The development of comment clauses[1]

Gunther Kaltenböck
University of Vienna

1. Introduction

This chapter investigates the recent development of comment clauses, more precisely 'main clause-like' comment clauses (Quirk *et al.* 1985: 1112), such as *I think, I suppose, I guess*, which have also been referred to as parenthetical verbs (Urmson 1952), reduced parenthetical clauses (Schneider 2007), epistemic/evidential parentheticals (Brinton 2008: 220), and complement-taking predicates (Thompson 2002). As illustrated by the examples in (1) – (3), comment clauses typically provide some epistemic qualification of the proposition in a host clause, and can occur in initial, medial or final position. In clause-initial position they may take a *that*-complementizer and can therefore be analysed as matrix clauses, although their syntactic status is far from clear (cf. Kaltenböck 2011). Functionally, initial comment clauses have been shown to have secondary status like in non-initial position (e.g. Thompson 2002, Kärkkäinen 2003).

(1)　Uhm <,> **I think** I was <,,> probably possessive and jealous of my mother <A15/ICE-GB:S1A-072#53>

(2)　Uhm <,> the other thing is **I guess** <,,> to ask whether you've also considered the sort of occupational psychology areas <,> as well as the clinical <A08/ICE-GB:S1A-035#144>

(3)　It was that sort of time of the year **I suppose** <B22/LLC:S-02-10#1006>

In recent years comment clauses have received a considerable amount of attention from various research angles, such as grammaticalization theory (e.g. Thompson and Mulac 1991, Brinton 1996, 2008, Traugott 1995a, Fischer 2007, Van Bogaert 2006, 2009, Boye and Harder 2007), and various historical perspectives (e.g. Palander-Collin 1999, Bromhead 2006), descriptive corpus linguistics (e.g. Stenström 1995, Mindt 2003, Kearns 2007), functional–pragmatic perspectives (e.g. Aijmer 1997, Simon-Vandenbergen 2000, Ziv 2002, Thompson 2002, Kärkkäinen 2003, 2007, 2010, Kaltenböck 2010), Relevance Theory (e.g. Blakemore

---

1990/1991, Ifantidou 2001), prosodic analysis (e.g. Wichmann 2001, Kaltenböck 2008, Dehé and Wichmann 2010), language acquisition (Diessel and Tomasello 2001), or from a cultural perspective (Wierzbicka 2006).

What makes comment clauses interesting as a linguistic category is their ambivalent character, which stems from a discrepancy between usage and structure: structurally they represent clauses, but functionally they are like disjunct adverbials conveying secondary information. This indeterminacy can be attributed to the ongoing process of grammaticalization they are subject to (cf. Section 2). As grammaticalizing elements they are in a state of latent instability and particularly susceptible to change. This is evidenced, for instance, by the adoption of new pragmatic functions (e.g. Aijmer 1997, Kärkkäinen 2003, 2007, Kaltenböck 2008, 2010, Van Bogaert 2006), which signal a shift away from use as marker of epistemic stance to use as a general pragmatic marker. Comment clauses have also been claimed to undergo a process of expansion from their prototypical 'first person form' (e.g. *I think*) to variant forms such as *I would think, I'm thinking* (Van Bogaert 2010b).

The present study investigates to what extent there is corpus evidence for such changes in recent decades, especially with regard to signs of further grammaticalization of the most central of all comment clauses, *I think*, and a possible extension of the class of comment clauses to variant forms. It shows that *I think* has advanced on the path of grammaticalization and is changing from a marker of epistemic modality, typically expressing lack of speaker commitment, to a pragmatic marker with important textual and discourse-organizational functions. In partial compensation for this loss of epistemic function, formal variants such as *I'm thinking, I just think* are increasingly used as comment clauses with modal meaning. The paper argues, finally, that these changes can best be accounted for in a constructional framework which considers taxonomic links to related constructions.

As database for the empirical investigation the study makes use of the *Diachronic Corpus of Present Day Spoken English* (DCPSE), which consists of two parallel subcorpora with data from the *London-Lund Corpus* (LLC), compiled from 1958 to 1977, and from the *British Component of the International Corpus of English* (ICE-GB), compiled in the early 1990s (cf. www.ucl.ac.uk/English-usage/projects/dcpse/design.htm and Aarts, Close and Wallis, this volume). It thus covers a period of roughly 30 years and its spoken data make it particularly useful for studying ongoing change, as spoken language is well-known to be at the forefront of linguistic innovation. Owing to its relatively small size of a total of 885,436 words, however, DCPSE does not always yield sufficient instances for the different comment clause types. The study therefore focuses on *I think* as the most frequent and most

prototypical of all comment clauses and supplements the DCPSE data with data from other corpora, notably the *Corpus of London Teenage Language* (COLT), comprising 500,000 words of London teenage speech from 1993 (cf. Stenström, Andersen and Hasund 2002), and the *Corpus of Historical American English* (COHA, beta version), comprising 400 million words of written American English from 1810 to 2009 (cf. http://corpus.byu.edu/coha). The use of these corpora, of course, raises methodological issues, as their differences in make-up (text types), mode (spoken, written), and social/regional variation (British English, American English, London teenage speech) do not allow for direct comparisons between them. In the case of DCPSE and COLT the differences can be somewhat minimized by restricting the comparison to the DCPSE text type 'Informal face-to-face' (category B), which corresponds with the prevailing register in COLT. Nonetheless, it is clear that the results from the different corpora show only limited comparability. In the absence of sufficient spoken data from previous periods, however, drawing on other corpus types seems justified (with the appropriate methodological caveats). Despite the obvious differences between them, the COHA data can usefully complement DCPSE, for instance by providing insights into the development of the less frequent variant forms and their correspondence with *I think*.

The chapter is structured in the following way. After a brief overview of the presumed historical development of comment clauses in Section 2, Section 3 discusses general frequencies of the comment clauses *I think, I suppose, I believe, I suspect, I expect, I understand, I imagine, I guess, I reckon* in DCPSE. Section 4 focuses on the most frequent and prototypical of these, *I think,* and examines evidence for its further grammaticalization, viz. positional distribution (4.1), use of the *that*-complementizer (4.2), scope (4.3), collocation patterns (4.4), and other uses (4.5). Section 5 traces the development of the less frequent variant forms of *I think* in COHA, which suggests a functional shift of epistemic meaning from *I think* to *I'm thinking, I just think, I'm guessing*. Section 6, finally, argues for a constructional account of comment clauses as an appropriate model to capture their diachronic development. The conclusion in Section 7 sums up the main findings.

2. Origin and diachronic pathways

Various syntactic pathways of development, involving different processes of change, have been proposed for epistemic comment clauses. The difficulty in tracing their trajectory through time lies, not unexpectedly, in the scarcity of data from older periods of English and the unavailability of authentic spoken data, i.e. the mode preferred by comment clauses. To

compensate for this, one approach is to project backwards from synchronic findings, as has been done by Thompson and Mulac (1991). In their influential study of the present-day epistemic parentheticals *I think* and *I guess* they propose a cline from a matrix clause with a *that*-complementizer, to omission of *that*, and finally to a parenthetical disjunct in non-initial position. This process of grammaticalization thus results in a reversal of the matrix clause/complement clause structure with the original matrix clause *I think* being reanalysed as a ''unitary epistemic phrase'' and the original complement clause being reanalysed as the matrix clause (cf. also Traugott 1995b: 38-9). Although intuitively appealing, this ''matrix clause hypothesis'' (Brinton 2008: 246) has been shown to be in conflict with actual historical data. According to Brinton (1996: 239-54), diachronic evidence suggests that first-person epistemic comment clauses such as *I think, I guess, I suppose* originated not in a matrix clause but an adjoined adverbial/relative structure of the type *as I think*. A similar view is expressed by Fischer (2007: 304-5), although she identifies the anaphoric connective element as an adverbial derived from a demonstrative.

As noted in the introduction, the process involved in the development of epistemic comment clauses is generally thought to be one of grammaticalization.[2] Various studies, both synchronic and diachronic, have shown that they undergo many of the changes characteristic of grammaticalization (cf. Kärkkäinen 2003, Van Bogaert 2009, 2010b, Thompson and Mulac 1991, Palander-Collin 1999, Brinton and Traugott 2005, Brinton 1996, 2008, Boye and Harder 2007, Traugott 1995a, Kaltenböck 2008). These changes include ''semantic bleaching' (Traugott 1982) or ''desemanticization' (Heine, Claudi and Hünnemeyer 1991), i.e. loss of the original concrete meaning, 'pragmatic strengthening' (Traugott 1988), i.e. the acquisition of discourse/pragmatic functions, 'subjectification', i.e. increased subjectivity (Traugott 1988, 1995b: 38-9), positional mobility, and possible 'phonological attrition' (Lehmann 1995). As noted by Brinton (2008: 242) comment clauses also conform to Hopper's (1991) principles of grammaticalization, viz. layering, divergence, specialization, decategorialization (cf. also Van Bogaert 2010b). Where comment clauses seem to diverge from prototypical grammaticalization is with regard to some of Lehmann's (1995) parameters, notably condensation (i.e. reduction in scope) and fixation (i.e. loss of syntactic variability). These parameters, however, have been challenged as necessary features of grammaticalization (e.g. Tabor and Traugott 1998, Fischer 2007, Brinton 2008: 244-45 on scope; Van Bogaert 2010b on lack of internal fixation).

---

[2] A different view is expressed by Fischer (2007: 311), who sees parentheticals like *I think* as formulaic tokens undergoing lexicalization.

In terms of their semantic development, it has been noted that comment clauses follow a path which involves the reduction of semantic content (bleaching) while adopting more pragmatic meanings (pragmatic strengthening). This semantic-pragmatic cline has been described as a unidirectional development from propositional to expressive or interpersonal meaning (Traugott 1982) and has subsequently been elaborated into a more complex concept of unidirectional change which includes the following tendencies: from truth-conditional to non-truth-conditional, from conceptual to procedural, from non-subjective to subjective and intersubjective (Traugott and Dasher 2002). Given their increasingly pragmatic function it is not really surprising that comment clauses have also been described as cases of pragmaticalization rather than grammaticalization (Erman and Kotsinas 1993, Aijmer 1997). In a comprehensive definition of grammar, however, which includes pragmatic meaning, comment clauses can still be appropriately described in terms of grammaticalization (cf. Brinton and Traugott 2005: 139).

3. Frequencies in the corpora

As noted in the introduction, the present study is limited to what Quirk *et al*. (1985: 977) term 'main clause-like' comment clauses. The most frequent of these are *I think, I guess, I believe, I suppose, I understand, I reckon, I imagine, I expect, I suspect,* although frequencies vary considerably for the different predicates, with *I think* being by far the most frequent of all (cf. Table 1).

