Hans-Jörg Schmid

Is usage more than usage after all? The case of English *not that*.[1]

Abstract: This paper combines the findings of a detailed corpus study and the results of a small-scale experiment in an attempt to support the idea that grammatical knowledge includes knowledge of usage patterns. The linguistic structures investigated are English *not that* sentences of the *type Other people have friends. I have enemies. Not that I care* (example taken from the British National Corpus). An in-depth investigation of 1,759 uses of this structure extracted from the British National Corpus reveals that neither their lexico-grammatical environments nor the pragmatic functions they serve are distributed in a random or unpredictable way, but rather pattern to an astonishing extent. This finding prompts the question as to whether the observed patterns are merely a matter of repeated usage driven by recurring pragmatic needs, or instead are represented as part of the linguistic knowledge of individual speakers, i.e., as a part of grammar. In order to answer this question a simple experiment was designed aiming to investigate whether there is a match between the data observed in the corpus (reflecting usage) and the knowledge-based behavior of individual speakers in an offline task (tapping into grammar). 30 native speakers of English participated in a test which put them under severe time pressure while they filled in gaps in short texts. Four gaps targeted uses of *not that* representing patterns with different frequencies in the corpus. The results of the experiment indicate that the participants were faster and found it easier to activate the structure *not that* for stimuli targeting frequent patterns than for those targeting very rare types of usage. This is interpreted as evidence for the assumption that frequency-related knowledge about lexico-grammatical and even pragmatic patterns is a part of grammatical knowledge.

Keywords: not that sentence, grammatical knowledge, usage-based models, corpus study, lexico-grammatical patterns, pragmatic patterns, entrenchment, conventionalization, Cloze test

Hans-Jörg Schmid: Institut für Englische Philologie, Ludwig-Maximilians-Universität München, Schellingstr. 3, D-80799 München, Germany. E-mail: hans-joerg.schmid@lmu.de.

1 Introduction

In recent years, the long-standing controversy between formalist and functionalist accounts of grammar has gradually been superseded by the divide between models that keep up the Saussurean distinction between knowledge of language and use of language, on the one hand, and usage-based models, which see grammar as being informed by and emerging from usage, on the other. In a paper entitled "Grammar is grammar and usage is usage", Newmeyer (2003) defended the former position that linguistic usage must not play a role in accounts of linguistic knowledge. In doing so, he makes it particularly clear that corpus-based frequency counts and statistical information drawn from them are irrelevant for grammar: "To sum up, probabilistic information drawn from corpora is of the utmost value for many aspects of linguistic inquiry. But it is all but useless for providing insights into the grammar of any individual speaker" (2003: 698). Casting serious doubt on this position, representatives of the usage-based camp – including, for example, Bybee (2006, 2007, 2010), Croft (2010), Dąbrowska (2010), Goldberg (2006), Hopper (1987) and Langacker (e.g., 1988, 2008) – have repeatedly drawn attention to effects of usage on represented grammatical knowledge. The latest showdown pitching the two factions against each other can be witnessed in the volume edited by Boye and Engberd-Petersen (2010).

My paper makes an attempt to provide empirical evidence supporting the usage-based model. It takes the generative position epitomized in Newmeyer's statement as a zero hypothesis, which assumes that there is no match between the frequencies of usage patterns in corpora reflecting conventionalized language use, on the one hand, and the entrenched linguistic knowledge that individual speakers have to activate while solving experimental tasks, on the other. As it will transpire, this hypothesis has to be rejected, while an alternative usage-based hypothesis predicting the existence of a link between usage patterns and entrenched knowledge finds support in the data presented in this study.

Meeting the demands of the two-pronged research question – addressing usage as well as grammar – the methodology used to test the two hypotheses relies on two types of data: usage data extracted from a corpus, and experimental data designed to tap into the mental representation of linguistic knowledge. The linguistic structures investigated are sentences including the sequence *not that* which instruct addressees to reject a possible inference of what was said before

(cf. Delahunty 2006; see Section 2 for more details). In the first methodological step, attestations of *not that* sentences are collected from the British National Corpus and analyzed with regard to frequencies of patterns of occurrence; this provides an account of conventionalized usage as represented in the corpus. Secondly, four patterns instantiated with different frequencies in the corpus, ranging from very frequent to very rare, are selected to serve as stimuli in a psycholinguistic experiment testing the linguistic knowledge of individual speakers. The rationale behind this methodology is that if no match between the corpus data and the experimental data is found, then Newmeyer's position that knowledge of usage patterns is not part of grammatical knowledge is confirmed. If, on the other hand, such a match can be observed and a correlation between usage patterns and grammar is shown, this casts serious doubt on the separation of usage and grammar and supports the alternative hypothesis proposed by usage-based models.

In Section 2, the linguistic structure under investigation is introduced. In Section 3, *not that* is described from a semantic, pragmatic, syntactic and historical perspective in order to lay the foundation for the detailed investigation of the lexico-grammatical and pragmatic usage patterns surrounding this structure, which are presented in Sections 5 and 6 respectively, after the corpus material has been introduced in Section 4. Section 7 builds on the results of the corpus study to formulate the rule-based zero-hypothesis and the usage-based alternative hypothesis. Section 8 describes the design and implementation of the small-scale Cloze test and presents its results. Section 9 concludes with a rejection of the rule-based hypothesis in favor of the alternative usage-based entrenchment hypothesis.

2 The linguistic structure under investigation

The subject of this investigation is a construction involving the sequence *not that* in uses of the type illustrated in (1) and (2), (both, as well as all other examples rendered in this paper, are taken from the *British National Corpus* unless indicated otherwise):[2]

---

2 It is for a variety of other reasons that the negator *not* can find itself placed before the complementizer *that* in an English sentence. Structures producing this possibility include negated *that-clauses* as in (a), *wh*-cleft sentences with negated *that*-clauses as in (b), negated nominal topic constructions (cf. c) and inverted negated conditionals (cf. d):

(a) *Indeed, looking back over the years, I have become surprised, **not that** people murder each other but that, given our love of bloodshed, they don't do it more often.* (H90 284)
(b) *What she found hardest to forgive was **not that** Luke had betrayed her. It was that he had made her betray herself.* (H8S 3989)

(1) *. . . no one will fancy him now.* **Not that** *I fancied him before but . . .*
    (KP3 1662)

(2) *At Silverstone, Hunt was the national hero.* **Not that** *he had only supporters –*
    for *there were many who thought he was, and behaved like, a spoiled brat –*
    but . . . .
    (CD9 1159)

While this construction has hardly been studied so far, it is remarkable that a
range of linguists with distinctly different interests have found themselves suffi-
ciently intrigued to comment on it: grammarians (Jespersen 1917; Huddlestone
and Pullum 2002) and lexicographers (*OED*, *Collins Cobuild*, *Longman*, see below)
as well as linguists with a pragmatic bent (Horn 2001; Delahunty 2006). In his
groundbreaking treatment of negation, Jespersen (1917) states the following:

> Inversely [to *not* functioning as an object as, e.g., in *I hope not* or *I suppose not*] we have a
> negative adverb standing for a whole main sentence, *not that* meaning 'I do not say that' or
> 'the reason is not that' as in Sh. Cæs. III. 2. 22. Not that I Lou'd Cæsar lesse, but that I lou'd
> Rome more [. . .]. (Jespersen 1917: 54)[3]

Huddleston and Pullum, in *The Cambridge Grammar of the English Language*
(2002), discuss *not* + *that*-clauses in a section on "*not* as a marker of nonverbal
negation" and describe their meaning as "this is not, however, to say/suggest that
. . .", adding that "the *not* calls up a proposition that might be naturally assumed
or expected in the context, and denies that it is in fact true" (Huddleston and
Pullum 2002: 811).

Somewhat less convincingly, the first edition of the *Collins COBUILD English
Grammar* (Sinclair 1990) states in the section on concessive clauses:

(c) *Their concern was* **not that** *others might come here one day and identify the place as having
    been the kidnappers' hideout.* (CAM 1641)
(d) *Ludens's own relations with Marcus might have been called, from Ludens's point of view,
    promising, even satisfactory, were it* **not that** *they were also precarious.* (APM 753)

None of these structures will be dealt with in this paper, however.
3 Ironically, the passage from *Julius Caesar* quoted by Jespersen is not an example of the kind
of *not that* of interest here, because it elliptically takes up the cataphoric pointer *this is my
answer.* Confer the preceding cotext: "If then that friend demand why Brutus rose against
Caesar, this is my answer: Not that I loved Caesar less, but that I loved Rome more". An
expanded version would therefore read something like 'my answer is not that I loved Caesar
less, but . . .', which does not match the denying function characteristic of the uses of *not that*
at issue here. For a controversial discussion of another use of *not that* in Shakespeare's work
(*Othello,* III.iv), see Horn (2001: 435) and Delahunty (2006: 221).

> You can use 'not that' instead of using 'although' and a negative. For example, instead of saying 'I have decided to leave, although no one will miss me', you can say 'I have decided to leave—not that anyone will miss me'. (Sinclair 1990: Section 8.68)

The online edition of the *Oxford English Dictionary* (Simpson n.d.) labels *not that* as being elliptical and provides the glosses "I do not say this because . . .", "[i]t is not the fact that . . ." and "[o]ne must not suppose that . . ." (s.v. *that*) in order to capture the semantic range of *not that*. Highlighting the usage conditions of *not that* rather than explaining its meaning, the fourth edition of the *Longman Dictionary of Contemporary English* gives the somewhat unsatisfying account that it is "used before a sentence or phrase to mean the opposite of what follows it, and to make the previous sentence seem less important" (Summers 2005, s.v. *not*).

In the only study exclusively devoted to *not that* of which I am aware Delahunty (2006) provides an account of the construction couched in the framework of relevance theory (Sperber and Wilson 1995). Studying 88 examples collected unsystematically from a variety of sources, Delahunty argues that the *not that* structure "encodes a procedure to the effect that the hearer is to reject the proposition represented by the S of the NTS [i.e., *not that* sentence, HJS] as a conclusion derived from local context" (2006: 238). Delahunty convincingly rejects Horn's (2001: 435–436) brief remarks that *not that* sentences are synonymous with inferential negatives and can be regarded as cases of metalinguistic negation. Significantly, Delahunty claims that generalizations which are more specific than the one rendered in the quote are neither necessary nor possible, since he was unable to detect "further structure" (2006: 239) in his limited dataset. Suggesting a phrase marker accounting for the syntactic structure of *not that* sentences (rendered in Section 3 below), Delahunty essentially claims that the interpretation of *not that* sentences can be accomplished by means of regular and compositional syntactic processing and by following the procedural instructions encapsulated in the quote above. In his account, then, *not that* is a matter of grammar and general pragmatic principles. To appreciate the validity of this view, I will take a closer look at the syntax, semantics and history of *not that*.

