



Long distance dependencies in the Germanic sandwich

Jack Hoeksema¹ and Ankelien
Schippers²

(1: Rijksuniversiteit Groningen, 2: Carl von Ossietzky Universität Oldenburg)

Outline

- Long-distance dependencies in English, German and Dutch
- A notorious sentence from the King's Song
- Historical developments in German and Dutch
- *That*-trace effects
- Conclusions

Long-distance movement

- Movement over a finite clause boundary (typically signalled by the presence of *that* (*dat*, *dass*)
 1. [_{CP} Who do you think [_{CP} that Bill saw t_{who}?]
- Main features (cf. Chomsky, 1977):
 - Embedded finite clause contains a gap
 - Filler of this gap is in Spec-CP of a higher (matrix) clause
 - Movement is sensitive to so-called island constraints
 - Occurs in *wh*-questions, relatives, clefts, comparatives and topicalization construction
 - Occurs in quite restrictive contexts (cf. Ross, 1967)

The case of Dutch

- Line from the song composed for the inauguration of King Willem Alexander in 2013:



“De dag die je wist dat zou komen is hier”

The day that you know that would come is here

Den Tag der du wusste dass kommen wurde ist hier

- Example of an LD-relative clause
- Created an enormous uproar in Dutch media, sentence considered to be “ungrammatical” and song “bad Dutch”

Some public opinions

- The Taalprof blog (by Peter Arno Coppen): the sentence is not wrong, similar constructions can be found in (older) literary works (see also Schippers, 2013)
- Reactions to the blog :
lul niet 'taalprof', tekst klopt van geen kanten, 'De dag die je wist dat zou komen', je kan t wel proberen te verdedigen met je goede gedrag, maar t slaat nergens op

“don’t talk crap ‘language prof’, the text does not make any sense “the day that you knew that would come” – you can try and defend it with your good behavior, but it does not make any sense”

De tekst bevat veel te veel Anglicismen. Het lijkt of Ewbank eerst een Engelse tekst had en die door de vertaler van Google gehaald heeft

“The text contains too many anglicisms. It seems as if Ewbank first had an English text and then ran it through Google Translate”

Interview with the culprits

Interviewer: *De dag die je wist dat zou komen, that is not proper Dutch, is it?*

John Ewbank: *Well de dag waarvan je wist dat die zou komen didn't work well (in this song)*



Who is right?

- Everyone!
- LD-relatives are indeed highly unusual in present day Dutch (PDD) , but are frequently attested in older stages of the language.
- The alternative construction with *waarvan* (called “resumptive prolepsis” is indeed much more common in PDD.
- LD-relatives indeed appear to be much more common in English (and thus may sound like an ‘anglicism’).
- Structurally speaking, there is nothing wrong with the sentence...
- ...but it does violate certain “soft” constraints on LD-movement.

Constraints on LD-movement

- Sensitive to the *that*-trace effect [** that* t_{wh}] (effectively banning LD-subject movement over a complementizer)
 - Leads to ungrammaticality in English, but also causes degraded acceptability in Dutch and German (cf. Maling & Zaenen, 1978; Bennis, 1980; Reuland, 1983; Featherston, 2005 and Den Dikken, 2007)

De dag die_i je wist dat t_i zou komen

The day REL you know that would come

- Sensitive to so-called islands (not discussed here)

