In this book authorities are limited also by...).
In (2) and (3), on the other hand, the meaning of the *be said to* pattern is “qualificational”. These sentences do not report on a specific utterance act. The meaning of *be said to* is “off-stage” here and you can safely remove the pattern, so that *said* can be argued to no longer be a matrix verb, but to be part of an auxiliary-like construction that modifies the meaning of the infinitive. (2) and (3) have in common that the NCI is not used for information structural reasons (since it can be left out: *...a flour which gives them...*, *A market is a place...*), but they differ in that they illustrate two different form-meaning pairings. In (2) the modifying construction has an evidential function, i.e. its writer uses the pattern to indicate that s/he has a *source* for the information s/he is conveying, so that s/he is not the (sole) judge of the factuality of the statement that American ring doughnuts are light, fluffy, and non-greasy. We have called this “the evidential NCI construction”. Other frequent instantiations of the “evidential NCI construction” in Present-day English are *be alleged to, be assumed to, be believed to, be claimed to, be considered to, be deemed to, be estimated to, be expected to, be felt to, be found to, be held to, be known to, be reported to, be seen to, be shown to, be supposed to, be taken to, be thought to and be understood to*. 

In (3) the modifying construction connects a description with a descriptum. This we have called the “descriptive NCI construction”.

**Believe-type “raising to subject”, or NCI, patterns**

1. In this book authorities are *said to be* limited also by the kinds of reasons on which they may or may not rely in making decisions and issuing directives, and by the kind of reasons their decisions can pre-empt.

2. AMERICAN ring doughnuts from The Delicious Donut Co are made from a flour which *is said to give* them a light, fluffy, and non-greasy consistency.

3. A market can *be said to be* a place where buyers and sellers meet to make an exchange of goods (or services).
the deontic NCI

4. Pupils are expected to use their Maths to solve problems.
5. You were supposed to do six and you only did four!
6. Standards for exposure to benzene are expected by the UK government to be set in 1993: a level of 3 ppb is under consideration, although according to the World Health Organization there is “no known safe threshold dose”.
7. The mechanism supposed by Miller (1948) to underlie acquired equivalence is that introduced by Hull (1939) with his notion of secondary generalization.
8. The Japanese economy is expected to grow by only 3.8% during fiscal 1991 compared with this year’s 5.2%.
9. Napoleon is supposed to have said “An army marches on its stomach.”

Two of the patterns that can instantiate the evidential NCI construction in English, be expected to and be supposed to, can carry the additional meaning illustrated in (4) and (5).

In these examples be expected to and be supposed to instantiate a “deontic NCI construction”. They can be paraphrased with “Pupils should use their Maths to solve problems” and “You should have done six and you only did four”.

Note that the same patterns can also be the realization of a plain passive, as illustrated in (6-7), as well as the evidential NCI construction, (8-9).
the counterfactual NCI

10. The race was supposed to be taking place in blazing sunshine, but the sun refused to come out and there was even some rain.

11. I didn’t think much of the first time really. It wasn’t how I imagined it. I thought you were supposed to enjoy it, and it was half and half really.

A fifth NCI construction can only be realized by be supposed to.

Visconti (2004) terms this the “epistemic” use of be supposed to, because it “evokes a possible world, a state of affairs which would be expected to occur but does not”, often in co-texts containing “counterfactual signals” like but, in fact, in reality, ...

To our mind, however, “counterfactual” is the operative word here, rather than “epistemic”, because no judgement of the (un)certainty or probability of a proposition (cf. van der Auwera & Plungian 1998) is involved.
In Dutch the NCI is a far less frequent and less productive pattern than in Present-day English. The only patterns occurring today with any frequency are geacht worden te (‘be considered/supposed to’), verondersteld worden te (‘be supposed to’) and, to a lesser extent, verwacht worden te (‘be expected to’).

Like the English NCI, the Dutch NCI is not simply a passive. Arguably, the examples in (12-13) are plain passives in that the meanings of achten and veronderstellen are on-stage: these instances report on someone—the organisers of the election in (12), the 19th century practitioners of phrenology in (13)—considering or supposing something to be the case.