3 Some observations on the syntax, semantics and history of *not that*

In their treatment of *not that*, Huddleston and Pullum remark that

> [t]he syntactic analysis is somewhat problematic. In terms of function the construction occupies a nonembedded position, like a main clause. In terms of its structure, we might take

> *not* as modifying the content clause (as in *not all* it modifies *all*, and so on). If so, the whole construction will have the form of a subordinate clause even though it is not functionally subordinate; as with other cases of this kind (such as *That it should have come to this!*) there is implicit rather than explicit functional subordination." (2002: 811)

From a syntactic point of view, then, the question arises as to whether the somewhat exotic clause-initial combination of the negator and a *that*-clause can be traced back to an "underlying" regular phrase marker. From a semantic point of view, this raises the question to what extent the meaning(s) of this combination can be accounted for without taking recourse to noncompositional information.

An obvious candidate for an underlying structure motivating *not that* is suggested in the OED entry quoted above:

> *not that . . .* ← [*it is*] *not* [*the fact*] *that . . .*

Following the lead provided by the OED, *not that* could then be characterized as an elliptical negated focus construction, since the expanded version has the effect of topicalizing the negator *not* at the beginning of the clause and putting the proposition expressed in the *that*-clause into end focus (cf. Delahunty 2001). Example (1), repeated here as (3) for the sake of convenience, could thus be treated as an elliptical version of (4):

(3) *. . . no one will fancy him now.* **Not that** *I fancied him before but . . .*

(4) *. . .* no one will fancy him now. It is **not** the fact/case **that** I fancied him before
    . . .

If grammatical description is to be parsimonious, and if, following Occam's razor, "[*s*]*enses are not to be multiplied beyond necessity*" (Grice 1989: 47; [emphasis original]), this would have an appealing advantage: clause-initial, "plain" *not that* and the expanded copula structures *it is not that, it's not that, it isn't that* and *it wasn't that* (cf. Delahunty 2001) could all be traced back to the same regular phrase structure. However, for a number of reasons, this account is not satisfying. Two strong arguments are put forward by Delahunty (2006: 224; contra Horn 2001: 435–436). Firstly, sentence-initial *not that* and the copula structures – referred to as *negative inferentials* by Horn (2001: 435) – are not, at least not always, mutually substitutable. Only about half of the *not that* examples collected by Delahunty allowed for a substitution by *it is not that*, the other half resisted it. This is confirmed by my data. For example, a substitution of (3) above by the

copula version in (5) neither preserves the communicative impact of (3), nor does it produce a coherent piece of discourse:[4]

(5) . . . no one will fancy him now. It's not that I fancied him before . . .

Secondly, Delahunty (2006: 225) shows that plain *not that* sentences behave in a fundamentally different way from *it is not that* sentences with regard to possible continuations with positive inferentials such as *it is just that, but that* or simply *that*, which frequently follow *it is not that*. Again, this is also confirmed by my data: the proportion of examples like (6), where a *not that* utterance is followed by *it's just that*, is very low (around two percent):

(6) *Caroline declined the coffee but accepted a magazine.* **Not that** *she read it; it was just easier to bury her face in its glossy pages than to have to endure Nicolo's stony-faced silence.*
(JY7 1919)

These observations seem to indicate that the plain *not that* type and the copula type *it BE not that* are not variants of one form from a semantic and pragmatic point of view.

    The supposition that plain *not that* is not an elliptical variant of *it is not that* is further corroborated by historical evidence. As shown in greater detail in Schmid (2011), the attested examples of *not that* found in the quotations to all entries in the *OED* indicate that plain *not that* predates *it is not that* by more than two centuries. The first quotation of plain *not that* to be found in the *OED* is a translation of Lat. *non quod*[5] in the Wycliffe New Testament quoted in the entry on *sufficience*:[6]

---

4 The copula versions of the type *it's not that* have the additional property, not shared by plain *not that*, of creating a focus construction highlighting the negator which can often felicitously be paraphrased by *the point is not that*. Cf. the example . . . *in the country where Great-Granny comes from they don't eat some of the things we eat.* **It's not that they**'**re fussy***, it's because they think it's wrong* (FRH 1343), which can be rendered as *the point is not that they're fussy*.
5 The *vulgata* version reads: *non quod sufficientes simus cogitare aliquid a nobis quasi ex nobis sed sufficientia nostra ex Deo est.* The Greek bible has *ουχ οτι* literally also 'not that, because, since'.
6 Interestingly, the earliest attested example provided in the sub-entry on *not that* (s.v. *that*, II.2.b) dates from 1601. That the quotation from the Wycliffe New Testament may indeed be the first attestation of *not that* in the function investigated here is corroborated by an analysis of the related Old English forms *ne þæt* and *na þæt* in the Toronto corpus of Old English texts. None of the 280 occurrences of these sequences has meanings comparable to modern *not that*.

1382 WYCLIF *2 Cor.* iii. 5 Not that we ben sufficient for to thenke ony thing of vs, as of vs, but oure sufficience is of God.
(*OED*, s.v. *sufficience*)[7]

However, the first attestation of *it is not that* that I know of occurs in Shakespeare's *As you like it* (1599), Act III, Scene 2, where Phebe generously reacts to Silvius' confession of his love for her:
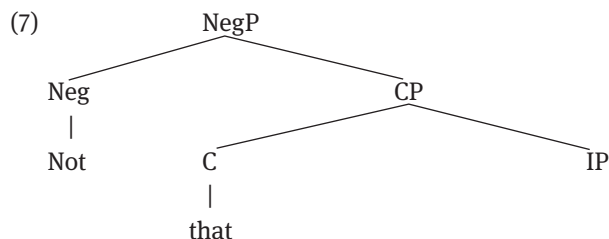
Silvius, the time was that I hated thee;
And yet it is not that I bear thee love;
But since that thou canst talk of love so well,
Thy company, which erst was irksome to me,
I will endure; and I'll employ thee too.

It is worth mentioning in passing that this quotation comes from a period where focus constructions depending on abstract nominals of the "shell noun" type (cf. Schmid 2000) such as *the idea is that, the reason is that* and *the point is that* begin to occur in larger numbers (cf. Mantlik 2011). In short, what the historical and comparative evidence collected in Schmid (2011) suggests is that plain *not that* has entered English – and many other European languages as well as other languages affected by Bible translations – by means of loan syntax from Latin, while copula *not that* comes much later and may have to do with the emergence of nominal focus constructions.

On the whole, then, there is converging evidence from syntax, semantics and language history indicating that plain *not that* is not an elliptical form of *it is not that*. However, this does not automatically mean that plain *not that* is noncompositional and therefore has to find its place in the famous dumping ground for idiosyncrasies which is the lexicon. A simple way of saving *not that* from this fate would be to come up with a regular phrase marker tailor-made for *not that*, as suggested by Delahunty (2006: 216). His account is rendered in (7):

What is also remarkable is that *not that* occurs in marginal glosses in the Old Testament in the Wycliffe Bible with the aim of rejecting possible misinterpretation of passages in the scripture (see Schmid 2011).
7　The passage preceding this sentence in the Wycliffe New Testament reads: "and made open [and *ye be* made open], for ye be the epistle of Christ ministered of us, and written, not with ink, but by the Spirit of the living God; not in stone tables [not in stony tables], but in fleshly tables of heart. For we have such trust by Christ to God; Not that . . ." (quoted from www.biblegateway.com).

(7)

```
                    NegP
           ┌─────────┴─────────┐
         Neg                   CP
          │            ┌───────┴───────┐
         Not           C               IP
                       │
                      that
```

On the semantic side, this option would be accompanied by a compositional account analyzing the meaning of plain *not that* simply as "negation of proposition encoded in the *that*-clause." As already observed in Section 2, the conventional semantico-pragmatic properties of *not that* could be accounted for by the procedural instructions proposed by Delahunty (2006). This reasoning would allow for a very economical description of *not that* as an essentially rule-based and compositional structure whose interpretation is guided by general pragmatic principles giving rise to conventional interpretations in local contexts. This may not do justice to *not that*, however, as I will try to show in the following sections.

4 Survey of the corpus data

The material for this investigation was retrieved from the British National Corpus, which contains roughly 100 million words. About 10 percent of the material are transcriptions of originally spoken utterances, with approximately 4.5 million words recorded in everyday casual situations (representing the so-called *demographically sampled* part) and 5.5 million words garnered from more institutionalized settings such as meetings, debates, etc. (*context-governed* part). The written material, accounting for 90 percent of the corpus, "includes, for example, extracts from regional and national newspapers, specialist periodicals and journals for all ages and interests, academic books and popular fiction, published and unpublished letters and memoranda, school and university essays, among many other kinds of text" (BNC website, http://www.natcorp.ox.ac.uk/corpus/index.xml, last accessed 5 July 2011).

    The corpus was accessed using the BNC*web* version and client (cf. Hoffmann et al. 2008). The search string "not_XX0 that_CJT" (which translates as '*not* POS-tagged as a negator directly followed by *that* POS-tagged as subordinating conjunction *that*') yielded a recall of 1,991 instances of *not that*; the string "n't_XX0 that_CJT" recalled 110 results. Due to homonymic structures of the types mentioned in the introduction above (cf. Note 2), manual post-processing was necessary to eliminate corpus hits that were clearly not instances of the *not that* type

under investigation here. 1,687 valid hits of *not that* and 72 valid hits of the contracted form *n't that* survived this selection process, which together (n = 1,759) make up a fairly low relative frequency of 17.89 occurrences per million words.[8] 379 of these uses were instances of the "expanded" type discussed in Section 3, where *not that* occurs as a complement in copula sentences preceded by *it is* (101 instances), *it's* (155 instances) or *it was* (51 instances). An additional 63 instances are of the type *it wasn't that*, and 9 more of the type *it isn't that*.[9] The examples of these expanded types were not included in the further analysis.