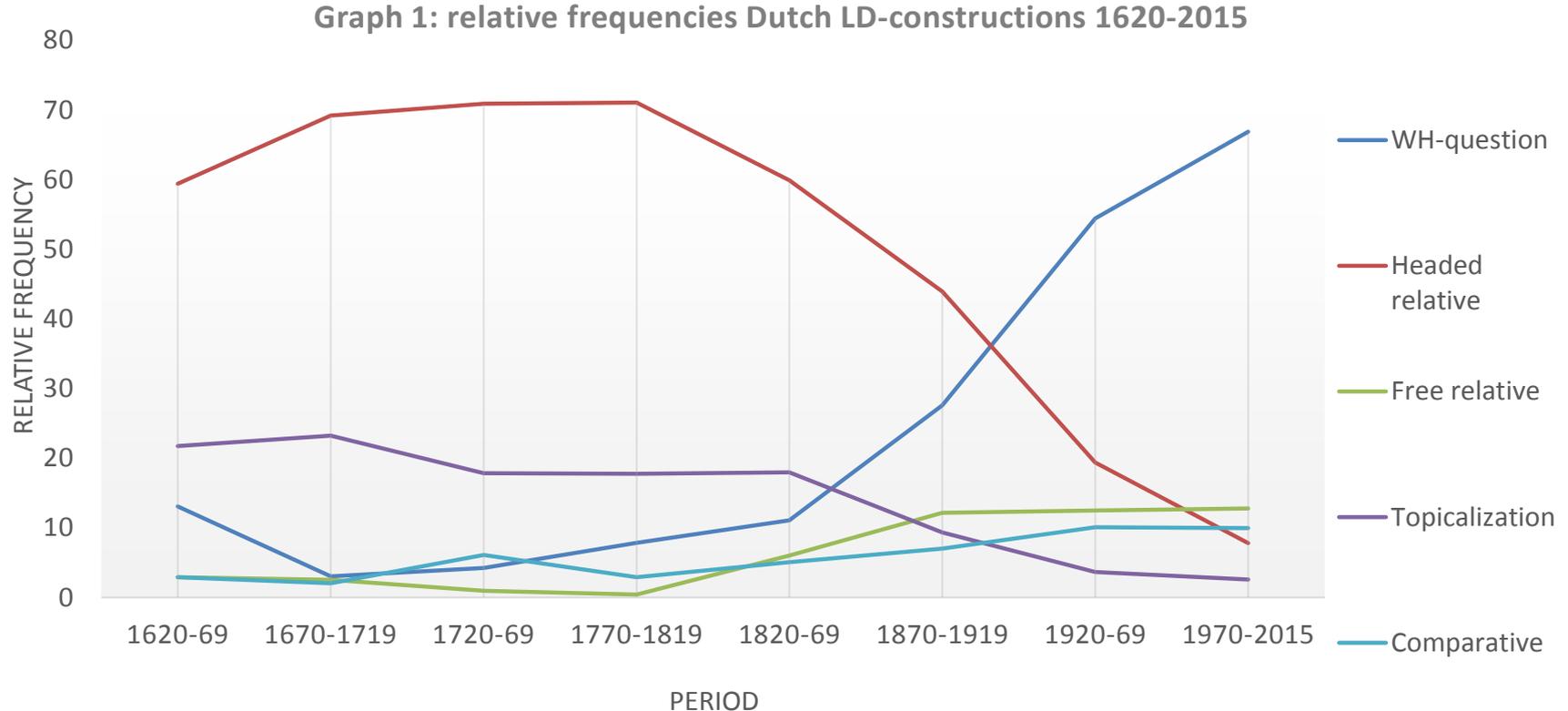
Corpus data on LD wh-movement

- Database of manually collected examples of LD movement in Dutch (ca. 2200, from 14th century to now) and English (ca. 1100, from 18th century to now) (cf. Hoeksema and Schippers, Schippers and Hoeksema and Schippers, 2012)
- Mostly from literary works
- Examples of all main types of LD-dependencies:
 - Wh-questions
 - Relatives (free and headed)
 - Topicalization constructions
 - Comparatives

Dutch data

period	WH	HR	FR	TOP	COMP	Total
< 1620	4	45	1	10	2	62
1620 – 1669	9	41	2	15	2	69
1670 – 1719	6	138	5	46	4	199
1720 – 1769	11	157	2	41	15	226
1770 – 1819	19	172	1	43	7	242
1820 – 1869	24	130	13	39	11	217
1870 - 1919	59	94	26	20	15	214
1920 - 1969	135	48	31	10	25	249
1970 - 2015	478	55	91	18	71	713
Total	744	881	172	242	152	2191

Diachronic developments in long-distance extraction



- Two most frequent constructions (wh-questions and relatives) show an opposite development: wh-questions increase compared to relatives.
- Major change around second half of the 19th century.
- Topicalization constructions patterns with headed relatives.