Nowadays, Dutch NCI examples most often instantiate a deontic NCI construction, less often an evidential NCI construction and occasionally also a descriptive NCI construction. The counterfactual occurs as well.

‘Heads of departments are supposed to/should prevent excessive and useless surfing.’

15. Iedereen wordt dan verondersteld met het Klein Gevaarlijk Afval naar de containerparken te gaan.

‘Everyone is then supposed to/should then take the Small Toxic Waste to the recycling centre.’
16. Ze werden populair bij atleten en wielrenners omdat meer rode bloedlichaampjes worden geacht een "zuurstofvoorsprong" te geven.

'They [EPO hormones] became popular with athletes and cyclists because more red blood cells are thought to give an "oxygen advantage".'

17. 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.

'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.'
18. Kan de rechterlijke macht, nu in ons constitutioneel staatsbestel alle machten van de natie uitgaan, geacht worden de natie te vertegenwoordigen?

'Since in our constitutional system all powers are vested in the nation, can the judicial power be supposed to represent the nation?'

19. [...] ook de programmering van de radio- en televisieprogramma's mag verondersteld worden een breed publiek aan te spreken.

'... also the scheduling of the radio and television broadcasts may be supposed to appeal to a large audience.'
counterfactual

20. [Het tv-programma bestaat uit] studiogesprekken met ‘Prominenten’ van het vierde garnituur op de bank, die geacht worden interessante dingen over hun seksleven te vertellen aan een zekere Verona Feldbusch.

‘The television show consists of studio talks with fourth-class ‘celebrities’, who are supposed to tell interesting things about their sex lives to one Verona Feldbusch.’ (but it is clear from the context that the writer finds these sex live stories uninteresting)

21. De aanklager werd verondersteld de kinderen te verdedigen.

‘The prosecutor was supposed to defend the children.’ (but it is clear from the context that he did not)
In neither Dutch nor English linguistics the question of whether the NCI is native to Dutch and English or whether it was introduced into these languages as a result of contact with Latin is completely settled yet.

It is undeniable, however, that there have been two waves of Latin influence, a medieval (pre-1500) influence from religious writings, and a 16th-17th-century wave of influence from Classical Roman writing (Noël 2008, Colleman & Noël 2009, 2012).

We have argued in previous work that the presence of the evidential NCI in both English and Dutch is the result of this second wave of Latin influence.

We have already said that the evidential NCI developed into a very productive, schematic construction in English. In Dutch we only have the three patterns just illustrated, but more types of the pattern used to be quite common at one time, the one with the highest frequency being gezegd worden te, the Dutch cognate of be said to.

This shown in this table taken from a previous paper. It displays the frequency in a historical corpus of verbs said in the literature to still occur occasionally in ACI/NCI patterns today and/or to have been most frequent in these patterns in the 17th century + *verwachten*

Notice the frequency peak in the 18th century.

*Gezegd worden te* is responsible for almost half of the total in that period and it indeed looked like becoming entrenched as an evidential expression. Two examples on the next slide.
The entrenchment of *gezegd worden te* as an evidential expression did not carry through, however, and the frequency of the pattern went downhill fast in the 19th century and early 20th century, and has completely disappeared now.

We attributed this in this 2012 paper to competition in Dutch from other Dutch evidential constructions – a kind of competition which the English evidential NCI did not face.


22. *Lelf 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 “grote 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.’

23. *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.’
Where does the deontic NCI come from?

In English this is not a result of language contact.

We have argued in Colleman & Noël (2014), however, that in Dutch it is a result of contact with English: at a time when *geacht worden te* and *verondersteld worden te* were used as evidential constructions, speakers of Dutch who had contact with English connected these two patterns with the English pattern *be supposed to* and they also started using the Dutch patterns with the extra meaning the English pattern had, i.e. a deontic meaning.

We said in this 2014 paper that this influence from English that led to the deontic NCI in Dutch, not only accounts for why we have *geacht worden* and *verondersteld worden* in Dutch today, it also explains why these can still be used as evidentials today, whereas *gezegd worden* cannot anymore. Evidential *geacht worden* and *verondersteld worden* could continue by virtue of the deontic use of the patterns, while *gezegd worden*, which obviously did not acquire a deontic use, disappeared as an evidential as well.