5 Lexico-grammatical patterns in grammatical slots in *not that* clauses

In this section, patterns in the actual usage of *not that* on the linguistic levels of lexico-grammar and pragmatics will be identified and discussed. The notion of *pattern* embraced for this undertaking is "pre-theoretical": a pattern is simply defined as a recurrent sequence of linguistic elements or moves on any level of linguistic description which serves an identifiable function. While I do not intend

---

8 This score puts *not that* behind the superficially comparable two-word sequences *in addition* (80 occurrences per million words in the BNC)*, after all* (63) and *above all* (23), but ahead of *even so* (14) and the famous *let alone* (13; cf. Fillmore, Kay and O'Connor 1988).

9 Examples of these patterns are:

(i) *I have never had a serious girlfriend. I just think it is odd at 17 to spend all your time with one person. I would have to be in love with someone to do that.* **It is not that** *I haven't got time for girls. I haven't got time to meet girls. There is a big difference.* (CH1 5654)

(ii) *Fielding it with one hand, she sobered, and, putting her glass down on the fender, she hugged the cushion on her knees.* '**It's not that** *I'm reluctant,' she denied slowly, 'I just find it so incredible that you want to marry me.'* (H9V 3048)

(iii) *Thomas Huxley, the great biologist, whose household was dominated by a long series of cats over a period of forty years, described how one of them, a young tabby tom-cat, developed the alarming game of jumping on the shoulders of his dinner-guests and refusing to dismount until they fed him some titbit.* **It was not that** *the animal was hungry. It was the shock impact of the game that provided the reward.* (BMG 946)

(iv) *The knowledge caused her genuine distress and, in the face of Constance's increasing truculence, she turned to Louise for counsel.* '**It isn't that** *Constance is naughty, or has tantrums,' she said to her cousin on the telephone, 'it's as if she has just turned her back on me.'* (CEY 751)

(v) *The holiday was not a total success.* **It wasn't that** *Katherine wasn't enamoured of Paris.* (FNT 3418)

to go into a discussion of what is to count as being *recurrent*,[10] three aspects of this definition should be highlighted: firstly, the definition allows patterns to be observed both on the "surface" level of lexico-grammatical *elements* and on the "deeper" level of pragmatic and conversational *moves*. Secondly, it stipulates that patterns must not be the result of chance juxtapositions of elements, but emerge because their components serve a common purpose. To take a random example, although we find that the sequence of the words *paper* and *so* occurs 29 times in the British National Corpus, this still will not count as a pattern, as there is no identifiable function behind the recurrence of this sequence. Admittedly, this view of patterns is subjective, and it must be acknowledged that a whole range of statistical methods exist for determining the significance of a recurrent sequence and the strength of the attraction of the component elements (cf., e.g., Church and Hanks 1990; Clear 1993; Stubbs 1995; Evert and Krenn 2001; Manning and Schütze 2001). But as these are not always applicable, especially not on the level of pragmatics, they will only be invoked where appropriate. And thirdly, it should be emphasized that this notion of pattern has no theoretical implications at this stage, but merely reflects empirical observations.

The following discussion will focus on uses of *not that* in the two most prominent categories, "spoken" (169 occurrences) and "fiction and verse" (755 occurrences). Since spoken usage is deemed to be even more dissociated from represented grammatical knowledge than written usage, this seems to be justifiable on methodological grounds, since it would mean that matches between usage tendencies and grammatical knowledge would be expected to a lower degree for spoken than for written language. I will first look at the distribution of the elements in subject position in *not that* clauses, and then at verbs in *not that* clauses. The question throughout is whether it is possible to identify recurrent patterns.

5.1 Subjects of *not that* clauses

Table 1 lists the results of an R1 collocation analysis – i.e., an analysis of the items immediately following *not that* – in the 169 occurrences of *not that* in the spoken

---

10 Frequency thresholds have been a matter of dispute in corpus-based collocation studies. What is uncontroversial is that they depend on the size of the corpus investigated. Kjellmer (1982: 26), for example, uses a threshold of "more than once" in the one-million-word Brown Corpus, and Clear (1993: 277) three occurrences in a corpus of 25 million words. For Jones and Sinclair, collocations are recurrent when they occur together "more often than their respective frequencies and the length of text in which they appear would predict" (1974: 19). See also Stubbs (1995) on this question.

section of the BNC. This is a reasonable approximation to a search for the first item of the phrase realizing the subject constituent in the *not that* clause. Since the frequencies of occurrence of the elements found in this slot depend, as always, on their overall frequencies in the corpus, the table not only provides information on the observed frequencies of all items, but includes their total frequencies in the spoken texts, plus a calculation of the expected frequencies and the log-likelihood statistics. Loosely speaking, this is a measure of the strength of the collocational association between *not that* and the respective item, based on a comparison of expected and observed single and combined frequencies.

Item	Observed collocate frequency	Percentage of all tokens in slot R1	Total no. in spoken texts	Expected collocate frequency	Log-likelihood value
I	88	52.07%	309,557	4.366	413.7451
you	19	11.24%	268,463	3.786	32.3164
we	12	7.10%	108,427	1.529	29.1726
he	9	5.33%	75,396	1.063	22.9535
they	8	4.73%	96,562	1.362	15.3203
both	2	1.18%	3,093	0.044	11.4127
there	6	3.55%	73,008	1.030	11.3587
she	4	2.37%	42,873	0.605	8.3937
it	9	5.33%	253,864	3.580	5.9323
their	1	0.59%	13,232	0.187	1.7346
any	1	0.59%	15,929	0.225	1.4393
people	1	0.59%	21,295	0.300	1.0093
that	4	2.37%	227,029	3.202	0.1881
this	1	0.59%	58,179	0.821	0.0369
[unclear]	1	0.59%	152,636	2.153	−0.7798
the	2	1.18%	409,714	5.778	−3.3994
,	1	0.59%	583,403	8.228	−10.5608

Table 1: Results of an R1 collocation analysis of *not that* in the spoken corpus section

What is particularly remarkable about this list is the massive proportion of occurrences of *I*, which fills the subject slot in more than 50 percent of the 169 instances. As the difference between the expected frequency of occurrence of 4.3 and the observed one of 88 is huge, the log-likelihood statistics are highly significant. To confirm that there is indeed something special about this distribution, occurrences of *I* following *that* in **all** *that*-clauses in the spoken section of the BNC were retrieved as a benchmark. Here the proportion of first person singular *I* amounts to only a little more than 7% (4,761 occurrences in a total of 65,072 *that*-clauses).

The large share of *I* as subject in the *not that* clause is thus clearly specific and significant. Mainly as a consequence of the large share of *I*, the proportion of subjects realized by personal pronouns amounts to a staggering 88.76 percent. With regard to the subject position, then, *not that* clauses make up an almost closed system in the spoken subcorpus.[11]

A similar, though somewhat less extreme, situation is found in the material from the subcorpus "fiction and verse", where roughly 75% of the subjects in plain *not that* clauses are realized by personal pronouns (cf. Table 2). This contrasts with a mere 43% in all *that*-clauses in the section "fiction and verse."

Item	*not that* total = 755 occurrences	all uses of *that* total = 88,605 occurrences
I	25.30%	6.73%
she	16.16%	10.40%
it	12.72%	5.18%
he	12.58%	11.36%
they	3.05%	3.24%
you	3.05%	3.97%
we	2.12%	1.63%
total personal pronouns	74.98%	42.51%

Table 2: Results of an R1 collocation analysis of plain *not that* in the corpus section "fiction and verse"

While the first person singular pronoun *I* is again substantially more frequent following *not that* than after *that* alone, the difference is not as pronounced as in the spoken subcorpus. Other remarkable contrasts to be gleaned from Table 2 are the comparatively large proportions of occurrences of *she* and *it* in *not that* cotexts (i.e., linguistic contexts). As will be shown later, the high proportion of occurrences of *it* is mainly caused by cotexts of the type *not that it mattered*, *not that it's any of your business* and *not that it makes any difference*. As far as the striking number of *she* subjects is concerned, many of these are authorial comments on the troubles, feelings and motives of female protagonists (cf., e.g., *She recognized him immediately, but though his gaze passed over her there was no recognition. Not that she expected there would be . . .* , [C85 1472]).

---

[11] The proportion of all personal pronouns in the slot R1 following *that* in the whole spoken corpus amounts to 43%.

5.2 Verbs in *not that* clauses

To construct an automatic, and thus objective and replicable, search of verbal elements in *not that* clauses, the BNC*web* option for the retrieval of collocations was set at a span of 2 right to 4 right and the collocate search was restricted to verbs. This means that the corpus processing software looks for all verbs following *not that* within a distance of 2 to 4 items and checks their observed frequency in the collocation against the expected frequency calculated from individual and combined occurrence. To exclude rare collocates, the minimum number was set to 3. As one would not expect that the system of possibilities for the verbs is as closed as the one observed for the subject slots, a different way of assessing the results was chosen, one that exploits the dissimilar properties of two types of statistics indicating strengths of associations: while the log-likelihood measure introduced above is known to balance the overall frequency of occurrence of an item in the corpus and the frequency of occurrence in a given collocation, a second measure, called *mutual information*, by virtue of the way it is calculated, favors those elements that specialize so to speak in a given combination, even if they are fairly infrequent in the overall material (cf. Church and Hanks 1990: 23–24; Evert and Krenn 2001; Manning and Schütze 2001: 166–171). Typically, rank lists of collocates ordered in terms of the two measures show very little or no overlap. This is demonstrated by the two lists (cf. Table 3) of the top 15 collocates of all *that*-

Rank	According to log-likelihood	According to mutual information
1	*was*	*forfeiting*
2	*had*	*denoted*
3	*would*	*disinherited*
4	*could*	*presaged*
5	*been*	*encompassed*
6	*be*	*girdled*
7	*should*	*precedes*
8	*might*	*characterized*
9	*were*	*haunts*
10	*have*	*befell*
11	*'d*	*erred*
12	*did*	*interwoven*
13	*must*	*radiate*
14	*wanted*	*radiated*
15	*is*	*emanated*

Table 3: Log-likelihood and mutual information rank lists of verbal collocates in the span 2R to 4R following *that* in the subcorpus "fiction and verse"

clauses in the subcorpus "fiction and verse" ranked in terms of log-likelihood, on the one hand, and mutual information, on the other. As can be seen, high-frequency auxiliary and modal verbs prevail in the left-hand column, and fairly rare lexical verbs in the right-hand one; not a single form occurs in both rank lists.