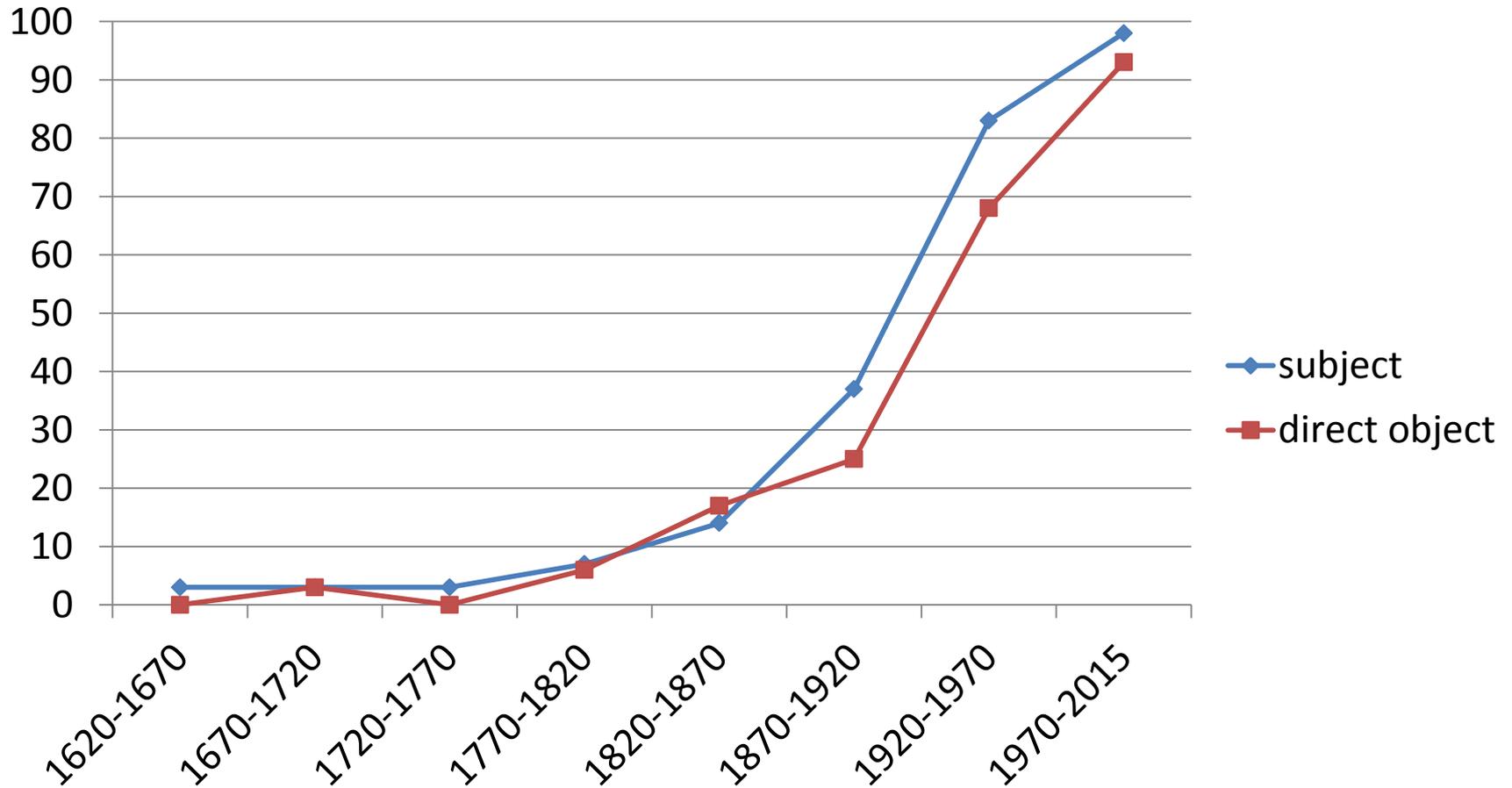
The demise of LD relatives and topicalization

- Hoeksema & Schippers (2012), Schippers & Hoeksema (2010) and Schippers (2012): LD relatives and topicalization constructions disappear at the expense of resumptive prolepsis (here illustrated for LD relatives)

De dag [_{CP} **waarvan** je wist dat **die** zou komen is hier]
the day where.of you knew that [RP] would come is here
(RP = resumptive pronoun)