Table 1. Frequencies of comment clauses in DCPSE (normalised per 1 million words)

Comment clause	LLC (~1960s) (464,074 words)		ICE-GB (~1990s) (421,363 words)		Change in frequency
	raw	pmw	raw	pmw	%
I think	1,379	2,971.5	1,187	2,817.0	-5.20
I suppose	202	435.2	141	334.6	-23.12
I believe	51	109.8	33	78.3	-28.69
I suspect	22	47.4	6	14.2	-70.04
I expect	15	32.3	8	18.9	-41.49
I understand	15	32.3	3	7.1	-78.02
I imagine	12	25.8	6	14.2	-44.96
I guess	10	21.5	16	37.9	+76.28
I reckon	9	19.3	5	11.8	-38.86
Total	1,715	3,695.5	1,405	3,334.4	-9.77

If we compare the results from LLC with those from ICE-GB, we note a decrease in frequency (per million words) for all predicates, except *I guess*, whose frequency has increased in ICE-GB by 76.28 per cent. This increase may be attributed to an influence of American English, which is well-known for making greater use of *I guess* (e.g. Van Bogaert 2009: 421; Kärkkäinen 2007). However, the increase of *I guess* has not tested as statistically significant, neither as per million word change nor relative to the total set of comment clauses (see Appendix Figure A1.1 and A1.2 respectively; cf. Also Aarts, Close, Wallis, this volume). Conversely, the only predicates showing significant decrease in absolute terms (per million words) are *I suppose, I suspect*, and *I understand* (cf. Figure A1.1). This means that in their decision to employ a comment clause, speakers increasingly disfavour these three predicates.

Overall, the whole set of comment clauses shows a significant decrease, too. However, while comment clauses are declining in use in DCPSE, within this decline some forms are significantly increasing their share. These are *I reckon* and *I suppose* (as can be seen in the changes relative to the total set in Figure A1.2). On the other hand, *I suspect* and *I understand* not only fall significantly in absolute terms but also in terms of their relative share, together with *I imagine* (cf. Figure A1.2). The changing preferences within the set of comment clauses can be explained, at least in part, by the generally attested colloquialization of English, i.e. a shift to greater informality (e.g. Mair 2006: 183-93, Leech, Hundt, Mair and Smith 2009: 239-49). *I suspect, I understand* and *I imagine* are the more formal variants (with more semantic content), whereas *I reckon* and *I suppose*, as the more informal (and more semantically bleached) variants, are being preferred. What is difficult to account for at this stage is the general decline of comment clauses (in absolute terms). A possible explanation might be their replacement as epistemic markers by other, variant forms (as discussed in Section 5).

A comparison of the figures in DCPSE with the *Corpus of London Teenage Language* (COLT) confirms the general downward trend of comment clauses. Although such a comparison has to be taken with considerable caution owing to the difference in make-up (text type) and speakers (as noted in Section 1), the young age of the speakers and their location in a centre of linguistic innovation can provide clues for the further development of comment clauses. As illustrated by Figures 1.1 and 1.2, COLT shows further reduction in frequency for all comment clauses with the exception of *I reckon* (cf. Figure A2 in the Appendix). The high frequency of *I reckon* is in line with the shift noted within the DCPSE set of comment clause and may be an indication that *I reckon* is a regional variant for

American English *I guess*, which is conspicuously rare in the COLT data.[3] As noted by Fischer (2007: 311), parenthetical phrases such as *I think* are also important markers of personality and group identity. Regional and group-specific differences are perhaps therefore not really surprising.

Figure 1.1. Normalised frequencies (per million words) of comment clauses *I think* and *I suppose* in LLC, ICE-GB and COLT
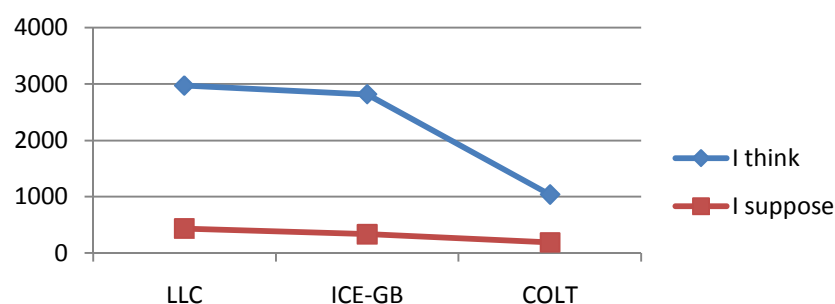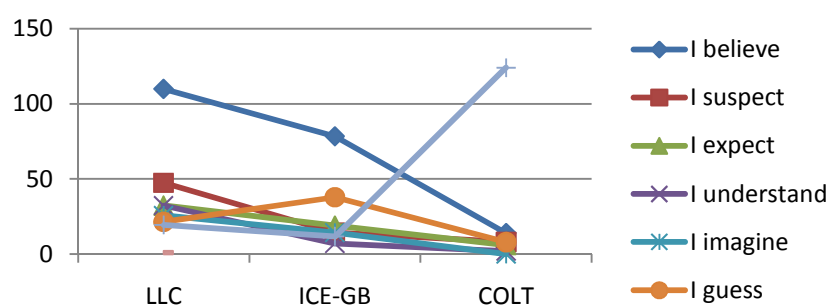

Figure 1.2. Normalised frequencies (per million words) of comment clauses *I believe, I suspect, I expect, I understand, I imagine* and *I guess* in LLC, ICE-GB and COLT

As can be seen from Table 1, the numbers for the individual comment clauses in DCPSE (as well as in COLT) are, however, very low, with the exception of *I think*. The small number of samples makes it difficult to draw any conclusions about the diachronic development of these comment clauses. I will therefore focus mainly on the development of *I think* as their most frequent and most prototypical representative.

4. *I think*: evidence for further grammaticalization?

Although the decrease in frequency could suggest otherwise, *I think* has been claimed to undergo a process of grammaticalization, developing from an epistemic marker into

---

[3] The exact figures for the COLT data are the following, normalised per 1 million words (raw frequencies in brackets): *I think* 1,038 (519), *I suppose* 190 (95), *I believe* 14 (7), *I suspect* 8 (4), *I expect* 6 (3), *I understand* 2 (1), *I imagine* 0, *I guess* 8 (4), *I reckon* 124 (62).

something like a discourse marker (e.g. Mindt 2003) or a textual-interactional device fulfilling various pragmatic purposes (e.g. Ziv 2002, Kärkkäinen 2003, Kaltenböck 2010). This section investigates the question of further grammaticalization by looking at a number of different parameters, namely positional distribution (4.1), complementizer use (4.2), semantic-pragmatic scope of *I think* (4.3), collocational patterns (4.4), and other attested uses of *I think* (4.5), leading to a conclusion in 4.6. The basis for the analysis is provided by the data in Table 2 and the corresponding diagrams in Figures 2.1 and 2.1, which provide the confidence intervals for absolute per million word change and relative change within the set respectively.

Table 2. Overall frequencies of different uses of comment clause *I think* in DCPSE (normalised per 1 million words)

I think	LLC (464,074 words)		ICE-GB (421,363 words)		Change in frequency		
	raw	pmw	raw	pmw	%	χ^2 partial	X vs ¬X 2 x 2 χ^2
Initial (-*that*)	966	2,081.5	829	1,967.4	-5.48	0.00	0.01
+*that*	94	202.5	81	192.2	-5.09	0.00	0.00
Medial	145	312.4	141	334.6	+7.11	1.06	1.20
Phrasal	52	112.0	68	161.3	+44.02	**5.23 sig<0.05**	**5.49 sig<0.05**
Final	122	262.8	68	161.3	-38.62	**8.38 sig<0.01**	**9.05 sig<0.01**
Total	1,379	2,971.5	1,187	2,817.0	-5.20	**14.67 sig<0.01**	

Figure 2.1. Absolute pmw change of different uses of comment clause *I think*, including 95% Newcombe-Wilson confidence intervals (see Appendix 1, Aarts *et al*., this volume). Where the confidence interval does not cross the zero axis the change is considered to be significantly different from zero (i.e. 'significant').
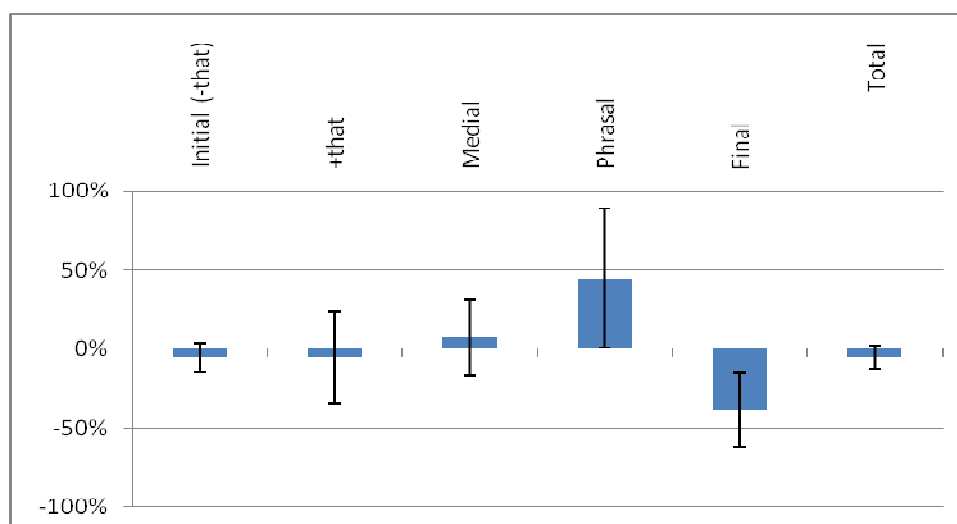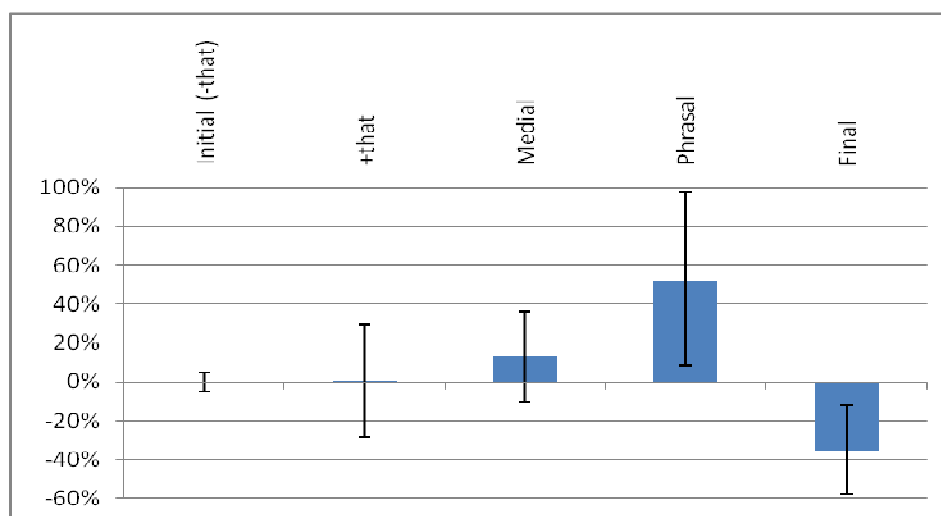
Figure 2.2. Relative change of *I think* within the set of comment clauses overall.