The situation is entirely different in the case of *not that* sentences. Carrying on with the subcorpus "fiction and verse" first, Table 4 provides the same type of juxtaposition based on identical corpus processing commands. For *not that*, no fewer than 9 of the 20 items listed can be found in both rank lists. These are highlighted in boldface print.

Rank	According tolog-likelihood	According to mutual information
1	***mattered***	***mattered***
2	*would*	***minded***
3	*'d*	***matters***
4	***cared***	***cared***
5	*was*	*complaining*
6	*had*	*deserve*
7	***minded***	***bothered***
8	***matters***	*blame*
9	*know*	*recall*
10	*'m*	***believed***
11	*have*	***care***
12	*'ve*	*bother*
13	***wanted***	*intended*
14	***needed***	***needed***
15	*'s*	*makes*
16	***care***	*expected*
17	*seen*	*noticed*
18	***believed***	*expect*
19	***bothered***	***wanted***
20	*could*	*worry*

Table 4: Log-likelihood and mutual information rank lists of verbal collocates in the span 2R to 4R following *not that* in the subcorpus "fiction and verse"

For the data collected from the spoken section of the BNC, the match between the two rank lists is even closer, since all 20 items occur in both lists (cf. Table 5). Admittedly, this is partly made possible by the relatively small number of overall items (169), but it is still highly significant for our assessment of patterns in the right-hand side cotext of *not that* clauses.

Rank	According to log-likelihood	According to mutual information
1	've	recall
2	'm	noticed
3	can	wish
4	recall	remember
5	noticed	'm
6	remember	've
7	know	'd
8	wish	want
9	'd	can
10	want	used
11	's	come
12	're	know
13	come	're
14	used	could
15	would	would
16	think	think
17	could	had
18	be	's
19	have	be
20	had	have

Table 5: Log-likelihood and mutual information rank lists of verbal collocates in the span 2R to 4R following *not that* in the subcorpus "spoken"

What has been shown so far, then, is that in both the subject and the verb slots, the clauses introduced by plain *not that* exhibit a remarkable degree of regularity in the corpus sections investigated. The next section will look more closely at specific lexically filled patterns observable in the corpus data.

5.3 Semifixed lexical patterns in *not that* clauses

The patterns presented here are arranged in groups sharing similar meanings. Each group will be referred to by a label reflecting a prototypical wording of the pattern.

5.3.1 NOT THAT I KNOW OF

This is a very specific pattern in several respects. For one thing, it is typically found at or towards the beginning of an answer to a question (rather than as a

continuation of a turn by one speaker). Secondly, the semantic and pragmatic impact of this pattern seems to differ substantially from the *not that* uses in the examples quoted so far, as it does not lend itself to glosses of the type 'this is not to say/suggest that'. On the contrary, superficially, *not that I know of* seems to correspond fairly well to *I don't know* (but see the more detailed discussion in Section 6.2.2.2 below). A typical example of this pattern is provided in (8):

(8) *Did he have a gun? – Not that I know of.*
     (GW3 2871)

Given its place in conversational structures, this pattern is more or less restricted to dialogue (in fiction and everyday conversations). Frequencies of occurrence of several variants of this pattern are shown in Table 6:

	Spoken	Fiction and verse	Other written sources	Total
not that I know of	8	12		20
not that I know about	1			1
not that I know (not followed by *of*)	3	8	2	13
not that I knew	1	1	1	3
not that I recollect/recall/remember	2	3		5
not that I can remember	3	2	1	6
not that I can recall	3	2		5
not that I can think of	2	1		3
not that I could see	3		1	4
not that I've [*ever*] *seen*	1	3		4
not that I ['*ve*] *noticed*	1	5	1	7
not that I think	1	5	3	9
not that I thought		2	2	4
not that I'm/am aware of	2	2	1	4
not that you'd know		3		3
not that I've [*ever*] *come across*	2			2
Σ	33	49	12	93

Table 6: Occurrences of the pattern NOT THAT I KNOW OF in various sections of the BNC

5.3.2 NOT THAT IT MATTERED

Loosely speaking, the pattern dubbed NOT THAT IT MATTERED has the function of expressing that propositions uttered in the previous cotext are not important. The pattern is illustrated in (9):

(9) *She simply couldn't keep house – **not that** that mattered so much – but she
    couldn't cope at all with the children. It wasn't that she didn't love them – or
    want to – she just couldn't.*
    (HH9 937)

As Table 7 indicates, this pattern is very rare in the spoken section of the corpus
and comparatively frequent in fiction and verse

	Spoken	Fiction and verse	Other written sources	Total
not that it matters	1	6	2	9
not that that matters		2		2
not that it much matters			1	1
not that it mattered		9	2	11
not that it really mattered		3		3
not that that mattered		2	1	3
not that it makes any difference		3		3
not that it would have made any difference	1	1		2
not that it would make much difference		1	1	2
not that it's any of your (damn(ed)) business		4		4
not that there was any (need, harm, danger) etc.		7	1	8
Σ	2	38	8	48

Table 7: Occurrences of the pattern NOT THAT IT MATTERED in various sections of the BNC

5.3.3 NOT THAT I CARE

Apparently very similar to the previous pattern, the pattern NOT THAT I CARE high-
lights a somewhat indifferent reaction on the part of the speaker or a third-person
referent to the propositional content of a preceding utterance (cf. 10).

(10) *Other people have friends. I have enemies. **Not that** I care.*
    (ALH 2640)

On the lexico-grammatical level, this difference is reflected in the choice of
subjects: whereas *it* is a marker of uses of the type NOT THAT IT MATTERS, the
uses subsumed under NOT THAT I CARE have animate first-person or third-person
subjects.

	Spoken	Fiction and verse	Other written sources	Total
not that I care	0	4	2	8
not that X *cared*	0	13	3	16
Σ	0	17	5	22

Table 8: Occurrences of the pattern NOT THAT I CARE in various sections of the BNC

As indicated in Table (8), no instances of this type were found in the spoken section.

5.3.4 NOT THAT I'M AGAINST

The fourth pattern subsumes uses of *not that* denying the assumption that the referent of the *not that* clause objects to a previously mentioned proposition. With regard to its distribution, the pattern is similar to the previous one (cf. Table 9), although a small number of uses were identified in the spoken subcorpus. A typical example is given in (11):

(11) *Ian Fleming's superhero first appeared in Issue 4 of ZZAP!, much to the dismay of the reviewing team.* **Not that** *they had anything against the original character or film; far from it.*
(C87 620)

	Spoken	Fiction and verse	Other written sources	Total
not that I'm against		4	1	5
not that X *have/had anything against*		7	2	9
not that X [*have,'ve,has,'s*] *got anything against*	1	2	3	6
not that X *had minded*		2		2
not that X *mind/minds/minded*	2	10	3	15
Σ	3	25	9	37

Table 9: Occurrences of the pattern NOT THAT I'M AGAINST in various sections of the BNC

5.3.5 NOT THAT I'M COMPLAINING

Manifestations of this pattern deny intentions which could possibly be attributed to the speakers on the basis of what they have said before, as in (12):

(12) *But it is a strange life when one sees one's friends for perhaps one month in the year, and then we all go our separate ways until we can grasp another opportunity – sometimes between flights at airports!* **Not that** *I'm really complaining! I'm lucky enough to have a diary full of opera engagements, . . .* (BMC 298)

	Spoken	Fiction and verse	Other written sources	Total
not that X *would/'d want*	2	3	3	8
not that X [*is,'s,am,'m,are,'re*] [*really*] *complaining*		2	2	4
not that *I'm saying*	1	1		2
not that *I'm suggesting*		1	2	3
not that X [*is,'s,am,'m,are,'re*] *blaming*		1	1	2
not that X *'d blame*		4		4
Σ	3	12	8	23

Table 10: Occurrences of the pattern NOT THAT I'M COMPLAINING in various sections of the BNC

Taken together, the lexico-grammatical patterns listed in this section account for a considerable proportion of the data from the BNC. The semifixed lexical patterns alone add up to 223 occurrences, which corresponds to more than 16% of the material. In the spoken corpus section, the instances listed in the tables account for a quarter of the full material, and in the fiction and verse section for 18%. While these proportions do not appear to be huge, one has to bear in mind that the recurrence of patterns of identical lexical items decreases dramatically with increasing length of strings. From this perspective, the amount of patterning observed is clearly impressive and renders it extremely unlikely that the observed regularities are due to chance.

What is more, these automatically identifiable lexico-grammatical and lexical patterns are just the tip of an iceberg of pragmatic regularities to be found in the corpus data. As will be shown in the next section, these regularities go beyond the pragmatic descriptions quoted in Section 1 above – cf. Jespersen's *I do not say that*, Huddleston and Pullum's *this is not, however, to say/suggest that . . .* ,

the *OED*'s *one must not suppose that* . . . and Delahunty's relevance-theoretical description.

6 Pragmatic patterns

6.1 General characterization of the pragmatics of *not that*

As argued by Delahunty, *not that* sentences instruct the interpreter "to reject the proposition represented by the S [sentence] of the NTS [*not that* sentence] as a conclusion derived from local context" (2006: 238). Delahunty rightly emphasizes that the assumptions to be rejected do not have to "exist in the mind of the interpreter prior to the processing" of the *not that* sentence (2006: 238). While he makes this point with reference to the rare cases where the *not that* clause precedes rather than follows the main clause – as in his fabricated example *Not that I want to embarrass you, but you have egg on your tie* – it can be claimed that this is the rule rather than an exception. More often than not, *not that* sentences are deployed by the speaker to ask the hearer to reject an assumption which the latter would probably never have formed if it had not been for the *not that* sentence.

It is thus the speaker who presents, or indeed even stages, the information provided in the *not that* sentence as a potential inference[12] of what was said before. This becomes clear when we compare the *not that* construction in (13) to a paraphrase not using *not that* (cf. 14):

(13) *It was a hopeless war from the start, in spite of the gallantry of the Finnish Army, and most of the fighter planes available were outdated.* **Not that** *the Russians were much better, but they did have a few of the new German FW190S which Hitler had promised to Stalin as a goodwill gesture over the Poland deal.*
(HTW 2553)

(14) It was a hopeless war from the start, in spite of the gallantry of the Finnish Army, and most of the fighter planes available were outdated. The Russians were **not** much better, but . . .

---

12 Note that the conveniently neutral term *inference* is used here deliberately to steer clear of lengthy and basically unnecessary discussions of the kinds of implicatures or presuppositions involved in each case.