However, this is an explanation that works better for *verondersteld worden* than for *geacht worden*, because *veronderstellen*, like *suppose*, means ‘to believe or assume as true’ (an epistemic judgement), while *achten* means ‘to consider/deem to be’ (an opinion).

In other words, we need a more extensive account for the presence of *geacht worden te* in Dutch, especially because it is considerably more frequent than *verondersteld worden*. What explains this higher frequency?
For this paper we thought we would try and find this explanation using a resource we recently found out about, an online ngram query tool that accesses the digitized newspaper archive of the Dutch Royal Library, containing the full text of a collection of Dutch-language newspapers published between 1840 and 1995.
The ngrams we searched for are two-grams: *geacht te*, *verondersteld te* and *verwacht te*, the three NCI patterns that still occur in Dutch today.
This search produces a graph which visualizes the relative frequencies of these two-grams, relative to all other two-grams present in the texts that were searched.

For our purposes this visualization is not very useful, and in fact rather misleading, because what is displayed here are the relative frequencies of unfiltered query hits. And these two-grams do not always instantiate NCI patterns. This is especially the case for verwacht te, so we should not take the green line very seriously.

What is confirmed by the filtered data, however, is that geacht te is the most frequent NCI pattern (the violet line) and that verondersteld te is much less frequent (the blue line).

When you click on any of the dots in the coloured lines you get to a list of texts containing the query hits and you can then manually filter the hits by looking for them in each text and determine whether it is an NCI or not. This is a very slow and cumbersome process, because this is not a corpus-linguistic tool and you cannot quickly export the query results as a set of concordance lines. So we had to make an initial selection of data to look at.

Obviously the blue line helped in making this selection: it shows that there are years in which verondersteld te hardly occurred at all in the texts searched, so apart from the very first year covered (1840) we only looked at years in which there were more than five hits for verondersteld te ➞ 1869 and 1884 ➞ 15 years between these, and we tried to stick to this interval as closely as possible.
So what did a qualitative examination of the results reveal in terms of why we still have these three NCI patterns in Dutch today, and how they came to be deontic expressions?

Let’s start with the most frequent NCI pattern, *geacht worden te*. 
Initially, in the earliest of the volumes looked at, only plain passive and
descriptive instances occur, which can be translated with ‘to be considered
to be’.

The context is often one of rules and laws, so that ‘statutorily’, or ‘officially’, can
be added to the paraphrase: ‘to be officially considered to be’.

We thought this typical statutory context could be the key to why we have *geacht worden te* still in Dutch while other NCI patterns have disappeared. Our guess is
that the influence of French on Dutch legal and administrative language may
have been at play here.

To this day, Dutch legal language very much bears the stamp of French, and the
cause is obvious: the French civil code, the so-called Code Napoléon. At the start
of the 19th century the Netherlands were French and the Dutch Civil Code
introduced after the French had gone was very much based on this Napoleonic
Code.

The present Dutch Civil Code contains many instances of the kind of *geacht
worden te* illustrated on this slide. The equivalent of it in the Code Napoléon is
the expression *être censé* + adjective/past participle/infinitive. In a juridical
context this means ‘to be considered by law to be the case’/’to be deemed by
the law to be’.
27. L’héritier qui renonce est censé n’avoir jamais été héritier.
   'The heir who renounces a succession is deemed never to have been heir.'

28. Le fossé est censé appartenir exclusivement a celui du côté duquel le rejet se trouve.
   'The drain is deemed to belong exclusively to him on whose side the drainpipe is located.'

These are two examples from the Code Napoléon.

This French pattern came to be used with dynamic infinitives and later grammaticalized into a deontic modal expression.

Did the same happen in Dutch, i.e. did the plain passive “juridical” use of geacht simply grammaticalize into a deontic expression, without foreign influence?