4.1 Positional distribution

A key characteristic of comment clauses is their positional flexibility with regard to the host construction. Identifying the position of a comment clause is, however, not always as straightforward as it might seem and needs to take into account prosodic realisation. Compare, for instance, the following example, where prosodic binding to the right (marked by round brackets) indicates initial position.[4]

> (4)   And then you go up through HArtney WIntney <,,> and it's the same nUmber all
>   the way (**I think** it's the <,> B Three uh one one <B11/LLC:S-01-11#613>

Initial position in the present classification is equivalent to pre-nuclear position, i.e. pre-subject position, which allows for preceding adverbials, but disregards discourse markers (e.g. *well*) and vocatives (e.g. *Peter*). Final position, on the other hand, includes only clause-final position, but allows for afterthoughts, for instance in the form of a right-dislocation. Medial position subsumes all remaining positions except phrasal and initial +*that* (cf. 4.2 and 4.3 for discussion). Medial thus includes *I think* immediately following a relative element (e.g. *The first point which **I think** I've got to resOlve...* <J04/LLC:S-12-04#61>) since this position typically precludes the use of a *that*-complementizer. Instances of *I think* immediately

---

[4] Capital letters in the the *London-Lund Corpus* indicate the nucleus (tonic), i.e. the prosodically most prominent syllable in a tone unit.

following other subordinators (e.g. *...althOUgh **I think** it will reshApe itsElf* <B09/LLC:S-01-09#385>) have been classified as initial, as they do permit complementizer use.

The most striking development from LLC to ICE-GB is the dramatic fall in the number of clause-final *I think* in absolute terms, which drops by 38.62 per cent (from 262.8 to 161.3 occurrences pmw) and has tested as statistically significant in terms of chi-square (cf. Table 2) and Newcombe-Wilson intervals (cf. Figure 2.1). In relative terms, i.e. compared to the other positions within the set, the proportion of final uses has also decreased significantly (as shown by the confidence interval in Figure 2.2). Medial position, by contrast shows a slight, non-significant rise in absolute terms (7.11%; cf. Table 2, Figure 2.1) as well as in its relative share (cf. Figure 2.2). Initial position drops slightly, too, but remains relatively stable with minus 5.48 per cent (Table 2), which corresponds with an unchanged relative proportion in Figure 2.2. (The remaining uses of *I think* will be discussed further below.)

How can we interpret the substantial decrease of final position? A possible explanation lies in the predominant function performed by clause-final *I think*. While initial and medial *I think* play an important role in online planning, for example as a stalling device bridging a hesitation phase (Stenström 1995), or as a discourse link (Ziv 2002), final position seems to be more apt for expressing an epistemic afterthought with the purpose of mitigating a previous statement (cf. also Conrad and Biber 2000: 72). Of course, final *I think* (like the other positional variants) may fulfil a range of pragmatic functions, such as signalling completion and pursuing a response (cf. Kärkkäinen 2003:161-70), but its final (i.e. focus) position seems to give particular emphasis to its original semantic meaning and to foreground its epistemic function, downtoning the previous statement. Compare, for instance, examples (5) and (6).

(5) Uhm <,> well <,,> MArlowe was <,,> a lIttle <,> uh a lIttle After ShAkespeare <,> **I think** <A01/LLC:S-03-01#503>

(6) Yes <,> but it Also is a vEry good nOvel <,,> **I think** <,,> <A01/LLC:S-03-01#712>

Final position also represents the prototypical position for a comment: first you express a state of affairs (Posner's 1973 'commentatum'), then you comment on it. Conversely, initial use may be more prone to semantic bleaching, owing to its high frequency, and is therefore more likely to have its epistemic meaning eroded to that of a 'mere' starting point function (cf. Kärkkäinen 2003: 160) or filler function (cf. Kaltenböck 2010). As noted by Kärkkäinen

(2003: 145) for initial *I think*, 'the degree of uncertainty and doubt expressed by this marker ... is in fact quite low ... this marker may at times simply be doing some rather *routinized work* in conversation organization' (emphasis in original). The decrease of clause-final *I think* in DCPSE can therefore be taken as an indication of a functional change of comment clause *I think* towards a more semantically bleached use with reduced epistemic function. Initial position, as the more typical repository of bleached uses, is consequently not so much affected by the decrease.

4.2 Use of the **that-***complementizer*

Omission of the *that*-complementizer is generally seen as a sign of increasing grammaticalization of initial comment clauses. This view has been expressed, for instance, by Thompson and Mulac (1991), who argue that frequently used main clauses such as *I think* are being reanalysed as 'unitary epistemic phrases' with the omission of *that* as a strong concomitant. Similarly, Torres Cacoullos and Walker (2009: 17) take zero *that* to be 'a measure of the development of discourse formulas'. Although it is doubtful whether comment clauses such as *I think* actually started out in the history of English as matrix clauses with a complementizer, as assumed by Thompson and Mulac (cf. Section 2), historical studies have noted an overall decline of the complementizer at least from the Late Middle English period, with some fluctuation and register variation (Rissanen 1991, Finegan and Biber 1995, Pallander-Collin 1999).

The DCPSE data, however, do not show any significant change in the use of the *that*-complementizer with only a slight fall of 5.09 per cent from 202.5 occurrences per 1 million words in LLC to 192.2 occurrences in ICE-GB (cf. Table 2, Figure 2.1) and an unchanged relative proportion within the set (cf. Figure 2.2). This rather stable development can be attributed to the relatively short time span of 30 years covered by DCPSE and a process of grammatical persistence, as discussed by Torres Cacoullos and Walker (2009: e.g. 34), who argue that just as lexical meaning may persist in grammaticalizing constructions, grammatical properties, too, may persist in the development of discourse formulas (cf. also Section 6).[5] In an attempt to extend the time frame covered, the ICE-GB results for text category B (informal face-to-face) were compared with COLT (cf. Table A1 in the Appendix), which shows a large (and statistically significant) drop of 91.56 per cent (from 118.5 occurrences pmw to only

---

[5] Cf. also Kaltenböck (2011), who shows that the *that*-complementizer after *I think* has an important filler function in spoken discourse, which may also account for its relatively persistent use.

10.0 occurrences). [6] This trend is also reflected in the relative change for *that* within the total set given in Figure A3 in the Appendix.[7] Despite the methodological shortcomings of a comparison with COLT (cf. Section 1), these data provide at least some indication of a change in the speech of younger language users.

4.3 Scope

Although comment clauses have traditionally been assumed to have scope only over a clausal host construction, recent corpus studies have shown that their semantic-pragmatic scope may also be over a phrasal (i.e. non-clausal) constituent (e.g. Kaltenböck 2008, 2009, Kärkkäinen 2003, 2010, Van Bogaert 2006). Typical examples of such uses are given in examples (7) – (10).

(7)   Well of course the two hundred and fIfty pounds which the LAbour gOvernment insIsted on <,> in **I think** nIneteen sIxty-sIx sixty-sEven sIxty-sEven <I01/LLC:S-11-02#71>

(8)   Uh in the uhm <,> **I think** October issue of Computational uh Linguistics there's an attempt to do something of this type <A05/ICE-GB:S1A-024#105>

(9)   We're going to have a very small set (**I think** at Brave) for Edward <,> <E05/ICE-GB:S1B-045#110>

(10)   The best count we've Ever done in a day is uh (nine ten elEven **I think**) <,,> uhm <F08/LLC:S-10-08#231>

Apart from elliptical/incomplete examples, phrasal scope is signalled either by positioning *I think* within a phrase, such as the PP in example (7) and the NP in example (8), or by prosodically binding *I think* to a constituent on its right or left as in (9) and (10) respectively. Pragmatically, these phrasal scope uses are interesting since they often function not so much as epistemic markers (shields)[8] but more like approximators (Prince *et al*. 1982), making

---

[6] The same drop in the use of *that* can be observed by comparing the incidence of 'Initial' with '+*that*', which are also cases of initial *I think*. Such a comparison shows a similar stable development from LLC to ICE-GB (cf. figures in Table 2) and a large fall of +*that* from ICE-GB to COLT (5.82% > 1.2%, cf. figures in Table A1 in the Appendix)

[7] Interestingly, this relative drop of *that* from ICE-GB to COLT constrasts with a sharp rise of initial (-*that*) uses (cf. Figure A3).

[8] In the terminology of Prince at al. (1982: 86) shields indicate lack of speaker commitment. As such they differ from approximators, which indicate non-prototypicalness with respect to class membership.

referring terms or predicates less precise. Typically these are nouns of some semantic 'precision' permitting scalar approximation, such as numerals. In examples (7) and (8), for instance, *I think* can be replaced by prototypical approximators: *in about/around/approximately nineteen sixty-six, in the (something) like October issue* (cf. Kaltenböck 2010 for further discussion).

The corpus data show that such phrasal uses of *I think* have significantly increased in DCPSE both in absolute and relative terms. In absolute terms the ICE-GB component shows a rise of 44.02 per cent compared to the London Lund component (from 112.0 to 161.3 occurrences pmw; cf. Table 2). This change also tested as statistically significant using Newcombe-Wilson confidence intervals for absolute pmw change and relative change within the total set of different uses of comment clause *I think* (cf. Figure 2.1 and 2.1 respectively). A comparison of ICE-GB text category B with COLT shows a similar increase in the share of phrasal uses (cf. Figure A3 in the Appendix). The data are thus indicative of a functional shift within the paradigm towards more frequent phrasal scope uses of *I think*.

The increase of phrasal uses suggested by the data is particularly interesting from a grammaticalization perspective, as scope is an important (and debated) parameter in grammaticalization theory. Although originally grammaticalization was thought to involve narrowing of (structural) scope (cf. Lehmann's 1995 parameter of condensation), it is now accepted as including widening of scope too, especially in the development of pragmatic markers (e.g. Tabor and Traugott 1998, Traugott 1995a, Fischer 2007: ch. 6). The increase of phrasal uses of *I think* would seem to fit in with a narrowing of scope from clausal to phrasal. However, since we are dealing here with semantic-pragmatic rather than syntactic scope, it is more appropriate to see phrasal scope as an indication of more varied uses of *I think* along the lines of 'context generalization' (Heine 2003, Himmelmann 2004).

Brinton (2008: 247) seems to suggest the possibility of scope widening for *I think* when she notes that 'the parenthetical now has scope over the clause and ultimately over larger units of discourse'. Although it is possible for *I think* to have pragmatic scope over more than one clause (cf. Kärkkäinen 2003: 116, 162), it does not normally have scope over larger stretches of discourse. What can be found, however, are cases of unclear scope, where *I think* does not express an epistemic qualification of a host construction but is inserted primarily for purposes of online planning not unlike a filler (cf. Kaltenböck 2010) or for 'routinized work in conversational organization' (Kärkkäinen 2003: 145). Consider, for instance, the following example, where the repeated use of *I think* functions as a turn-taking signal (rather than an epistemic qualifier), expressing A's repeated attempt to gain the floor.