While it is of course still possible to establish coherence for (14), this version is certainly less cohesive, and the denying aspect is much less noticeable here than in (13). (13) has the effect of suggesting that the idea that "the Russians were much better" is a natural inference of the preceding statement. An additional effect highlighted by the juxtaposition of (13) and (14) is a distinct parenthetical element in the *not that* version in (13), which presents the information as some kind of afterthought or additional spontaneous reflection.

In short, from a general pragmatic perspective, uses of *not that* can be described as staged denials of inferences that could be made on the basis of the previous cotext, which are, or are intended to come across as, seemingly spontaneous afterthoughts. With this general characterization in place, we can now move to pragmatic patterns to be observed in the data. Since pragmatic patterns cannot be identified automatically in a corpus the way lexico-grammatical patterns can but must be isolated via time-consuming manual interpretations taking into account the cotext and context of each individual occurrence, the size of the dataset to be investigated had to be reduced. As it is in spoken language that we would expect to find the most "natural" and spontaneous uses of *not that*, the following discussion of the patterns will focus on this section of the corpus.

6.2 Types of pragmatic patterns

The recurrent pragmatic constellations in the data can best be described using the following three pragmatic parameters (which will become clearer in the following discussion):

–   The place of the *not that* clause in conversational structure: continuation of same speaker's turn vs. turn-initial second-pair part ("answer").
–   The type of perspective resulting from the choice of referring item functioning as topic in the *not that* clause: participant role (first person vs. second person) vs. third person reference.
–   Aspects of utterance denied by the *not that* clause: propositional-content condition, preparatory condition, sincerity condition, reason for utterance.

6.2.1 Conversational structure and perspective

With regard to conversational structures and perspectives, the data show a very strong tendency to pattern in characteristic ways. These can be gleaned from

Table 11. In the large majority of occurrences in the spoken corpus (roughly 73%),[13] *not that* is used as a part of one speaker's turn, usually by way of the kind of afterthought or parenthesis described above. In three quarters of the corpus examples (75%) the participant roles *I, we* and *you* function as topics and subjects in the *not that* clause; almost two-thirds (63%) are first-person references (cf. Section 5.1). The proportion of first-person references is particularly large (78%) in those cases where *not that* is part of an answer. The perspective of, and on, the speaker thus turns out to be very dominant.

What is highly remarkable about the third-person references is that there is not a single case of reference to concrete objects. While the majority of the 42 corpus examples of this type refer to persons by means of *she, he* and *they* (23), the rest are either anaphoric uses of *it* with extended reference or reference to fact (Halliday and Hasan 1976: 52–53), i.e., abstract references encapsulating preceding propositions, or existential *there*-constructions.

	First person reference	Second person reference	Third person reference	Total
second-pair part ("answer") by new speaker	32 78% of row total 34% of column total	4 10% of row total 29% of column total	5 12% of row total 12% of column total	41 27%
continuation by same speaker	63 57% of row total 66% of column total	10 9% of row total 71% of column total	37 34% of row total 88% of column total	110 73%
total	95 63%	14 9%	42 28%	151

Table 11: Survey of distribution of occurrences of *not that* with regard to selected pragmatic parameters

Examples of the patterns with the largest frequencies of occurrence emerging from this table are given in (15) to (17). Not surprisingly, some of these most frequent pragmatic patterns correspond to the semifixed lexico-grammatical patterns discussed in Section 6 above:

---

13 Note that 16 of the 169 uses found in the spoken corpus section had to be excluded from the analysis because their structures were so fragmented that a sound pragmatic interpretation was not feasible.

- Continuation by same speaker with first person reference (63 occurrences):

(15) *. . . you're only working a few days but it's surprising the teaching days I do [pause] really, that really does cut into my time at home. [pause]* **Not that** *I mind but it just means . . .*
(KBG 1873)
[lexico-grammatical pattern NOT THAT I HAVE ANYTHING AGAINST]

- Continuation by same speaker with third person reference (42 occurrences):

(16) *Ian Fleming's superhero first appeared in Issue 4 of ZZAP!, much to the dismay of the reviewing team.* **Not that** *they had anything against the original character or film; far from it.*
(C87 620)
[= Example 11 above, lexico-grammatical pattern NOT THAT I HAVE ANYTHING AGAINST]

- Answer by new speaker with first person reference (32 occurrences):

(17) *Did he have a gun? –* **Not that I know of.**
(GW3 2871)
[= Example 8 above, lexico-grammatical pattern NOT THAT I KNOW OF]

6.2.2 Patterns of denial

A more detailed view of pragmatic patterns emerges from an analysis of the *not that* clauses with regard to their relations to the preceding context. As has been shown, all *not that* clauses in my data constitute reactive speech acts which deny inferences that could potentially be made on the basis of preceding, or, very rarely, following, utterances. This basically confirms Delahunty's (2006) view.[14] As is typical of relevance-theoretical approaches, Delahunty argues that there is neither a need for nor a gain in trying to produce a more fine-grained description of possible types of inferences. Being interested in a description of systematic patterns of linguistic behavior, I disagree with Delahunty here and will try to show that certain types of inferences are much more common than others. My account will be based on Searle's (1975) generalizations capturing the inferential mechanisms involved in the interpretation of indirect speech acts. I argue that

---

14 In the framework proposed by Geurts (1998), which pursues different aims, *not that* achieves what he refers to as "implicature denials" (1998: 294–299). The only exception is example 25 below, which would count as a denial of presupposition.

speakers using *not that* instruct the hearer to deny inferences which result from very similar ways of reasoning. The major difference to Searle's approach is that his only concern are inferences which trigger or explain indirect *illocutionary acts* (e.g. requesting by means of asserting), whereas the users of *not that* exploit a wider range of potential inferential patterns. They not only deny inferred illocutionary acts but also other types of inferences potentially activated by a previous utterance. What indirect speech acts and *not that* denials have in common, however, are the preconditions for speech acts which can be exploited for inferences and are described by Searle (1975: 72, 81–82) in terms of generalizations: the preparatory condition, the propositional-content condition, the sincerity condition, and the "reasons for doing the act" (Searle 1975: 71–72, *et passim*). In indirect speech acts, these preconditions for the felicitous performance of speech acts are exploited as triggers for inferences, in uses of *not that* they serve as potential sources of inferences that are denied. These considerations yield the following set of four patterns of relations between the *not that* clauses and the preceding cotext:

– not *that* denying inferable propositional content
– not *that* denying aspects inferable from the preparatory condition
– *not* that denying aspects inferable from the sincerity condition
– *not that* denying inferences concerning the reason for making a previous utterance.

These four types of mechanisms can serve to explain pragmatic patterns underlying all uses of *not that* found in the spoken corpus section. However, as *not that* utterances are assertives/representatives, the category "reason for doing the act" does not refer to a future act outside the speech act itself (as is the case for the directives and commissives discussed by Searle 1975), but concerns reasons for, and the intentions in, making the earlier utterance, from which inferences are rejected. The patterns will be discussed in the following subsections. It should be noted that some examples resist an unambiguous classification.

6.2.2.1 *Not that* denying inferable propositional content

With an overall number of 85 instances, more than 55% of the data are instances where *not that* denies inferences that are related to the propositional content of the preceding utterance. This pattern has already been illustrated in examples (15) and (16). In (15), the possible sequence of inferential steps is: 'if something cuts into time spent at home, this is an unpleasant situation' > 'we normally mind unpleasant situations' > 'I deny that I mind'. In (16) the *not that* clause (*not that they had anything against*) rejects a possible inference triggered by the phrase *to*

the dismay of the reviewing team. In each of these two examples, and indeed in the large majority of *not that* clauses denying aspects related to the propositional content, the target is an inferred psychological state attributable to the speaker or a third-person referent on the basis of what was said. Questions which the *not that* clause answers, so to speak, are: 'how does/did the referent feel about it', 'what does the referent have in mind', 'what is the referent's attitude'. With as many as 39 occurrences (i.e., 26%) of all valid and analyzable spoken instances, this is a fairly strong pragmatic pattern.

Three more subpatterns of propositional-content denials can be identified in addition to those construing psychological states. They are based on personal traits, habits or actions inferred from the propositional content of the previous utterance. The logic of these types can be explained with the help of everyday deductive reasoning.

The first type follows the rationale that the behavior of a person can be a significant clue for a personal trait. This is illustrated in (18):

(18) Joan   *I sang that for* <voice quality: laughing> *hours on end* <end of voice quality> *to Tasha.*

    Heidi   *Yeah it wouldn't work with her.* <pause>

    Joan   ***Not that*** *I can sing but it was er an apology for singing.*

        (KC3 3023)

Here the *not that* clause denies a quality attributable to the speaker which is a natural inference of the previous utterance referring to an action. The inferential logic is something like 'if someone sang at a given time, it is very likely that they have the quality of being able to sing'. A similar kind of reasoning motivates the denial of inferred habits: 'if someone behaves in a certain way on one occasion, it may be the case that they behave like that on a regular basis'. This is illustrated in (19), where two women talk about "another pile of bloody washing" (the symbols <-|-> mark overlapping speech):

(19) Judy   <voice quality: yawning> *I've got all mine out* <-|-> *to dry* <unclear> <end of voice quality> .

    Doreen   <-|-> *Take them upstairs.* <unclear>

    Judy   <yawn>

    Doreen   ***Not that*** *I* <-|-> *normally carry* <unclear> <-|->

        (KCR 83)

Even though Doreen's *not that* utterance is not complete, we can assume that she rejects the inference that she always carries the laundry upstairs, which one

could make on the basis of the request *take them upstairs*. With 13 instances, this pattern is much less frequent than the one denying a personal trait (23 instances).

A fourth type of denial related to propositional content is based on inferences motivated by the question 'how was the referent reacting to the situation described?' For example, in (20), the speaker is telling a story about a situation where another girl's breasts were visible to her and rejecting the likely inference that her glance was attracted by this unusual sight:

(20) *God! Constance has got a skirt like that* <pause> *and her boobs are just like hanging out! I swear! But **not that** I was looking or anything like that*!
(KCW 2838)

6.2.2.2 *Not that* denying aspects inferable from the preparatory condition

Searle's preparatory condition captures the epistemic basis of assertions, i.e., the question 'does the speaker have sufficient evidence for his or her assertion?' Its role in *not that* clauses can be illustrated by example (21), where the speaker is apparently talking about a hotel:

(21) *Night porter, Q is lifts to all floor, and S is swimming pool* <pause> *and I know there's a swimming pool because* <pause> ***not that** I ever used it, but* <pause> *ma-- it, it was when I booked Francis.*
(KBD 5181)

The *not that* clause in this sequence rejects one possible epistemic foundation licensing the earlier explicit assertion *I know there's a swimming pool*, namely that the speaker has actually used the pool himself.