Or is there evidence in the data from the KB historical newspaper archive which supports our polysemy copying hypothesis? We’ll see in a minute.
Let’s first return to *geacht worden te*. Before turning to evidential and deontic examples, we will mention for the sake of completion that in addition to the legalistic examples of the plain passive use of *geacht worden te* which we have already illustrated, there are also examples in which it signals an opinion of an individual, or certain individuals.
The first evidential examples in the data set we have considered appear near the end of the 19th century. These look like “indigenous” examples: there is nothing to suggest contact influence.

At that time, the end of the 19th century, gezegd worden te (‘be said to’) was still around as well, so these evidential examples do not need a special explanation. They are instances of the evidential construction that had been part of Dutch grammar for a few centuries already.
The first undeniably deontic example, however, is found in an advertisement for an American product dating from 1914 and one could reasonably suspect influence from English here.

The next deontic examples of *geacht worden te* in our data occur only fifty years later, however, and they do not show obvious influence from English.
Only in the 1977 data do deontic examples become more frequent than the other uses.

So *geacht worden te* only becomes entrenched as a deontic construction well into the 20th century, and apart from that single advertisement example there is no indication of influence from English in the KB data we have looked at so far.

Things look somewhat different in the case of *verondersteld worden te*, however.
The earliest examples in the KB data we looked at are plain passive ones. These are clearly “indigenous” examples.

The earliest, late 19th-century, evidential examples, however, do reveal possible influence from English in the functional development of this pattern, in that they all occur either in Dutch newspapers published in the US, or in texts that go back to an English-language source.
evidential *verondersteld worden te*

37. Kapt Lapsley’s huis woei om en zijne schoonzuster werd onder de puinhoopen begraven en *wordt verondersteld te zijn* omgekomen. (1884; *De Grondwet* [‘The Constitution’], published in Michigan)

‘Capt Lapsley’s house blew over and his sister-in-law was buried under the rubble and is supposed to have been killed.’

38. Daar die wolk zich terzelfder tijd noord- en zuidwaarts uitbreidt en Krakatau bijna onder de linie gelegen is, werd de richting in het noordelijk halfrond een naar het noordwesten en in het zuidelijk halfrond naar het zuidwesten, ongeveer als de beide benen der letter V, waarvan de punt in Krakatau *wordt verondersteld te liggen*. (1884, *Algemeen handelsblad*, text starts from article in *Nature*)

‘As this cloud expands northwards and southwards at the same time and Krakatau is nearly under the equator, the direction in the northern hemisphere was one to the north-west, and in the southern hemisphere to the southwest, a bit like the two legs of the letter V, whose angle is supposed to lie in Krakatau.’

Here are just two examples, one from a Dutch-language newspaper published in the US, and one from a newspaper published in the Netherlands but from a text that obviously uses the journal *Nature* as its main source.

Examples like these show that English is likely to have been a factor in the survival of *verondersteld worden te* as an evidential pattern.

The first apparently “authentic” evidential example was found in the 1929 data.
Turning to deontic *verondersteld worden te* it is interesting that a different version of the cough remedy ad from 1914 suggests that *verondersteld* was not yet deontic at the time.

Notice the infinitive of the modal *moeten* (‘must’) after *wordt verondersteld te*. For this translator of what is very likely to be an originally English text *verondersteld worden te* on its own obviously did not yet have deontic force.

In the case of *veronderstellen*, though, more deontic, and also counterfactual examples, occur earlier, than in the case of *achten*, in spite of it being a less frequent pattern.

Most of these display an obvious English influence, only some do not.
The first deontic example in the KB data we looked at dates from 1929 and looks genuinely Dutch. What is important for our polysemy copying hypothesis, however, is that it postdates the earliest evidential examples, and that a relevant English model was established for those.

Many of the later examples occur in translations from English, however, which points to the role of English in the entrenchment of deontic *verondersteld worden te* as well.

The first example of a counterfactual NCI in the data occurs in a translation from English. (41)
Here’s two more deontic examples. Both are from translations of English serial novels.