(11) A: Yes Well **I think**

B: <unclear words> was there that's the wonderful thing about it but it's an

awfully long time ago

A: **I think** he **I think** the reason is that it's neither very alternative and free range

like Bryanston <,> and nor is it too stiff like Rugby <B44/ICE-GB:S1A-054#99>

Such 'textual' or 'interactional' uses of *I think* typically occur in initial and medial position (at key structural points in the utterance) and often co-occur with other fillers/discourse markers (e.g. Aijmer 1997: 26, Kärkkäinen 2003: 128), or disfluency features (pauses, hesitation sounds, repetitions, restarts; e.g. Kaltenböck 2008, 2010; cf. also Section 4.4), as illustrated by the following examples:

(12) I mean **I think** really uhm <,,> it's very difficult to to to produce any form of art unless you are driven <B11/ICE-GB:S1A-015#145>

(13) Uh and it it's certainly uh **I think** a fair assumption that uh when talking about innovation most people assume that innovation is something that's carried out within the industrial context <I05/ICE-GB:S2A-037#3>

As a concomitant effect of their semantic bleaching, these textual uses of *I think* show considerable phonetic reduction (to the extent of total omission of the pronoun: e.g. *Think it's a a vital <,> need* <A03/ICE-GB#59>) and are prosodically fully integrated into the host construction (i.e. lacking an accented syllable), especially if in medial position (cf. Kaltenböck 2008, 2010, Dehé and Wichmann 2010). As pointed out by Kärkkäinen (2003: 128), 'the unemphatic encoding (reduced and/or accelerated phonetic realization, lack of stress) of this marker make it a very poor marker of speaker uncertainty at best'.

4.4 Collocation patterns

As noted above, parenthetical *I think* tends to co-occur with other fillers and discourse markers, especially in initial and medial position, which can be seen as a sign of its own semantically bleached status and structural, rather than epistemic, function. An analysis of the occurrence of discourse markers (viz. *actually, well, you know, I mean, like, oh*) immediately preceding or following *I think* in DCPSE shows a slight increase from 22.33 per cent (163/730 instances) in LLC to 26.17 per cent (129/493 instances) in ICE-GB. At the same

time the number of short and long pauses immediately before or after *I think* has dramatically decreased from 51.1 per cent (371/730) to 20.69 per cent (102/493).[9] These figures lend support to a view of further grammaticalization of *I think* in so far as increased co-occurrence with other fillers suggests a similar function for *I think*. As a filling (stalling) device *I think* helps the speaker with online planning by bridging hesitation phases and thereby alleviates production difficulties, as is reflected in the reduction of disfluency features such as pauses.

Further evidence for a lack of epistemic meaning can be adduced from the collocation of *I think* with modal markers. As noted by Kärkkäinen (2003: 128) (cf. also Aijmer 1997: 26), *I think* often occurs with modality markers (such as *probably, maybe*), which can be explained by a lack of tentativeness of *I think*: 'Because it [*I think*] may not ... 'adequately' bring out the speaker's uncertainty ..., other epistemic markers can be argued to perform that function in the utterance' (Kärkkäinen 2003: 129).[10] An analysis of the DCPSE data shows a slight increase in the collocation of *I think* with *probably*, *possibly*, *perhaps*, *maybe* from 2.88 per cent (21/730) in LLC to 4.06 per cent (20/493) in ICE-GB.

4.5 Other uses

The corpus also contains a number of examples of *I think* which are difficult to classify since they occur as independent fragments, often at the beginning of a speaker's turn, without any clear link to a host construction. While some of these result from interruptions or restarts typical of spoken interaction, others can be identified as referring back to an utterance in a previous turn. Compare, for instance, the following examples, where *I think* occurs turn-initially as a separate tone unit (i.e. intonationally separated from the following utterance) and qualifies an utterance the same speaker has made in a previous turn and which has been acknowledged in some form (repetition or reaction signal) by the other speaker.

(14)   A: That should see you through until

       B: Till July

       A: Yes <,> July August

       B: Summer holiday time

---

[9] This change has tested as statistically highly significant (χ^2 = 112.65). The percentage of hesitation sounds (*uh, uhm*) immediately before or after *I think* has remained stable: 17.95% (131/730) in LLC, 17.85% (88/493) in ICE-GB.

[10] There may of course also be cases where *I think* is used to reinforce the modal meaning of *probably, maybe*, etc., as noted by a reviewer.

A: **I think** uh are you going to have a summer holiday this year <A18/ICE-GB:S1A-087#96>

(15) A: Right hand side when you uh get through

B: Right <,>

A: Well I mean yeah when you just walk in

B: Yes

A: **I think** But I might be wrong <B38/ICE-GB:S1A-046#365>

Although such uses are rare, they show that *I think* is flexible enough in its use to function even as an independent response (as an alternative to *I think so*). All of these examples occur in the ICE-GB component of DCPSE, which might be indicative of more recent usage.

Another rare use of *I think*, which can be found in the corpus, is that of a quotative or reporting verb, which reports the speaker's thoughts, often in the narrative part of a conversation. This type differs from comment clause uses by the form of its host clause, which is either interrogative or exclamative or, especially if declarative, preceded by a discourse marker such *oh* or *well*. Compare the following examples:

(16) When I <,> when I feel <,,> rEAlly deprEssed <,,> **I think** what a hOrrible lot <B12/LLC:S-01-12#507>

(17) And Other times **I think** oh well <,> it's quite plEAsant rEAlly becAUse they're all so odd <B12/LLC:S-01-12#424>

This use of *think* has retained much of its original semantic meaning of 'cogitation' (cf. Aijmer 1997: 12). With only 8 occurrences in LLC and 3 in ICE-GB this type is, however, extremely rare and seems to be becoming even less frequent.

Think also retains its full lexical meaning if complemented by a NP (e.g. *This is what I think*) or PP (e.g. *I think of myself as a smoker*). These uses are extremely rare, however, and decrease from LLC (9 NPs, 13 PPs) to ICE-GB (5 NPs, 5 PPs) in relative and absolute terms.

I think so, on the other hand, with its reduced semantic meaning (cf. Hooper 1975: 109), shows an increase from 24 instances in LLC (1.66% of all uses of *I think*; 51.7 occurrences pmw) to 33 instances in ICE-GB (2.65%; 78.3 pmw).

Although the low figures of the different uses discussed in this section do not warrant any far reaching conclusions, they are compatible with a view of continuous bleaching of *I think* and attest to its functional flexibility as a result of semantic erosion.

4.6 Conclusion

As is to be expected, the short time period of roughly 30 years covered by the DCPSE data is not likely to reveal any dramatic grammatical changes. Nonetheless, the corpus data provide a number of indications, all of which point in the direction of further grammaticalization of *I think*. It shows signs of further erosion of its original semantic meaning and is increasingly used not so much as an epistemic qualifier of a host clause proposition, indicating lack of speaker commitment, but as a more general pragmatic marker with important textual (structural) and interactional function. Its textual function consists in acting as a stalling or filler device, which provides time for online planning. Its interactional function includes a variety of functions, such as marking boundaries, introducing a different perspective, and has been discussed in detail by Kärkkäinen (2003: 105-82), who notes that '[i]n a majority of cases *I think* simply performs some routine (organizational) task in interaction, without conveying either clear uncertainty or certainty, or serving to soften or reassure' (Kärkkäinen 2003: 172). Essentially, the change *I think* is undergoing is one of becoming less conceptual and more procedural in meaning (cf. Blakemore 1990/1991).[11] This development is reflected in a change of positional preference (cf. the decrease of final uses), collocation patterns (cf. the increase of fillers and decrease of pauses), reduced use of the complementizer (at least in COLT), and a weakening of the link to a host construction, as attested by the increase in phrasal use and cases of unclear scope (as well as the occasional use as independent response). What does not fit in with a view of further grammaticalization, however, is the overall decrease in frequency found for *I think* (cf. Section 3). To account for it, it is necessary to take a wider look at *I think* which includes its formal variants. This will be the focus of the following section.

5. Variant forms

As argued by Van Bogaert (2009, 2010b), comment clauses such as *I think* should not be viewed in isolation but as part of a larger and expanding constructional network, which includes other variant forms, such as *I'm thinking, I would think, I should think, I was thinking, I thought, I do think, I should have thought*. These variant forms are part of a more schematic meso-construction (Traugott 2007) which has *I think* as its prototypical member, and to which the variant forms are linked by analogy. As the most frequent and most

---

[11] Very generally, conceptual meaning contributes to the propositional content whereas procedural meaning "helps the hearer draw the right kind of inferences from the embedded proposition" (Blakemore 1990/1991: 207).

entrenched of all comment clauses, *I think* functions as a template and pacemaker for the entire taxonomy, leading other variant forms (as well as other complement-taking mental predicates) to higher levels of entrenchment.

Although this is a very plausible mechanism, a constructional network view would seem to be at odds with the reduction in frequency noted for *I think* (cf. Section 3). It is possible, however, to account for the drop in frequency and thus maintain a constructional network view, if we incorporate a more fine-grained functional description of comment clauses. Instead of viewing transparent (i.e. comment clause) uses as a uniform functional category (which contrasts with non-transparent uses, as in Van Bogaert 2009, 2010b), it seems necessary to differentiate between different uses within the comment clause category, i.e. different functions of transparent *I think*. As demonstrated in Section 4, *I think* shows signs of further semantic bleaching and is increasingly used not so much as an epistemic marker that signals lack of speaker commitment, but as a more general pragmatic marker with important textual and interactional function. The reduction in frequency can be seen as a result of the increased semantic bleaching of *I think*, which makes it less effective as an epistemic marker. To compensate for this loss of modal meaning other forms are likely to be drafted in to take over the epistemic function of *I think*, which would explain the expansion of the taxonomy, as noted by Van Bogaert. Such an expansion would then be motivated not only by analogy but also by a functional necessity, that of filling a functional 'vacancy'. Likely candidates for recruitment as epistemic markers are *I'm thinking, I'd think* and *I just think / I'm just thinking,* owing to the notion of tentativeness signalled by the progressive, the modal, and the adverb respectively.[12]

Providing empirical evidence for the above hypothesis is, however, problematic because of the lack of sufficient spoken data from different periods of present-day English which would permit a diachronic perspective. DCPSE, for instance, contains only 6 instances of *I'm thinking* used as comment clauses (3 in LLC, 3 in ICE-GB). Nonetheless, comment clause use of *I'm thinking* is clearly attested in larger corpora such as the *British National Corpus* (BNC), whose spoken section yields a total of 193 instances, of which 41 (4.12 per 1 million words) qualify as comment clauses.[13] The only corpus available to date which is both

---

[12] Other formal variants such as *I do think* and *I thought* seem less suited for a replacement of *I think* as marker of reduced speaker commitment because of the emphatic (booster) function of the former and the potential past time reference of the latter.