- Construction shows none of the hallmarks of LD movement (no gap, no *that*-trace, no sensitivity to islands) but is functionally similar to it (cf. Salzmann, 2006).
- Only possible when the proleptic object (i.e. *de dag*) is individual denoting and *d*-linked
- This explains why it is excluded for comparatives and free relatives (where the proleptic object would be null)

Percentage of resumptive prolepsis in Dutch headed relatives with long-distance dependencies



Wh-questions

- Resumptive prolepsis is also possible for wh-questions, so why do these not decline?
- Schippers (2012): corpus data suggests that wh-questions are not d-linked in most cases
- Furthermore, wh-questions increasingly show less lexical variation and are not very productive in that respect (cf. Schippers, 2012; Verhagen, 2006 and Dąbrowska, 2008)
- In ca. 50-90% of the cases (depending on the corpus one looks at) have the form of:
WH denk je [dat ...]
WH think you that

German

- No corpus data available, but descriptive grammars and other sources suggest LD constructions have also by an large disappeared (see Schippers 2012, section 4.4 for a literature overview)
- Timing of the decrease appears to coincide with Dutch change: middle of the 19th century.
- Main difference between Dutch and German: in German, LD wh-questions have also disappeared (especially in (northern) Germany).
- German, unlike Dutch, has alternatives for LD wh-questions: so-called “partial” wh-movement and parentheticals (extraction from embedded V2 clauses)

German alternatives to LD wh-questions

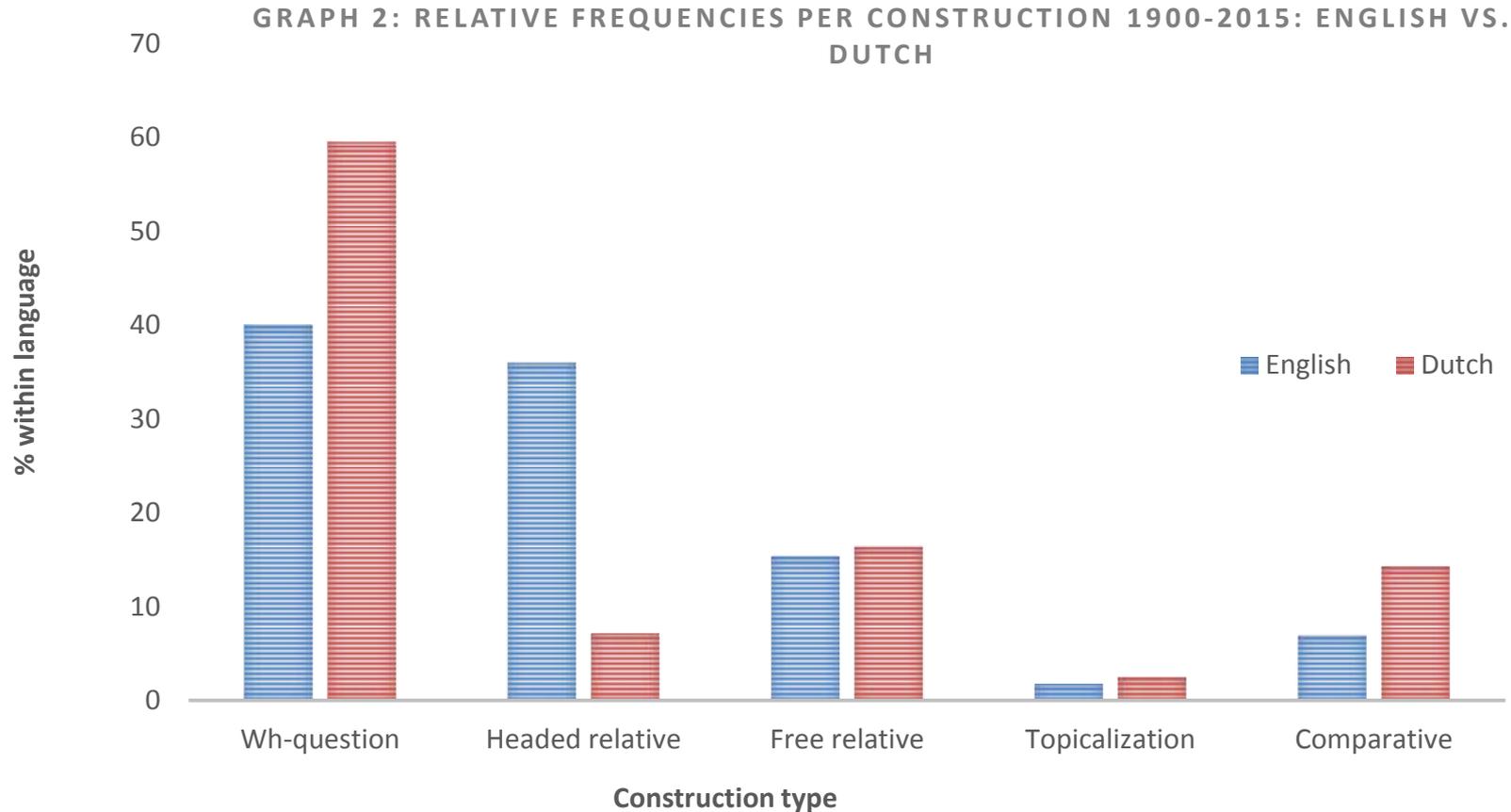
- Long-distance movement
%Wen meinst du dass Gunther t_{wen} angerufen hat?
- Partial wh-movement
Was meinst du wen Gunther t_{wen} angerufen hat
- (integrated) parenthetical (“extraction from embedded V2”, cf. Reis, 1995 & 1996)
Wen (meinst du), hat Gunther t_{wen} angerufen
- Whereas partial wh-movement and parenthetical questions are fully productive in German, they are much less common in Dutch and generally do not function as functional alternatives to LD wh-movement.