We are including these additional examples from 1948 and 1963 to kind of iconically convey that the less frequent pattern *verondersteld worden te* appears to have been more frequently used deontically before the more frequent pattern, *geacht worden te*, got going as a deontic construction. It could be that the deontic use of *verondersteld worden te*, the pattern which is a more straightforward translation of *be supposed to*, boosted the deontic potential of *geacht worden te*.

In other words, it was perhaps first and foremost *verondersteld worden te* which copied the polysemy of *be supposed to*, and subsequently *geacht worden te* followed suit, or rather, the example of *verondersteld worden te* may have unleashed the deontic potential of *geacht worden te*.

Naturally, this needs more empirical underpinning.
Finally, let’s turn to *verwacht worden te*.

Examples of *verwacht worden te* are even scarcer than those of *verondersteld worden te*, and they all, without exception, occur either in Dutch language newspapers published in the US, or in texts which are about things British or American and whose information is likely to be drawn from English language sources.

This is true of both evidential and deontic *verwacht worden te*, and it is also true of the later examples.

45 is about the New York Stock Exchange.
deontic verwacht worden te

45. Een vereeniging die in haar belofte God de eerste plaats geeft en die voorschrijft dat haar leden worden verwacht te behooren tot een godsdienstig genootschap en de verplichtingen ervan moeten nakomen, […], is niet neutraal te noemen. (1929; De Tijd, text about Scouting)

‘An organization which in its oath gives primacy to God and which prescribes that its members are expected to belong to a religious society and to abide by its rules, […], cannot be called neutral.’

46. Zoo werd Dr. Maurice Heusinkveld, van Greenleafton afkomstig, gevraagd om het werk te aanvaarden, ofschoon hij slechts 6 maanden taal studie had gehad instee van de twee jaren dat een nieuwe zendeling de taal wordt verwacht te studeeren. (1948; De Volksvriend, published in Iowa)

‘Thus Dr. Maurice Heusinkveld, originating from Greenleafton, was asked to accept the job, though he had had only 6 months of language instruction instead of the two years a new missionary is expected to study the language.’
concluding hypotheses

- As the influence of Latin on Dutch tapered from the 18th century, so did the frequency of the Dutch NCI.
- Only three patterns left in Dutch: geacht worden te, verondersteld worden te and verwacht worden te.
- Influence from French ensured the survival of geacht worden te.
- Influence from English (be supposed to) ensured the survival of verondersteld worden te as an evidential expression.
- Verondersteld worden te copied the polysemy of be supposed to and became deontic.
- Geacht worden te copied the polysemy of verondersteld worden te and became deontic.
- Verwacht worden te is a loan translation of English be expected to.

When, in the 18th century, even educated writers started to have less contact with Latin, the NCI pattern started to wither away in Dutch, even though before then a Dutch evidential NCI construction seemed to be taking off.

Only three patterns remain in Present-day Dutch, the most frequent one being geacht worden te. Less frequent is verondersteld worden te, and even less frequent verwacht worden te.

How come we still have these three patterns in Dutch today, while the Dutch cognate of be said to, gezeid worden te, has completely disappeared?

We think this is first of all the result of the influence from French in the 19th century, specifically French legal texts, which boosted the frequency of plain passive geacht worden te, the translational equivalent of être censé. This happened at a time when the Dutch evidential was still more productive than it is now and supported by plain passive geacht worden te evidential geacht worden te kept being a smouldering presence.

Plain passive, juridical, geacht worden te has the potential of developing into a deontic expression. Être censé did so in French, but probably at a time when French no longer exerted the same kind of influence on Dutch as it did in the first half of the 19th century. Did plain passive geacht worden te grammaticalize into a deontic expression in a similar way, on its own steam?

Probably not, because the less frequent pattern verondersteld worden te, seems to have been used more often in a deontic way before geacht worden te took hold as a deontic expression.

Evidential verondersteld worden te seems to have been kept alive through the influence of English, as a translational equivalent of be supposed to. When be supposed to came to be used more and more as a deontic expression, verondersteld worden te copied its polysemy.

Geacht worden te then copied the polysemy of verondersteld worden te.

Verwacht worden te was not around in Dutch until the late 19th century. Its presence in Dutch is entirely due to English influence.
References