[13] 29 initial, 7 +*that*, 4 phrasal, 1 final; as opposed to 8 +NP, 116 +PP, 5 intransitive uses, 23 'reporting thought' uses.

large enough and covers a longer time period is the *Corpus of Historical American English* (COHA).[14] Using COHA data is, of course, methodologically problematic as it covers exclusively written data (fiction, magazines, newspapers, non-fiction) and, unlike DCPSE, represents American English. Despite these differences COHA represents a valuable database for tracing the development of the less frequent variant forms of *I think* (with the appropriate methodological caveats).

For *I'm thinking* the COHA data show a steady increase from the 1930s (2.05 per million words) to the 2000s (8.19), which is in line with the attested rise of the progressive in general (e.g. Mair 2006: 89, Aarts, Close and Wallis 2010; Levin, this volume; Smith and Leech, this volume). It is necessary, however, to distinguish between comment clause uses and other uses of *I'm thinking*, such as those followed by a PP complement (e.g. *I'm thinking about you*), an NP complement (e.g. *That's what I'm thinking*), intransitive uses (e.g. *I'm thinking*), and the use as reporting verb (e.g. *I'm thinking, What the hell*). Comment clause uses, on the other hand, include clause-initial position with or without complementizer, medial position, phrasal scope and final position, as illustrated by the examples in (18) – (21) respectively.[15]

(18) It's an old idea. But **I'm thinking** I might know how to do it now. <COHA: 1994:FIC:A-P:Play:TrudyBlue>

(19) Revelation, **I'm thinking**, is a substance more real than sapphires <COHA:2000:FIC:C:FantasySciFi>

(20) I want a doctor to look at you just in case. **I'm thinking** a neurologist <COHA:1998:FIC:C:AntiochRev>

(21) 'A lie is the cruellest thing in the world, **I'm thinking**' Jason said with loud bitterness <COHA:1980:FIC:S:AnswerAsMan>

In all of these examples *I'm thinking* is interchangeable with *I think* without major change of meaning. There are also cases, however, which are indeterminate between a comment clause (epistemic) and a reporting verb reading (reporting the speaker's thoughts). In such unclear cases, where even the larger co(n)text does not provide any further clues for disambiguation,

---

[14] COHA covers the period from 1810 to 2009 and comprises a total of 406,232,024 words (7/09/2010). It is available online as a beta version at http://corpus.byu.edu/coha.

[15] There are two pro-form uses with *so*, e.g. *I'm thinking so just now* <COHA:1947:MAG:P:Atlantic>, in the corpus. They attest to comment clause use but have not been included in the data.

only those were classified as reporting verbs, where clear formal signals, such as orthographic marking (e.g. quotation marks, colon), interjections (e.g. *oh, sure*) introduce the 'host clause', or where the 'host clause' takes exclamative, interrogative, or imperative form.

As can be seen from Figure 3, the frequencies for both comment clause uses and other uses of *I'm thinking* remained relatively stable from the 1930s to the 1970s but have increased substantially after that, with a sharp rise in the 1990s. The relative frequencies within the set show a similar picture (cf. Figures A4 and A5 in the Appendix). This increase (which has tested as highly statistically significant as opposed to other uses) also brings about more varied comment clause use, in the form of phrasal scope, medial and *that*-complementizer uses, possibly by analogy to *I think* (cf. Table A2 in the Appendix).

Figure 3. *I'm thinking* in the *Corpus of Historical American English* (per 1 million words)[16]
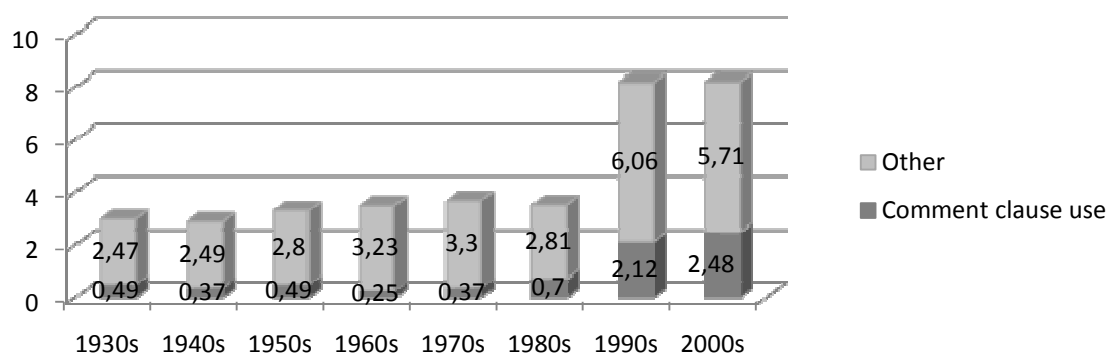
By comparison, the development of *I think* in COHA (illustrated in Figure 4) shows a slight but stable rise from the 1930s to the 1970s, which is more or less reversed from then on: the 1980s see a considerable dip, which is precisely the time when *I'm thinking* is beginning to take off. In the 2000s *I think* drops further, whereas *I'm thinking* peaks. This development is therefore consistent with a view that sees *I'm thinking* as taking over some of the epistemic function from *I think* (with a time lag of roughly a decade).

---

[16] At the time of access (3/08/2010) the subcorpora for the different decades had the following sizes: 1930s: 24,697,369; 1940s: 24,490,325; 1950s: 24,615,589; 1960s: 24,155,338; 1970s: 24,223,109; 1980s: 25,586,440; 1990s: 28,368,957; 2000: 31,503,769.

Figure 4. *I think* in the *Corpus of Historical American English* (per 1 million words)
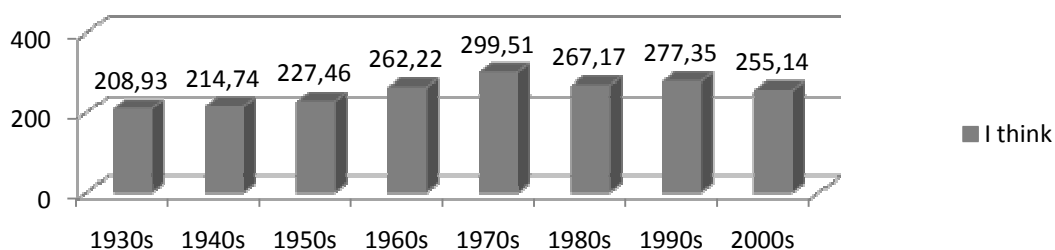
What about other variant forms that might qualify for the expression of epistemic meaning and hence as possible alternatives to *I think*? Another possible candidate is the modal form *I'd think* / *I would think* (cf. Van Bogaert 2009: 427). The corpus data, however, do not provide any clear evidence for such a development. In the COHA data for the 2000s (accessed 7/09/2010) only a minority of 12.5 per cent of all 48 instances of *I'd think* shows epistemic function which would be interchangeable with *I think*. An example is shown in (22).

(22) **I'd think** crazy guys are right down their alley
      <COHA:2007:FIC:W:DarkestEvening>

Similar results can be obtained for the uncontracted form *I would think*. Neither of the two forms, however, reveals a clear increase in frequency.

      The only other variant form that shows a clear tendency for increased modal use is *I just think*,[17] as in examples (23) and (24).

(23) As for Martin, Karl said: '**I just think** he's got to be one of our top players
      <COHA: 2005:NEWS:N:NYT>

(24) 'Ever see photos of the senator's oldest daughter?' Chase shook his head. 'Just a glimpse of the twins or a grainy background shot. Why?' Roman laughed. '**I just think** you'll like what you see' <COHA:2004:FIC:W:Heartbreaker>

As indicated by the results in Figure 5, the period from the 1930s to 2000s has seen a clear increase in the frequency of *I just think* together with its epistemic (comment clause) use (although it does not occur in final or medial position; cf. also Figure A5 in the Appendix for a similar relative change within the set). The development mirrors that of *I think* in that it

---

[17] Its progressive counterpart *I'm just thinking* yields only 5 occurrences in the 2000s.

shows substantial steps up in the 1980s and 2000s, which coincide with parallel drops in the frequency of *I think*.

Figure 5. *I just think* in the *Corpus of Historical American English* (per 1 million words)[18]
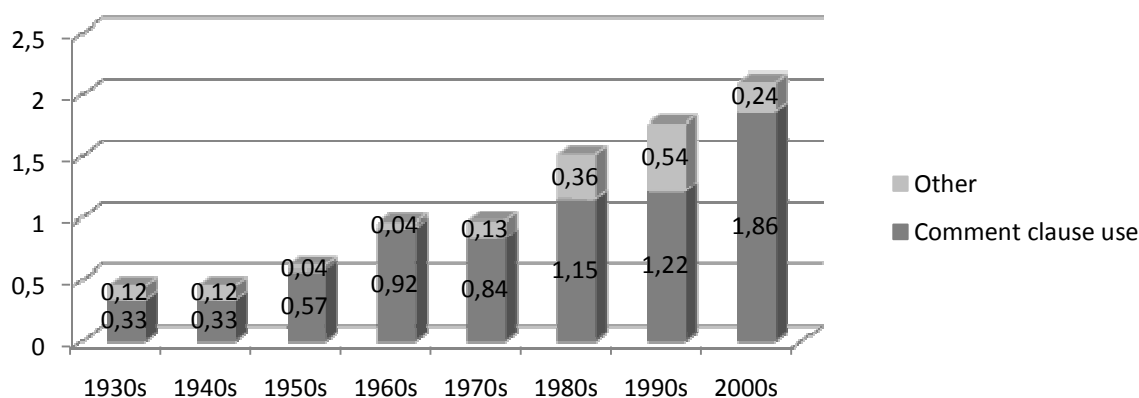

Other potential candidates for taking over from *I think* as marker of epistemic modality are variant forms involving different lexical predicates. The only form for which the corpus yields a substantial amount of data is *I'm guessing*, which has been noted to be used as a comment clause (cf. Van Bogaert 2010a). In COHA there is a clear indication that *I'm guessing* is used increasingly as a modal marker that is close in meaning to *I think*. As can be seen from Figure 6, there is a sharp rise of comment clause uses of *I'm guessing*, which includes initial, final, and phrasal uses, from the 1990s to the 2000s.[19] Typical examples are given in (25) – (27).

(25)   'My mom was a ballerina and a movie star.' # **I'm guessing** if your mom dies before you're born, she can have been anything <COHA:2000:FIC:SouthernRev>

(26)   'No, Milord, they started out that way, and **I'm guessing** they meant to at least make us wonder if they'd headed down it. <COHA:2005:FIC:WindridersOath>

(27)   'You've been a good boy all your life, **I'm guessing** <CHOA:2000:FIC:SouthernRev>

---

[18] The absolute figures are (comment clause use/other): 1930s: 8/3, 1940s: 8/3, 1950s: 14/1, 1960s: 22/1, 1970s: 20/3, 1980s: 29/9, 1990s: 34/15, 2000s: 55/7 (accessed 7/09/2010)..