What about English?

- English does not show the same asymmetry as Dutch
- LD movement still appears to be productive across the board (judging from the literature and native speakers judgements)

Construction	Frequency	% of total
Wh-questions	459	40.2
Headed relatives	412	36
Free relatives	176	15.4
Topicalization	21	1.8
Comparatives	79	7
It-cleft	4	0.3
Though-movement	1	0.1
Total	1143	100

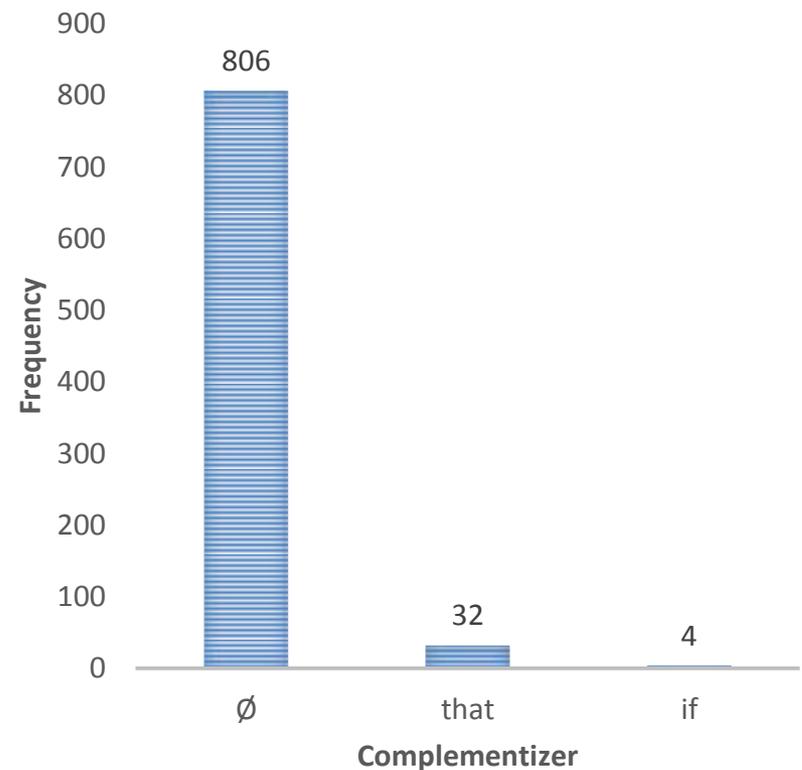
English compared to Dutch (20/21st century)



Is LD-movement really productive in English?