While denials of the epistemic basis of previous utterances by the same speaker (as in 21) are fairly rare, the strong pattern NOT THAT I KNOW OF (Section 5.3.1) can also be explained along these lines. This is particularly remarkable, as at first sight this pattern seems to be sufficiently lexically fixed, syntactically odd and semantically noncompositional as well as distinct from other *not that* uses for it to be explained independently as an idiom or fixed expression. Consider (22):

(22) None   *Are they here now?*
      Cassie   *No.* <pause>
      None   *Are they related to anyone?*
      Cassie   *No well **not that I know**.* (KP4 1882)

In this example, *not that I know* does not occur in turn-initial position but is preceded by another denial (i.e., *no*) and the marker of hesitation or repair-initiator *well*. If one sets up a simple scale of possible negative replies to alternative questions by the previous speaker (see (23)), *not that I know of* stands between the flat denial *no* and the evasive *I don't know*:

(23) *Are they related to anyone?*
    (a) *No, they aren't.*
    (b) *Not that I know (of).*
    (c) *I don't know.*

(23b) does not deny the proposition invoked by the previous speaker (i.e., it does not mean 'they are not related to anyone'), but denies the epistemic basis that would be required for a flat *no*. In line with other uses of *not that*, it can be paraphrased as 'I deny that it is the case that I know'. *Not that I know* functions as a disclaimer here, reducing the speaker's commitment to the certainty of the first denial, i.e., to its epistemic basis. While one could argue that this use of *not that* actually denies or tones down the previous *no*, an alternative explanation is that *not that I know (of)* denies the expectation, which is part of any serious question, that the interlocutor confronted with the question knows an answer. This provides a viable account of the substantial number of cases (cf. Section 6.1), in which *not that I know of* is a direct response to a yes-no question.

6.2.2.3 *Not that* denying aspects inferable from the sincerity condition

As far as the corpus data are concerned, this can hardly be called a pattern, since there are only two occurrences of *not that* which qualify as instances of a pragmatic link based on the truth of propositions. In one case, (24), the *not that* clause explicitly denies that a previous, hypothetical (cf. *suppose that* ...) proposition by the same speaker is true.

(24) ... *suppose that I think that one of the things that er makes me who I am is erm that I have a good head for numbers. Erm,* **not that that's true**. *Erm, but someone else might think this erm about themselves.*
    (HYD 181)

The second instance, rendered as (25), is particularly interesting from a pragmatic point of view, because here the *not that* clause cancels an existential presupposition of the previous utterance (cf. the reference to Geurts 1998 in Note 14).

(25) Unknown A: *Yeah, but they wouldn't use it, their manager or something like*
                 that, would say <-|-> <unclear> <-|->
      Unknown B: <-|-> *Their agent.*
      Unknown A: *Their agent, yeah.* **Not that** <voice quality: laughing> *they'd*
                 have one <end of voice quality> *but*
                 (JT4 109)

6.2.2.4 *Not that* denying inferences concerning the reason for making a previous utterance

In 11 instances in the data, all of them continuations of the same speaker with the first person pronouns *I* or *we* as subjects, the *not that* clauses deny an inference concerning the speaker's intentions in uttering a previous proposition. These come closest to denials of potential indirect speech acts in Searle's framework. Frequent means of achieving this are collected in Table 10 above under the label NOT THAT I'M COMPLAINING (cf. Section 5.3.5). Some of these denials actually make use of speech-act reporting verbs (cf. 26), others describe speakers' goals in different ways as illustrated in (27). Incidentally, Example (27) is very special in that the proposition the intention behind which is denied follows rather than precedes the *not that* clause. In short, it is an exceptional case of the cataphoric rather than anaphoric denial also mentioned by Delahunty (see Section 6.1 above):[15]

(26) *Er, the economic situation is is entirely their fault, er,* **not that I'm saying** *that*
      the Labour would do any better . . .
      (FX5 598)

(27) *but I think it's fair to point out that there is actually a difference of view, so it's*
      not an absolute position, that you can't identify more land within within
      Southern Ryedale, and indeed, erm, **not that** *I want to raise the Local*
      Government Commission's head again, but of course the Local Government
      Commission is proposing that York be a unitary authority expanded . . .
      (HVJ 193)

Six instances of *not that* found in the spoken corpus reach back even further, as it were, in the preconditions for felicitous speech acts and deny an even more fundamental aspect of utterances: the reason why the utterance was made in the first

---

15 Two anonymous reviewers have rightly pointed out that utterances of the cataphoric type *Not that it's any of my business, but . . .* appear to be fairly familiar. That they are not found in the BNC demonstrates the limits even of such an extensive database as this.

place. This touches upon the essential condition of assertive or representative speech acts, and it has to do with Grice's (1989: 27) maxim of *relation*. The point is that in producing an utterance, speakers commit themselves to the assumption that they have a good reason for saying what they are saying, and that what they are saying is sufficiently relevant, noteworthy and consequential to be worth communicating. It is this assumption that is denied by instances of the pattern NOT THAT IT MATTERED, which, as we have seen in Section 5.3.2, is frequent in the corpus section taken from fiction and verse sources. The logic of these uses is paraphrased in (28):

(28) If the speaker states p, the hearer assumes that the speaker wants the hearer to believe that p is relevant to the hearer (and/or the speaker), but this is not the case.

In the large numbers of instances in fiction an additional twist arises from the multilayered communication situation involving narrated events, on the one hand, and the stance of an authorial narrator, on the other. These two levels of communication are usually conflated in uses of *not that it mattered*, as it tends to remain open whether the proposition denied is potentially relevant to the narrator of the story, the reader of the story, the referent in the narrated event or indeed other, often unspecified figures playing a role in the narrated events (cf. Example 9 in Section 5.3.2: *She simply couldn't keep house – **not that** that mattered so much . . .*).

   In addition to the uses collected in 5.3.2, there are other means of denying a potential reason for having made a previous utterance, among them explicit rejections of their significance to the hearer, as in (29) and (30), as well as the recurrent pattern illustrated by (31):

(29) *Oh that's right, yes, well we'll come to that, yes er, well a byte is a bit **not that** that means a lot, er, what's a bit?*
   (HDV 69)

(30) *The first one* [pause] *is called geometrical isomerism* [pause] *in some books* [pause] ***not that** you ever read, will read any chemistry books* [pause] *but just in case you stumble over one day . . .*
   (F7U 923)

(31) "*It's an extraordinary step,*" Bloxham continued, "*calling anyone who isn't a member to a meeting. But these are extraordinary times. **Not that it's any of your business.**"*
   (CRE 490)

6.3 Summary of corpus findings

In summary, the corpus analysis has yielded the following main results:

– A range of semifixed recurrent lexico-grammatical patterns can be identified which include NOT THAT I KNOW OF, NOT THAT I MIND, NOT THAT IT MATTERS and others.
– To a large extent, the observed lexico-grammatical patterns can be traced back to pragmatic patterns.[16]
– Especially in the dataset taken from spontaneous spoken language, there is a fairly limited set of verbs functioning as predicates in *not that* clauses.
– In a significantly large proportion of the examples investigated, the first-person singular personal pronoun *I* occurs in subject position of the *not that* clause, reflecting the prominent role of the speaker in pragmatic patterns of denial.
– Third-person personal pronouns are also found significantly frequently in the subject slot in *not that* clauses, full noun phrases are rare.
– If *it* occurs as the subject in *not that* clauses, it tends to be used in extended anaphoric reference to propositional antecedents; as far as the spoken data are concerned, there were no examples where the antecedents referred to by *it* were expressions referring to concrete entities.

7 Deriving two competing hypotheses on *not that*

The discussion so far has presented two opposing views of the nature of *not that*. On the one hand, in Section 3 a rule-based and generalized description of *not that* was suggested, essentially on the basis of Delahunty (2006), which treats the structure as a conventional but compositional syntactic entity whose specific communicative impact can be computed online with the help of syntactic rules and general pragmatic principles. On the other hand, Sections 5 and 6 have shown that the authentic uses of *not that* culled from the BNC are not distributed randomly with regard to their environments, but show distinct regularities on various levels of description. A range of lexico-grammatical and related prag-matic patterns with different frequencies of occurrence were identified. While this finding *per se* is by no means opposed to predictions made by the rule-based account, a usage-based approach will be interested in probing the theoretical and

---

16 Cf. Schmid (forthcoming) for a discussion on the contribution of pragmatic associations to the emergence and stability of lexico-grammatical patterns of various types.

cognitive status of these patterns, whereas a rule-based one would simply not bother any further. A possible interpretation of the corpus findings in a usage-based framework could argue thus:[17] looking at the patterns found, it does not seem far-fetched to assume – in a first step – that different frequencies of patterns in corpora can be interpreted as being indicative of different degrees to which these patterns are collectively *conventionalized* in a speech community. More frequent patterns can be assumed to have a place in the repertoire of a larger number of speakers than less frequent patterns. That is why they are used more frequently in the first place. Frequency of usage is thus assumed to be an indicator of degrees of the diffusion (Schmid forthcoming), propagation (Croft 2000: 95–116, 166–195) or usuality (Blank 2001: 1597) of a given linguistic structure in a speech community. In a second and more daring step, it can be hypothesized that there is some match or even correlation between the frequencies of usage patterns and the way in which these patterns are *entrenched* in the minds of individual speakers. This assumption is based on a definition of entrenchment in terms of a continuously ongoing reorganization of individual communicative knowledge which is subject to the exigencies of the social environment and thus to usage effects such as frequency and recency (cf. Bybee 2006, 2010; Schmid forthcoming). In such an account, grammatical knowledge is not represented in a rule-based fashion, but in the form of more or less highly routinized associations to elements, chunks or schemas (cf. e.g., Goldberg 2006; Langacker 2008: 22–23; Bybee 2010: 33–56).[18] If one further argues that more routinized schemas are more automatized and therefore more readily activated than less routinized ones (cf. Langacker 2008: 16; Bybee 2010: 108 *et passim*), different degrees to which knowledge about given grammatical structures is entrenched can be

---

17 Cf. Schmid (forthcoming) for a more detailed discussion of the framework suggested here.

18 Two terminological notes are in order here. First, unlike Langacker (e.g., 2008: 21) and Bybee (e.g., 2010: 151), I emphasize the need to separate *cognitive* processes taking place in the minds of individuals from *social* processes taking place in speech communities. The former are subsumed under the label *entrenchment*, the latter under the label *conventionalization* (cf. Schmid forthcoming for more details). It is only the interaction of processes on both levels, the micro-level and the macro-level, which ultimately brings about grammar and is responsible for language change (cf. Croft 2009; Enfield 2010). Secondly, the non-trivial question whether *not that* and the patterns surrounding it are to be treated as *constructions* in terms of construction grammar (e.g., in Goldberg's sense), *chunks* (e.g., in Bybee's sense) or *schemata* (e.g., in Langacker's sense) will not be discussed here (see again Schmid forthcoming for more details). While there are important differences between these notions, for the purposes of this paper it will be sufficient to note that these unit-like symbolic form-meaning pairings are to be distinguished from the types of linguistic rules postulated, e.g., in generative grammar.

measured by the speed and ease with which these chunks or schemata are activated.