[19] There is also one occurrence with *not*, which was not included in the count but attests to the comment clause use of the progressive variant: *Ever heard of me?* **I'm guessing not** <COHA:2006:FIC:ManoloMatrix>. *I'd guess* also shows an increase in numbers from 1930 but with substantial drops in the 1980s and 2000s and few overall occurrences (e.g. 22 in 2000s).

Figure 6. *I'm guessing* in the *Corpus of Historical American English* (per 1 million words)[20]
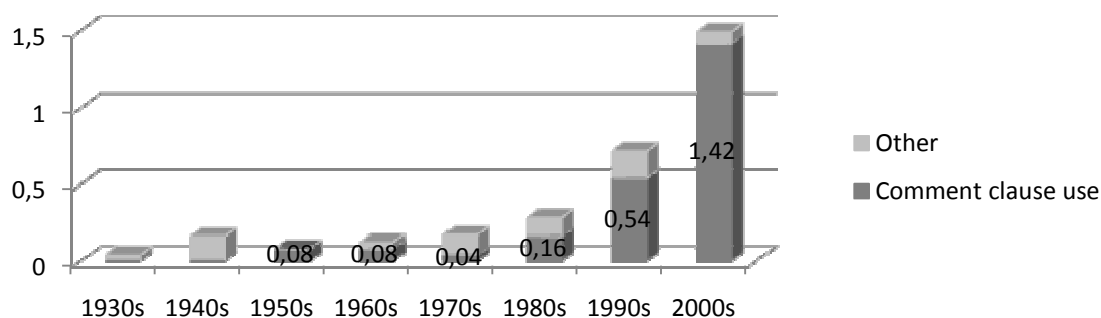
In contrast to *I'm guessing*, *I guess* shows a relatively stable frequency but with some fluctuation from the 1980s to 2000s, notably a decrease from 81.78 occurrences per 1 million words in the 1990s to 67.91 in the 2000s. Similar decreasing numbers can be found for *I suppose* (1930s: 78.04 > 2000s: 32.60 pmw), *I reckon* (1930s: 26.75 > 2000s: 3.08), *I expect* (1990s: 11.42 > 2000s: 6.56), and *I believe* (1930s: 59.34 > 2000s: 43.29). The rise of *I'm guessing* and *I'm thinking,* and *I just think* might therefore also be a compensation for the reduced frequency of other, invariant comment clauses (cf. COHA 12/09/2010).

In sum, the diachronic data derived from COHA for the period from 1930 to 2009 exhibit a substantial increase of some formal variants of *I think*, notably *I'm thinking* and *I just think*, as well as *I'm guessing*. This increase in frequency is due to their more frequent use as epistemic markers, signalling lack of speaker commitment, and runs parallel to the decrease of *I think*, which lends support to the hypothesis that variant forms take over some of the fading epistemic meaning of *I think*. Although the corpus covers written texts of American English and therefore does not directly compare with the spoken British English data in Section 4, it shows a clear trend which can be expected to be even more advanced in spoken language. It can be assumed, however, that regional varieties make different use of certain lexical predicates, with American English showing a greater preference for *I'm guessing* (parallel to *I guess*), and British English more readily adopting *I reckon* (cf. the COLT data in Section 4).

6. Comment clauses as constructions

The framework of Construction Grammar (cf. e.g. Goldberg 1995, 2006; Croft 2001, Croft and Cruse 2004; Östman and Fried 2005) has been shown to provide new insights for the

---

[20] The absolute figures are (comment clause use/other): 1930s: 0/1, 1940s: 0/4, 1950s: 2/0, 1960s: 2/1, 1970s: 1/3, 1980s: 4/3, 1990s: 15/5, 2000s: 42/2 (accessed on 7/09/2010).

description and development of constructions (for overviews cf. Fischer and Stefanowitsch 2007; Bergs and Diewald 2008; Trousdale and Gisborne 2008) and has recently been associated with the theory of grammaticalization (Trousdale 2008; Traugott 2007, 2008), which makes it particularly suited for comment clauses. Acknowledging the potential of a Construction Grammar approach, Brinton (2008: 255-56) briefly sketches out the historical development of epistemic parentheticals from a constructional point of view. A more detailed constructional account of comment clauses has been provided by Van Bogaert (2009, 2010b), who argues for an extension of the paradigm of grammaticalized epistemic predicate phrases to include variant forms such as *I would imagine, I'm guessing* (cf. Section 5 above).

These two constructional approaches provide useful accounts of the development of comment clauses, but still leave a few questions open, such as the persistent use of the *that*-complementizer, the ambiguous syntactic status of clause-initial comment clauses, and further degrees of grammaticalization / bleaching (as discussed in Section 4). To be able to address these questions it is necessary to place comment clauses in a constructional network, as demonstrated by Brinton and Van Bogaert, but to cast the net somewhat wider and include not only the comment clause construction in isolation but also related constructions of different degrees of schematicity. This section sketches out such a constructional network which complements the two existing models.

From a Construction Grammar perspective constructions are independent, but not isolated entities (cf. e.g. Fried and Östman 2004: 12, Croft and Cruse 2004: 262-64). They are linked with other, related constructions of different levels of schematicity in a larger taxonomic network of constructions. The nature of these links is still a matter of some discussion (cf. Croft and Cruse 2004: ch. 10), but can be assumed to include analogical relationships, i.e. based on the perceived similarity of two entities (cf. Van Bogaert 2009, 2010b, Traugott and Trousdale 2010). For comment clauses it is possible to identify analogical links to two different schemas, the 'matrix clause – object clause' and the 'sentence adverbial schema'.

The 'Matrix clause – object clause construction' codes speaker comment as a syntactically governing constituent. This syntactic foregrounding reflects the typical information structure of matrix + *that*-complement clause structures, where the subordinate clause has been noted to 'harbour, rather consistently, presupposed clauses' (Givón 1989: 132; cf. also Sadock 1984, Mackenzie 1984 for similar observations). This seems to be true especially with complements of cognition verbs (e.g. *I knew that she was there* cf. Givón 1989: 132). The pattern matrix clause + object clause can be assumed to represent a highly

24

dominant schema owing to its high token frequency (according to Greenbaum, Nelson and Weitzman 1996: 88-9 the most frequent type of clausal complementation) and owing to its taxonomic link with the more schematic Transitive construction ([Sbj] [TrnVerb] [ObjNP]; cf. Trousdale 2008 on the dominant role of the Transitive construction).

The 'Sentence adverbial construction', on the other hand, can take various forms (e.g. finite and non-finite clauses, PPs, NPs), with single adverbs being the most frequent representative, especially in spoken language (cf. Biber *et al*. 1999: 862). Adverbs, in turn, are related functionally and historically to another category, viz. that of pragmatic markers (e.g. uses of *indeed, only, actually*), for which adverbs represent the historical source out of which they have developed either directly, via sentence adverbials, or via conjunctions (e.g. Brinton 2008: 246, Traugott 1995a: 13). The category of sentence adverbials shares with pragmatic markers not only a similar function, viz. wide-scope evaluation (of a proposition or upcoming text respectively), but also similar coding as syntactically backgrounded: both are in a non-governing relationship to their 'host construction', which iconically reflects their secondary status.

Since constructions are form-meaning pairings (cf. e.g. Croft 2001: 18, Croft and Cruse 2004: 258), the links of comment clauses with their parent constructions will be of both a formal and a functional kind. Functionally, both constructions serve as repositories for speaker comment (stance). Given the reduced semantic meaning of comment clauses (cf. epistemic use), however, they would seem to be functionally more prone to coding as secondary comments, i.e. as sentence adverbials and ultimately (in their semantically reduced, pragmatically enriched form) as pragmatic markers. Formally, comment clauses display varying links. Their subject-predicate form is, of course, strongly reminiscent of main clauses and, together with clause-initial position (the typical position of main clauses), can be expected to activate the matrix-complement schema. With non-initial comment clauses the feature of positional flexibility may be more prominent and responsible for a strong link to 'coding as secondary comments', i.e. sentence adverbials, but still with some analogic link to matrix clauses, owing to their clausal form and potential for initial position.

These network relations of the comment clause construction ('cxn') can be represented in diagram form as in Figure 7. Note that the two 'parent constructions', which serve as analogical models, are also reflected in the two types of pro-forms found with comment clauses: viz. *so* (as in *I think / believe / suppose so*) and *it/that* (as in *I believe / suspect it*). The former is an instantiation of the adverbial link to a proposition, the latter is indicative of a governing (matrix clause) relationship over the following complement.

Figure 7. Taxonomic network for comment clauses
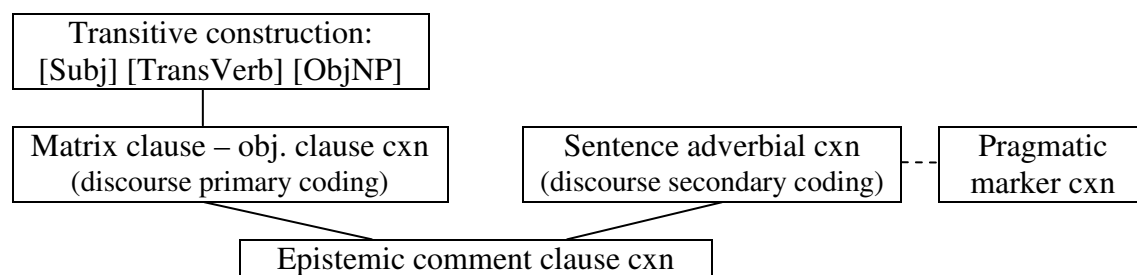
Positioning comment clauses in such a larger constructional network can help to account for their diachronic development with regard to both formal and functional features. The network model thus accommodates the functional development from epistemic to textual marker identified for *I think* in Section 4. The predominantly functional link of comment clauses to sentence adverbials (i.e. 'coding as discourse secondary') facilitates further grammaticalization (bleaching) along these lines owing to the already established pathway from (sentence) adverbials to pragmatic markers. Pragmatic marker function is simply a further possible development from adverbial usage (as evidenced by *you know* and *I mean*).

The model can also explain why epistemic comment clauses were associated with a *that*-complementizer. Despite their origin as independent clauses (cf. Section 2), comment clauses have come to be analogically construed by language users as instantiations of matrix clauses.[21] This is mainly the result of shared formal features, more precisely their clausal form and ability to occur in clause-initial position. Functionally, matrix clauses resemble comment clauses, too, since they also express speaker comment, even though it is typically discourse prominent. In initial position comment clauses have therefore inherited matrix clause features leading to complementizer use. With increased grammaticalization, this associative link with matrix clauses has, of course, considerably weakened. Nonetheless, the *that*-complementizer continues to be used on a low but fairly constant frequency level with even highly grammaticalized comment clauses. This retention of *that* can be attributed to grammatical persistence (Torres Cacoullos and Walker 2009: 34), which in turn can be motivated by a constructional network link to the 'Matrix clause – object clause' schema.