- LD-movement appears more productive in English... However:
- Corpus data reveals that LD-extraction almost invariably involves *that*-less clauses (95% of the cases)!
- Do these examples really involve LD-movement proper?
- Generalized *that*-trace effect?

ENGLISH DATA - PRESENCE OF COMPLEMENTIZER



That-trace in English

Schippers (2012): acceptability ratings on subject and object extraction with and without *that*:

Subject - \emptyset

Who did Robert think will blame Sandra?

Subject - *that*

Who did Sandra think that will blame Robert?

Object - \emptyset

Who did Robert think Sandra will blame?

Object - *that*

Who did Sandra think that Robert will blame?

- Allows us to investigate the effect of the presence vs. absence of the complementizer.
- Comparison with filler items without extraction (declaratives) shows whether extraction and presence of complementizer interact:

Declarative - \emptyset

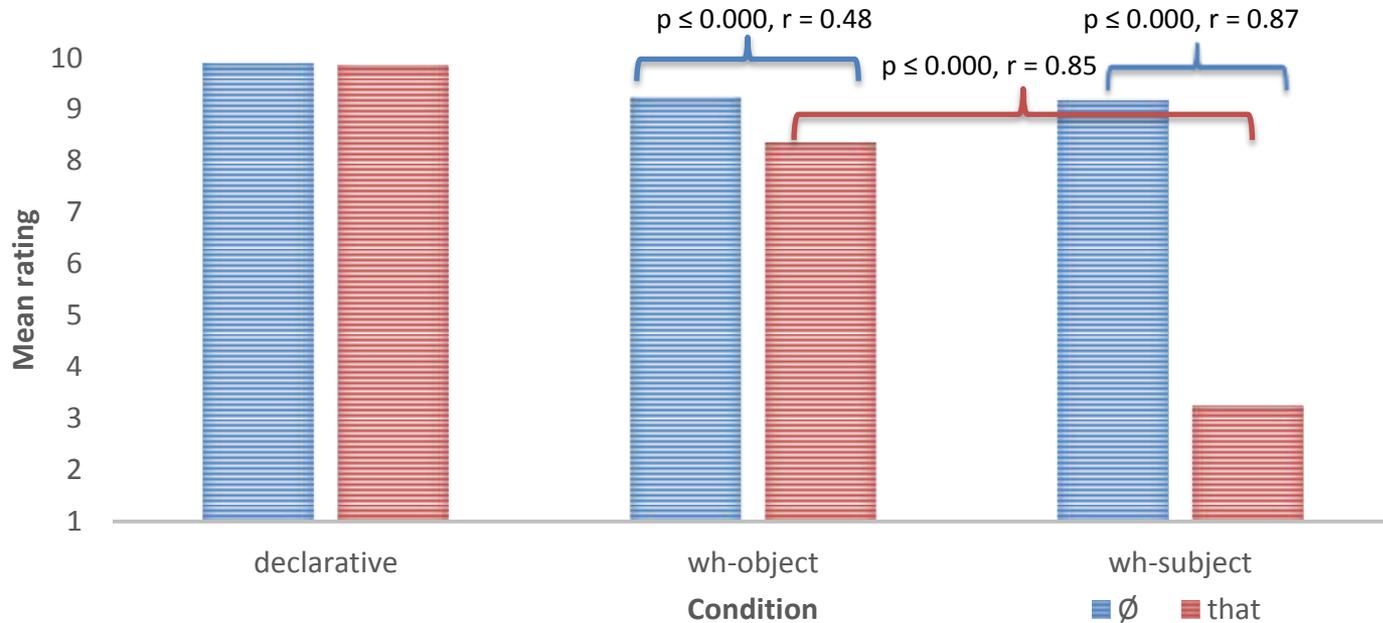
Roland believes Deanna may plant a tree

Declarative - *that*

Hector hopes that Maxine will sell her boat

Results *that*-trace

EXTRACTION AND *THAT* IN ENGLISH



- Extraction causes a significant drop in acceptability (difference between declaratives and extraction constructions always $p \leq 0.000$).
- Under extraction - presence of *that* causes a significant drop in acceptability (already for object extraction!)
- Under subject movement, presence of *that* creates the strongest effect

corpus data

- 1143 occurrences from English
- subject extractions: 341 without complementizer, 0 with complementizer
- all other extractions: 751 without complementizer, 51 with.