The contrast between the rule-based and the usage-based positions allows the formulation of two general competing hypotheses, the rule-based hypothesis H_0 and the usage-based one H_1.

H_0 (general): the patterns observed are just patterns of usage emerging from recurrent patterns of communication. They are irrelevant for a description of the grammar of individual speakers.

H_1 (general): the patterns observed are indicators of more or less highly routinized associations to chunk-like schemas in the minds of individual speakers, which constitute a part of their grammatical knowledge.

To test the validity of the two general hypotheses, they must be operationalized in a more specific way in the light of the corpus findings. In the case of the rule-based hypothesis H_0, this does not require too much effort:

H_0 (specific): the observed frequencies of usage patterns of *not that* do not have an effect on the ease and speed of the processing of *not that* sentences, since there is no representation of these patterns in the grammars of individual speakers.

The specific usage-based hypothesis H_1 has to be formulated on the basis of the corpus analysis and the frequencies and fixedness of individual patterns. The following assumptions can be derived from the corpus findings presented in the previous sections by selecting particularly distinct differences between the patterns identified (cf. Section 6.3):

H_1 (specific): other things being equal,
– the highly conventionalized pattern *not that I know of* is more routinized and therefore activated faster and more easily than less conventionalized patterns;
– uses with *I* in the subject position in the *not that* clause are more routinized and therefore activated faster and more easily than uses with *it* in this position;
– uses with *it* in the subject position with anaphoric reference to concrete objects are less routinized and therefore activated more slowly and less easily than uses with *it* in the subject position with abstract anaphoric reference.

8 Testing the hypotheses

8.1 Method, material, procedure and participants

As explained in the previous section, the notion of entrenchment and its key process of routinization are operationalized in terms of ease and speed of processing. The rationale behind the – admittedly rather crude and simple – method used in this study builds on this operationalization. It proceeds from the idea that in language production, pragmatic needs trigger the activation of linguistic structures which are suitable for the encoding of these pragmatic needs (cf. Schmid forthcoming). Given a certain communicative goal, the usage-based hypothesis H_1 predicts that more frequent and therefore more routinized structures will be activated more easily and rapidly than less frequent and less routinized ones. This way the speed of activation of contextually appropriate linguistic structures can be taken as an approximation to tapping into degrees of routinization and "depths" of entrenchment of associations to, or "mental representations" of, these structures. To test the hypotheses presented in the previous section, a simple Cloze test was designed, in which test participants were confronted with a short text serving as a pragmatic cue – or prime – and asked to fill a gap. In order to force test participants to activate the structures that would come to their minds first, they were put under severe time pressure. Rather than measuring the time participants needed to fill in the gaps – a measure which is subject to a large number of confounding factors including speed of reading and writing – the dependent variable was a simple count of correct, i.e., targeted, insertions into the gaps within the time period allowed.

Ten authentic extracts from the BNC and one fabricated passage were used as test stimuli. All passages were between 42 and 45 words long and were introduced by a brief note providing the context and the type of source. In all stimuli, there were two subsequent gaps towards the end of the text, which had to be filled by the informants. Each of the subjects was confronted with 11 texts and had thus to fill in 11 gaps consisting of two words. Three texts out of each set had *not that* as a target item, the other eight were distractors targeting other words. Each set began with a fairly simple and predictable gap – to be filled by *but also* which was primed by the prior occurrence of *not only* – to allow subjects to tune into the task. Other distractor texts targeted *let alone* and more or less unpredictable gaps. The three texts targeting *not that* were presented in different orders to the individual participants, but were always placed as the second, seventh and eleventh items presented. None of the participants reported that they had been aware of the fact that the study investigated the structure *not that*.

30 students from Nottingham Trent University took part in the study, all of whom were native speakers of English.[19] In order to control the time allotted for the completion of the 11 test items and implement the time pressure, each stimulus text was presented on a single page and subjects were asked not to turn pages before being told to do so (see test instructions rendered in the appendix). Time allowed for reading each text and filling in the gap was 25 seconds.

To test the specific hypotheses listed at the end of Section 7, the *not that* stimuli targeted the following four patterns:

A. *not that I know of*
B. *not that I mind*
C. *not that it mattered*
D. *not that it* with anaphoric reference to a concrete object.

These were operationalized by the following four passages, of which all except the last one were taken verbatim from the BNC. The passage targeting the rare usage D had to be invented as there was no instantiation to be found in the corpus. With the help of two native speakers of English it was ensured that this passage sounded as natural as possible.

A. *not that I know of*
   [Taken from fiction: two people talking about a man who disappeared]

   I think he'd got himself involved with the law. Don't ask me how, I didn't know and didn't want to. It was a relief to me when he went." – "Did he have a gun?" – "_____ _____ I know of."

B. *not that I mind*
   [Taken from casual conversation: about part-time work]

   you know, the point is you're only working a few days. But it's surprising, the few teaching days I do, they really do cut into my time at home. _____ _____ I mind but it just means . . .

C. *not that it mattered*
   [Taken from a biography: about someone called David]

   His father would now and again telephone me saying that David was down to his last few pounds and that he worried about money. _____ _____ it mattered very much because he was living in my flat and wasn't going to starve, . . .

D. *not that it* with anaphoric reference to a concrete object (fabricated)
[Taken from casual conversation: about a new restaurant]

Mary and I tried this new restaurant yesterday, the one on Morley Road, you know. You can't imagine the food they served us. I hardly touched my chicken salad – _____ _____ it wasn't fresh or anything, but the dressing was so awful . . .

The 30 participants were divided into three groups of ten, each of which was confronted with three of the four target passages in different orders:

Group 1: D – C – A
Group 2: D – B – C
Group 3: A – D – B

In this way, effects of the order of presentation were minimized, while ensuring that each target was presented to at least 20 informants and that the most interesting item, the nonattested type D, was presented to all 30 informants. As already mentioned, the order of the distractor texts was kept constant.

8.2 Predictions of the two hypotheses

The rule-based hypothesis H_0 predicts that all four gaps should be equally easy or difficult to fill, since all four stimuli require the activation of the same type of general knowledge about the syntax and pragmatics of *not that*. The usage-based hypothesis H_1 predicts that individual participants would find it easier to fill the gap targeting pattern A than that targeting B, the gap targeting B easier than C and C easier than D. In addition, it predicts that the proportion of informants who would be able to fill the gaps in these four target texts would decline from A to B, C and D, as pattern A was considered more conventionalized and routinized than pattern B, pattern B more than pattern C and pattern C more than pattern D.

8.3 Results and discussion

Table 12 provides a summary of the absolute and relative numbers of correct answers given by the informants.

Target	n	Correct	Proportion
A. "not that I know of"	20	20	100.00%
B. "not that I mind"	20	12	60.00%
C. "not that it mattered"	20	5	25.00%
D. "not that it" with reference to concrete object	30	5	16.66%

Table 12: Summary of results of Cloze test

As the table shows, the proportion of correct answers decreases from 100% for target A to 60% for B, 25% for C and 16.66% for target B. All differences between the four conditions are highly significant (according to a logistic regression with random person effects), except the one between condition C and D.[20] Overall, the predictions of the usage-based hypothesis H_1 thus seem to be borne out by the results of the Cloze test. The ease and speed of activation of *not that* triggered by specific contextual and cotextual cues varied systematically with the relative frequencies of the patterns found in the BNC and thus with their hypothetical degrees of conventionalization. What the proportions of correct answers seem to indicate is that the most "chunked" pattern *not that I know of* is routinized and entrenched as a lexically filled chunk in the minds of all test participants to such an extent that they are able to activate it with the help of suitable contextual cues even when put under severe time pressure. The target *not that I mind*, which is in line with the majority of uses in the corpus, since *I* is the subject in the *not that* clause, was activated by a larger proportion than *not that it mattered*, which also occurs less often in the corpus. And finally, the gap in D, for whose completion no pattern seems to be available, was filled correctly by only one-sixth of the informants, though it must be added that – partly due to the small number of participants in the test – the difference to condition C was not significant. The fact that five participants were able to work out the solution for target D indicates that speakers seem to be able to activate associations to a very general schema of *not that*, in addition to the more specific ones, which may very well "contain" the information captured in Delahunty's (2006) relevance-theoretic account of *not that* sentences. This only confirms that the network representing linguistic knowledge is "heteromorphic" (Wray 2008: 12, 20), i.e., marked by multiple associations routines and a considerable degree of redundancy (Nattinger and DeCarrico 1992: 23; Bybee 2010: 24; for neurological evidence, see Capelle et al. 2010: 198–199). For

---

20 Thanks to my colleague Helmut Küchenhoff from the Department of Statistics at Ludwig Maximilians University Munich, for doing the statistics.

more or less any given linguistic element many different routinized associations on different levels of schematicity compete for activation.

In short, in spite of its simple methodology the experiment carried out provides evidence that there is a correlation between observed discourse frequencies of usage patterns of linguistic structures and the speed and ease of their activation. This supports the most fundamental idea of usage-based linguistic models that the grammatical knowledge of individual speakers is influenced by frequency of exposure and casts serious doubt on the rule-based hypothesis. If the grammar stored in the mind of a given participant contained no specific knowledge about *not that* over and above the rule-based, compositional information that *not that* represents 'negation of proposition encoded in the *that*-clause' (cf. Section 2 above), then – *ceteribus paribus* – there would be no reason why he or she should be able to fill the gap in target A but miss out for targets B, C and/or D. Even if one has serious doubts regarding the validity of the test as a way of tapping into individual grammars, this finding is difficult to explain with a view of grammar which completely separates usage and grammar.