The taxonomic link to the Transitive / Matrix clause construction can also explain the wide-ranging differences for the use of *that* with different lexical predicates, such as 6.52 per cent for *I suppose* and 50 per cent[22] for *I understand* in the spoken part of ICE-GB: different verbs are cognitively associated with the transitivity schema to different degrees (which may

---

[21] On the importance of analogy for language change cf. e.g. Fischer 2007, Traugott & Trousdale 2010: 35-39.

[22] These figures are taken from Van Bogaert (2009: 378).

depend on the semantic content/weight of the verb) and therefore activate the matrix clause link, which triggers complementizer use, to varying extents. This closer association of some cognitive verbs with the transitive construction can be measured by their ability or a greater tendency to take direct object NPs (e.g. *I believe that/your story* vs. **I suppose that/a problem*). As illustrated by the figures in Table 3, the comment clauses most frequently associated with a *that*-complementizer in initial position are *I understand, I believe, I realise*. It is these verbs, *understand, believe, realise*, which also have the highest proportion of direct object NPs in ICE-GB (e.g. *He didn't understand the situation*). Conversely, verbs which in their comment clause use rarely take a *that*-complementizer (viz. *I reckon, I expect, I suppose, I think, I guess*) show a weaker association with NP objects.

Table 3. Frequencies of initial comment clauses +*that* (in spoken ICE-GB; from Van Bogaert 2009: 378) and frequencies of verbs and their transitive use with object NP (in total ICE-GB)

Initial comment clause	+ *that* (in spoken ICE-GB)	Verb frequency (in ICE-GB)	+ direct object NP (in ICE-GB)
I understand	50.00%	*understand* (187)	46.52% (87)
I believe	46.15%	*believe* (295)	15.93% (47)
I realise	33.33%	*realise* (87)	21.84% (19)
I guess	9.09%	*guess* (62)	8.06 % (5)
I think	8.78%	*think* (2,563)	0.31% (8)
I suppose	6.52%	*suppose* (237)	0.00%
I expect	0.00%	*expect* (124)	11.29% (14)
I reckon	0.00%	*reckon* (13)	0.00%

7. Concluding remarks

Overall, the DCPSE data show a noticeable decrease of comment clauses in British English between 1960 and 1990, with *I guess* being the only exception (cf. Section 3). This downward trend is also confirmed by data from COLT, although this London variety exhibits a particular preference for *I reckon*. How can we explain this development? The discussion of *I think*, the most frequent and most prototypical of all comment clauses (in Section 4), suggests that, at least in the case of *I think*, the reduction in numbers can be linked to an ongoing functional change as a result of further grammaticalization. The corpus data from DCPSE provide various indications for increased semantic erosion of *I think* and thus weakening of its epistemic meaning: (i) reduction of clause-final *I think*, the position most typically associated with speaker comment (4.1), (ii) weakening of the semantic-pragmatic bond with the host construction as evidenced in a diffusion of scope, which may cover not only clausal but

increasingly also phrasal constituents, or may even be unclear, as in the 'textual' uses of *I think* (4.3), (iii) an increase in the co-occurrence of discourse markers, suggesting similarity in function, and a decrease in the co-occurrence of pauses, suggesting effective use as a filling device (4.4), and finally (iv) an openness to new pragmatic uses such as approximative function of phrasal scope *I think* (4.3) or even use as independent response (4.5). Interestingly, the DCPSE data do not show any increase in *that*-omission, which is generally seen as a concomitant of increased grammaticalization (4.2). This can be explained, however, by grammatical persistence as a result of strong taxonomic ties of the comment clause construction with a superordinate Transitive construction / 'Matrix clause – object clause' construction (cf. Section 6). There are indications that this persistence may be waning, though, in the speech of teenagers, as suggested by the results from COLT.

I think thus shows signs of erosion of its typical comment clause meaning and is increasingly used not so much as an epistemic qualifier of some host clause, typically expressing lack of certainty or doubt, but as a general pragmatic marker with important textual or interactional function. These include stalling or filler functions to provide time for online planning as well as discourse-organizational functions, such as discourse linking/thematic highlighting (cf. Ziv 2002), marking boundaries or introducing a different perspective (cf. Kärkkäinen 2003). Such a development is not really surprising, as it continues the path from propositional to interpersonal meaning (cf. Section 2). The current development is just a step further down the cline of interpersonal meaning, reducing the traces of propositional meaning still present in epistemic evaluation. Essentially, the change of *I think* is one of becoming even less conceptual and more procedural in meaning (e.g. Blakemore 1990/1991), i.e. acting as a general processing instruction with a variety of potential pragmatic functions depending on contextually invited inferences.

The overall reduction in frequency of *I think* may be a result of its changing discourse function, which makes it increasingly unattractive as an epistemic marker. To compensate for its fading modal use it seems that other forms are being recruited to take over this function. The COHA data, which show a similar reduction of *I think* as DCPSE, indicate a substantial increase of the formal variants *I'm thinking, I just think* and *I'm guessing* for the period 1930 to 2009, both overall and in their uses as epistemic comment clauses (cf. Section 5). This increase is most pronounced from the 1980s, the decade which saw the first reduction of *I think* since the 1930s. The temporal development of the forms investigated thus suggests a shift of epistemic function from *I think* to certain variant forms. Given the conservative nature

of written texts, it may be hypothesized that this development is even more advanced in spoken language.

To account for the development of comment clauses a Construction Grammar model seems most appropriate. Such a model has been outlined by Van Bogaert (2009, 2010b), but, as argued in Section 6, it needs to be complemented by taxonomic links to related constructions, viz. the Transitive / 'Matrix clause – object clause' construction and the 'Sentence adverbial construction'. By placing comment clauses in such a larger constructional framework it is possible to account for formal and functional characteristics of their development, such as the advance of *I think* from an epistemic to a general pragmatic marker and the use of the *that*-complementizer.

Appendix

Confidence intervals in the following graphs are computed using the Newcombe-Wilson interval method with *p*(error)<0.05 (i.e. at a 95% confidence level). See Appendix 1 in Aarts *et al.*, this volume for more explanation. Where the confidence interval does not cross the zero axis the change is deemed to be statistically significant.

Figure A1.1. Per million word change over time (relative to LLC = 100%) for individual comment clauses and the total set.
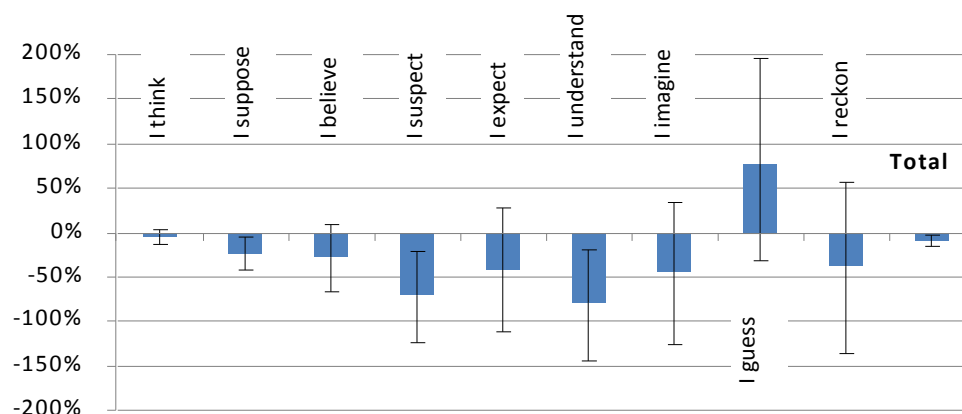

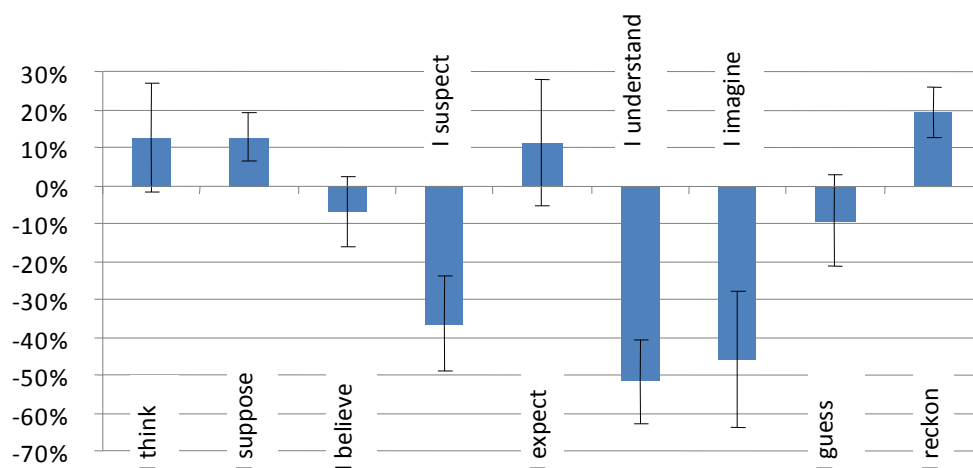Figure A1.2. Variation within the set of comment clauses: relative to the total set.