Comparison with COCA and BNC:

- COCA: *think that he*: 2059, *think he*: 25232: 92% *that*-drop
- BNC: *think that he*: 166, *think he*: 3067: 95% *that* drop
- *think + that* in LD data: 2, without *that* 428: 99.5% *that*-drop
- Difference COCA and LD data: chi square 29.8, $p < 0.0001$
- Difference BNC and LD data: Fisher's exact test, $p < 0.0001$
- For other verbs we checked, *say* and *hope*, also significant differences between LD cases and general usage data

Conclusions

- Dutch in between German (almost no long-distance movement) and English. English represents the most archaic system, German is the most innovative, and Dutch is in between.
- German has alternatives for WH-movement that can be found in Dutch, but are not yet common
- German and Dutch have replaced long-distance relativization by resumptive prolepsis
- Both Dutch and English have a weak island effect for extraction from headed finite CPs, which is strongest for subjects (“*that*-trace effect”) but can also be noted for extraction of nonsubjects
- English has the option of dropping the complementizer, Dutch replaces extraction from *dat*-clauses whenever possible by alternatives

References

- Chomsky, N. (1977). On Wh-movement. In P. Culicover, T. Wasow & A. Akmajian (Eds.), *Formal Syntax*. New York: Academic Press.
- Dąbrowska, E. (2008). Questions with long-distance dependencies: A usage-based perspective. *Cognitive Linguistics*, 19, 391-425.
- Dikken, M. den (2007). *Questionnaire study on Dutch that-trace effects: stimuli and results*. (Unpublished manuscript). CUNY, New York. Retrieved from http://web.gc.cuny.edu/dept/lingu/dendikken/docs/dutch_that_trace_results.pdf.
- Fanselow, G., Kliegl, R., Schlesewsky, M. (2005). Syntactic variation in German wh-questions. Empirical investigations of weak crossover violations and long wh-movement. *Linguistic Variation Yearbook*, 5, 37-63.
- Hoeksema, J. & Schippers, A. (2012). Diachronic changes in long-distance dependencies: the case of Dutch. In A. van Kemenade & N. de Haas (Eds.), *Historical Linguistics 2009: Selected papers from the 19th International Conference on Historical Linguistics, Nijmegen, 10-14 August 2009* (pp. 155 – 170). Amsterdam/Philadelphia: John Benjamins Publishing.
- Maling, J. & Zaenen, A. (1978). The Non-Universality of a Surface Filter. *Linguistic Inquiry*, 9, 475-497.
- Reis, M. (1995). Wer glaubst du hat recht? On so-called extractions from verb-second clauses and verb-first parenthetical constructions in German. *Sprache und Pragmatik*, 36, 27-83.
- Reis, M. (1996). Extractions from verb-second clauses in German? In U. Lutz & J. Pafel (Eds.), *On Extraction and Extraposition in German* (pp.45-88). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Reuland, E. (1983). *Movement Versus Merger: Relations between Inflection and Verb*. Paper presented at the 13th meeting of the Northeast Linguistics Society, Université du Québec à Montréal.
- Ross, J. (1967). *Constraints on variables in syntax*. Dissertation MIT, Cambridge, MA.
- Salzmann, M. (2006). *Resumptive Prolepsis. A Study in Indirect A' dependencies*. Dissertation, University of Utrecht, Utrecht.
- Schippers, A. & Hoeksema, J. (2010). Variation in long-distance dependencies. In F. Gregersen, J. Parrot & P. Quist (Eds.), *Language Variation – European perspectives III* (pp. 123-134). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Schippers, A (2012). Variation and change in Germanic long-distance dependencies. (Doctoral Dissertation). University of Groningen, Groningen. Groningen: Grodil.
- Schippers, A (2013). De dag die ik niet wist dat zou komen. MS. Url: sites.google.com/site/ankelienschippers/home/koningslied.
- Verhagen, A. (2006). On subjectivity and 'long distance Wh-movement'. In A. Athanasiadou, C. Canakis & B. Cornilli (Eds.), *Subjectification: Various Paths to Subject*