9 Conclusion and outlook

This paper has presented a detailed corpus-based analysis of *not that* sentences elucidating their syntactic, semantic and pragmatic characteristics. The findings of this analysis, especially those pertaining to the recurrence of semi-fixed patterns, were exploited to support the idea – epitomized in the label *usage-based grammar* – that the way in which linguistic knowledge is entrenched in the minds of individual speakers is sensitive to stochastic information about the frequencies and degrees of conventionality of linguistic structures as indicated by corpus data. Participants in a Cloze test performed better and faster in producing the structure *not that* if the pragmatic context and linguistic cotext prompted them to activate knowledge related to patterns of usage identified beforehand by means of the corpus analysis. More frequent and more chunk-like patterns were activated with more accuracy and greater speed than less conventionalized patterns.

While the corpus-related work presented in this study arguably has a very sound methodological basis, it should not be left to critics of the usage-based approach to point out that the experimental method is of course open to criticism. Most fundamentally, the objection could be raised that the Cloze design does not test anything that comes close to reflecting "represented linguistic knowledge", but rather targets the participants' ability to understand the pragmatic intentions suggested by the linguistic cotext provided in the task and to activate an appropriate linguistic structure. While this cannot be rejected offhand, it also seems indis-

putable that represented knowledge related to the targeted patterns plays a role in the task, especially if it is solved under time pressure, and has an effect on the performance of the participants. If grammar is to be defined as individually entrenched and collectively conventionalized knowledge of linguistic structures, then the results presented in this paper strengthen the view that usage may be more than usage after all. Potential future work pursuing the line of argumentation proposed here in a programmatic fashion should consider better experimental designs and try to recruit a larger number of participants.

References

Baker, Charlotte. 1975. This is just a first approximation, but. . . *Chicago Linguistic Society* 11. 35–45.

Blank, Andreas. 2001. Pathways of lexicalization. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals* Vol. II (= Handbücher zur Sprach- und Kulturwissenschaft 20.2), 1506–1608. Berlin & New York: De Gruyter.

Boye, Kaspar & Elisabeth Engberg-Pedersen (eds.). 2010. *Language usage and language structure*. Berlin & New York: De Gruyter Mouton.

Bybee, Joan. 2006. From usage to grammar: The mind's response to repetition. *Language* 82(1). 711–733.

Bybee, Joan. 2007. *Frequency of use and the organization of language*. Oxford: Oxford University Press.

Bybee, Joan. 2010. *Language, usage and cognition*. Cambridge & New York: Cambridge University Press.

Cappelle, Bert, Yury Shtyrov & Friedemann Pulvermüller. 2010. Heating up or cooling up the brain? MEG evidence that phrasal verbs are lexical units. *Brain and Language* 115(3). 189–201.

Clear, Jeremy. 1993. From Firth Principles: Computational tools for the study of collocation. In Mona Baker, Gill Francis & Elena Tognini-Bonelli (eds.), *Text and technology. In honour of John Sinclair*, 271–292. Philadelphia & Amsterdam: John Benjamins.

Church, Kenneth W. & Patrick Hanks. 1990. Word association norms, mutual information & lexicography. *Computational Linguistics* 16(1). 22–29.

Croft, William. 2000. *Explaining language change: An evolutionary approach*. London: Longman.

Croft, William. 2009. Toward a social cognitive linguistics. In Vyvyan Evans & Stéphanie Pourcel (eds.), *New directions in cognitive linguistics*, 395–420. Amsterdam & Philadelphia: John Benjamins.

Croft, William. 2010. The origins of grammaticalization in the verbalization of linguistic experience. *Linguistics* 48(1). 1–48.

Dąbrowska, Ewa. 2009. Words as constructions. In Vyvyan Evans & Stéphanie Pourcel (eds.), *New directions in cognitive linguistics*, 201–223. Amsterdam & Philadelphia: John Benjamins.

Dąbrowska, Ewa. 2010. The mean lean grammar machine meets the human mind: Empirical investigations of the mental status of linguistic rules. In Hans-Jörg Schmid & Susanne Handl (eds.), *Cognitive foundations of linguistic usage patterns*, 151–170. Berlin & New York: De Gruyter Mouton.

Delahunty, Gerald P. 2001. Discourse functions of inferential sentences. *Linguistics* 39(3). 517–545.

Delahunty, Gerald P. 2006. A relevance theoretic analysis of *not that* sentences: "*Not that there is anything wrong with that*". *Pragmatics* 16(2/3). 213–245.

Enfield, Nick. 2005. Micro- and macro-dimensions in linguistic systems. In Sophia Marmaridou, Kili Nikiforidou & Eleni Antonopoulou (eds.), *Reviewing linguistic thought. Converging trends for the 21st century,* 313–325. Berlin & New York: Mouton de Gruyter.

Evert, Stefan & Brigitte Krenn. 2001. Methods for the qualitative evaluation of lexical association measures. In *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics, Toulouse, France*, 188–195.

Fillmore, Charles J., Paul Kay & M. Catherine O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64. 501–538.

Firth, J.R. 1957. *Papers in linguistics 1934–1951*. London: Oxford University Press.

Geurts, Bart. 1998. The mechanisms of denial. *Language* 74. 274–307.

Goldberg, Adele E. 2006. *Constructions at work. The nature of generalization in language*. Oxford: Oxford University Press.

Greenbaum, Sidney. 1970. *Verb-intensifier collocations in English: An experimental approach*. The Hague: Mouton.

Grice, Paul. 1989. *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Herbst, Thomas. 1996. What are collocations: Sandy beaches or false teeth. *English Studies* 4. 379–393.

Hoffmann, Sebastian, Stefan Evert, Nicholas Smith, David Lee & Ylva Berglund Prytz. 2008. *Corpus linguistics with* BNC*web: A practical guide.* Frankfurt am Main: Peter Lang.

Hopper, Paul. 1987. Emergent grammar. *Berkeley Linguistics Society* 13. 139–157.

Horn, Laurence R. 2001[1989]. *A natural history of negation*. Stanford, CA: CSLI Publications.

Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Jespersen, Otto. 1917. *Negation in English and other languages*. Copenhagen: Bianco Lunos Bogtrykkerei.

Jones, S. & John M. Sinclair. 1974. English lexical collocations. *Cahiers de Lexicologie* 24. 15–61.

Kjellmer, Göran. 1982. Some problems relating to the study of collocations in the Brown Corpus. In Stig Johansson (ed.), *Computer corpora in English language research 1975–1981*, 25–33. Lund & Bergen: Norwegian Computing Centre for the Humanities.

Langacker, Ronald W. 1988. A usage-based model. In Brygida Rudzka-Ostyn (ed.), *Topics in cognitive linguistics*, 127–161. Amsterdam & Philadelphia: John Benjamins.

Langacker, Ronald W. 2008. *Cognitive grammar. A basic introduction*. Oxford: Oxford University Press.

Manning, Chris & Hinrich Schütze. 2001. *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.

Mantlik, Annette. 2011. *The historical development of shell nouns. A diachronic study of abstract noun constructions in English*. Munich: Ludwig Maximilians University dissertation.

Nattinger, James R. & Jeanette S. DeCarrico. 1992. *Lexical phrases and language teaching*. Oxford & New York: Oxford University Press.

Newmeyer, Frederick J. 2003. Grammar is grammar, and usage is usage. *Language* 79. 682–707.

Schmid, Hans-Jörg. 2000. *English abstract nouns as conceptual shells. From corpus to cognition*. Berlin & New York: Mouton de Gruyter.

Schmid, Hans-Jörg. 2011. Tracing paths of conventionalization from the Bible to the BNC: A concise corpus-based history of the *not that* construction. In Renate Bauer & Ulrike Krischke (eds.), *More than words. English lexicography and lexicology past and present. Essays presented to Hans Sauer on the occasion of his 65th birthday, part I*, 299–316. Frankfurt: Peter Lang.

Schmid, Hans-Jörg. forthcoming. Lexico-grammatical patterns, pragmatic associations and discourse frequency. In Thomas Herbst, Hans-Jörg Schmid & Susen Faulhaber (eds.), *Constructions – collocations – patterns*. Berlin & New York: Mouton de Gruyter.

Searle, John R. 1975. Indirect speech acts. In Peter Cole & Jerry L. Morgan (eds.), *Syntax and semantics,* vol. 3, 59–82. New York: Academic Press.

Searle, John R. 1976. A classification of illocutionary acts. *Language in Society* 5. 1–24.

Simpson, John. n.d. *Oxford English dictionary*, online edn. Oxford: Oxford University Press. http://dictionary.oed.com/entrance.dtl

Sinclair, John M. 1990. *Collins COBUILD English grammar*. London & Glasgow: Collins.

Sperber, Dan & Deirdre Wilson. 1995. *Relevance: Communication and cognition*. Oxford: Blackwell.

Stubbs, Michael. 1995. Collocations and semantic profiles. On the cause of the trouble with quantitative studies. *Functions of Language* 2. 23–55.

Summers, Della (ed.). 2005. *Longman dictionary of contemporary English*, 4th edn. Harlow: Longman.

Wray, Alison. 2008. *Formulaic language: Pushing the boundaries*. Oxford: Oxford University Press.

Appendix: Test instructions

Dear participant,

Many thanks for taking part in this questionnaire study. Your help is greatly appreciated. The purpose of this study is **not** to test your linguistic proficiency, but to test a hypothesis concerning the way language is processed. Of course, the study is anonymous.

> Please do not turn the pages before you are told to do so.
> This is very important.

On the next sheet you will be asked to provide information on your age, sex and native language. This sheet is followed by 11 pages containing very short excerpts from authentic texts. You will first find some information about the source and content of the passage and then the text itself.

Two words have been removed from each of the 11 texts. You are asked to guess what these words are and write them in the empty slots indicated. Please read the background information and the passages slowly and carefully and then make your choice. You will be instructed to turn the page after 25 seconds. Please do so immediately when asked, even if you have not managed to come up with a solution. Do not turn the pages before instructed to do so. If you can't think of anything sensible, please guess or just leave a blank.

Personal information

Age: _____

Sex: female ☐   male ☐

First ('native') language: _____