Figure A2. Relative change within the set of comment clauses (comparing LLC, ICE-GB, COLT), also with 95% Newcombe-Wilson confidence intervals.
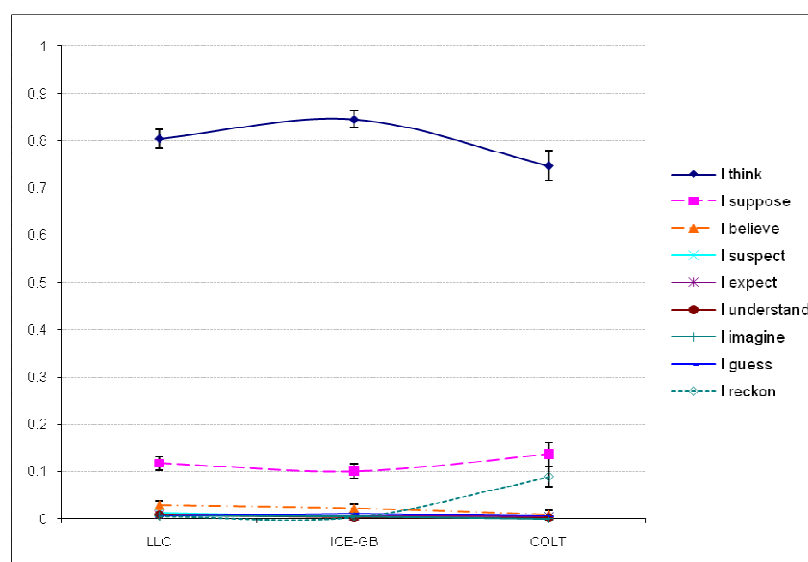

Table A1. Different uses of comment clause *I think* in ICE-GB (DCPSE) text category B (informal face-to-face) and COLT (normalised per 1 million words)

I think	ICE-GB cat. B (185,537 words)		COLT (500,000 words)		Change in frequency		
	raw	pmw	raw	pmw	*%*	χ^2 partial	X vs ¬X 2 x 2 χ^2
Initial	356	1,918.7	408	816.0	-57.47	0.24	**5.60 sig<0.02**
+*that*	22	118.5	5	10.0	-91.56	**10.32 sig<0.001**	**10.62 sig<0.001**
Medial	37	199.4	23	46.0	-76.93	**4.03 sig<0.04**	**4.28 sig<0.04**
Phrasal	35	188.6	41	82.0	-56.52	0.22	0.23
Final	43	231.7	42	84.0	-63.75	0.12	0.13
Total	493	2,657,1	519	1,038.0	-85.30	**17.34 sig<0.01**	

Figure A3. Relative change of different uses of comment clause *I think* within the set in text category B 'informal face-to-face' of LLC and ICE-GB, and in COLT.
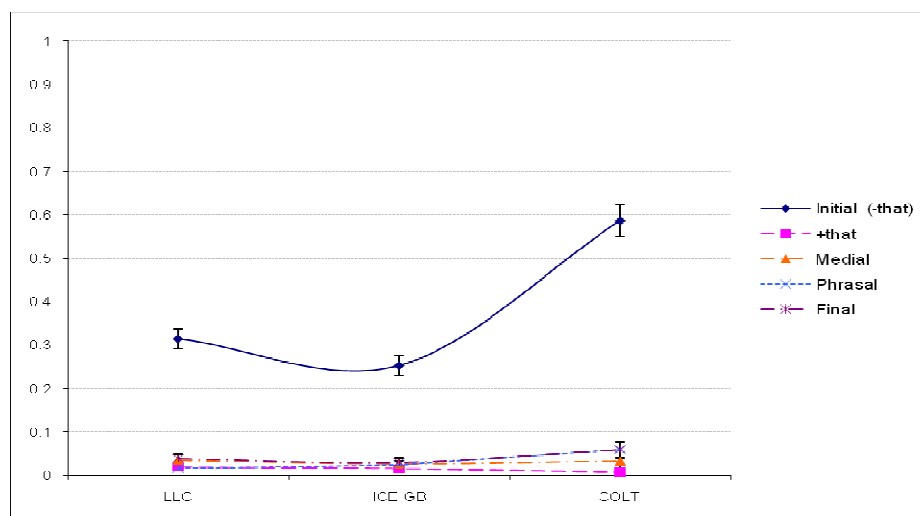

Table A2. *I'm thinking* in the *Corpus of Historical American English*

	30s	40s	50s	60s	70s	80s	90s	2000s
Initial	9	6	6	4	6	7	47	55
+*that*	-	-	2	2	-	-	6	9
Medial	-	-	1	-	-	4	-	1
Phrasal	-	-	-	-	-	-	2	4
Final	3	3	3	-	3	7	5	8
Total	12	9	12	6	9	18	60	78
Other uses	61	61	69	78	80	72	172	180

$\chi^2 = 38.43$, sign. < 0.001

Figure A4. Relative change of comment clause uses of *I'm thinking* compared to other uses of *I'm thinking* in the *Corpus of Historical American English*.
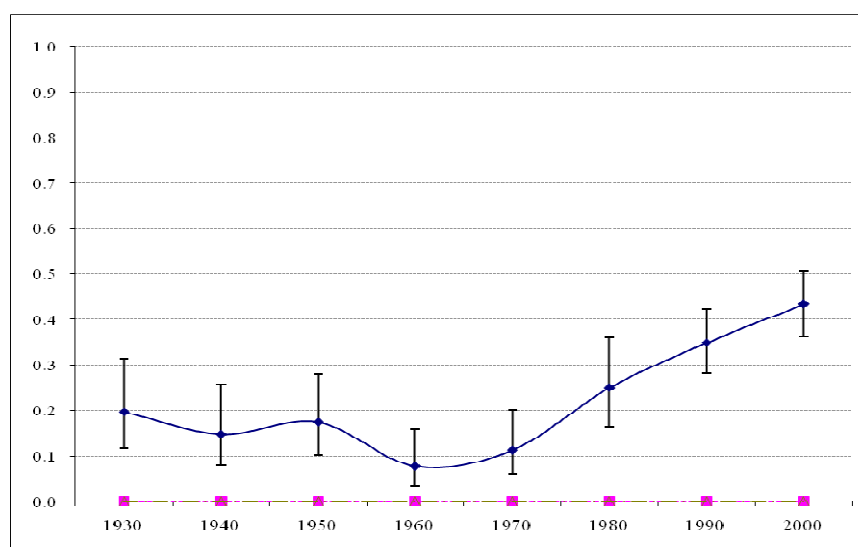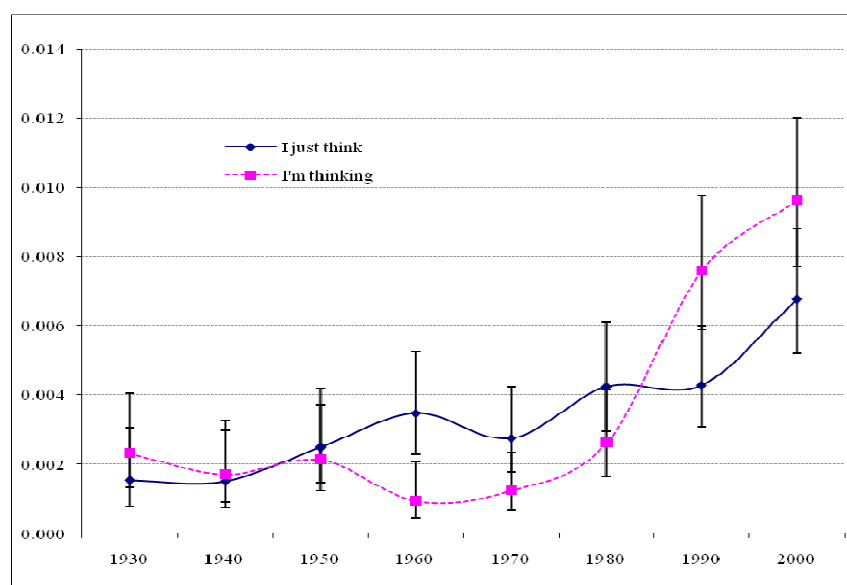
Figure A5. Relative change of comment clause uses of *I just think* and *I'm thinking* within the set (*I think, I just think, I'm thinking*) in the *Corpus of Historical American English*.

References

Aarts, Bas, Joanne Close and Sean A. Wallis 2010 Recent changes in the use of the progressive construction in English. In: B. Cappelle and N. Wada (eds.) *Distinctions in English grammar, offered to Renaat Declerck*. Tokyo: Kaitakusha. 148-167.

Aijmer, Karin. 1997. *I think* – an English modal particle. In Toril Swan & Olaf Jansen Westwik (eds.). *Modality in Germanic languages: Historical and comparative perspectives*. Berlin: Mouton de Gruyter, 1-47.

Bergs, Alexander & Gabriele Diewald. 2008. Introduction: constructions and language change. In Alexander Bergs & Gabriele Diewald (eds.). *Constructions and language change*. Berlin, New York: Mouton de Gruyter, 1-21.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. London: Longman.

Blakemore, Diane. 1990/91. Performatives and parentheticals. *Proceedings of the Aristotelian Society* 91, 197-213.

Boye, Kasper & Peter Harder. 2007. Complement-taking predicates: Usage and linguistic structure. *Studies in Language* 31, 569-606.

Brinton, Laurel J. 1996. *Pragmatic markers in English: Grammaticalization and discourse functions*. Berlin: Mouton de Gruyter.

Brinton, Laurel J. 2008. *The comment clause in English: Syntactic origins and pragmatic development*. Cambridge: Cambridge University Press.

Brinton, Laurel J. & Elizabeth Closs Traugott. 2005. *Lexicalization and language change*. Cambridge: Cambridge University Press.

Bromhead, Helen. 2006. *The reign of truth and faith. Epistemic expressions in 16th and 17th century English*. Berlin, New York: Mouton de Gruyter.

Conrad, Susan & Douglas Biber. 2000. Adverbial marking of stance in speech and writing. In Susan Hunston & Geoff Thompson (eds.). *Evaluation in text*. Oxford: Oxford University Press, 56-73.

Croft, William & D. Alan Cruse. 2004. *Cognitive linguistics*. Cambridge: Cambridge University Press.

Croft, William. 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Dehé, Nicole & Anne Wichmann. 2010. The multifunctionality of epistemic parentheticals in discourse: prosodic cues to the semantic-pragmatic boundary. *Functions of Language* 17 (1), 1-28.

Diessel, Holger & Michael Tomasello. 2001. The acquisition of finite complement clauses in English: A corpus-based analysis. *Cognitive Linguistics* 12, 97-141.

Erman, Britt & Ulla-Britt Kotsinas. 1993. Pragmaticalization: the case of *ba* and *you know*. *Studier I modern språkvetenskap*, New series 10, Acta Universitatis Stockholmiensis. Stockholm: Almqvist & Wiksell, 76-93.

Finegan, Edward & Douglas Biber. 1995. *That* and zero complementisers in Late Modern English: exploring ARCHER from 1650-1990. In Bas Aarts & Charles M. Meyer (eds.). *The verb in contemporary English. Theory and description*. Cambridge: Cambridge University Press, 241-57.

Fischer, Kerstin & Anatol Stefanowitsch. 2007. Konstruktionsgrammatik: ein Überblick. In Kerstin Fischer & Anatol Stefanowitsch (eds.). *Konstruktionsgrammatik. Von der Anwendung zur Theorie*. Tübingen: Stauffenburg, 3-17.

Fischer, Olga. 2007. *Morphosyntactic change. Functional and formal perspectives*. Oxford: Oxford University Press.

Fried, Mirjam & Jan-Ola Östman. 2004. Construction Grammar: a thumbnail sketch. In Mirjam Fried, Jan-Ola Östman (eds.). *Construction Grammar in a cross-language perspective*. Amsterdam, Philadelphia: Benjamins, 11-86.

Givón, Talmy. 1989. *Mind, code and context*. Hillsdale NJ: Lawrence Erlbaum.

Goldberg, Adele. 1995. *Constructions: A construction grammar account of argument structure*. Chicago: University of Chicago Press.

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

Greenbaum, Sidney, Gerald Nelson & Michael Weitzman. 1996. Complement clauses in English. In Jenny Thomas & Mick Short (eds.). *Using corpora for language research*. London: Longman, 76-91.

Heine, Bernd, Ulrike Claudi & Friederike Hünnemeyer. 1991. *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press.

Heine, Bernd. 2003. Grammaticalization. In Bryan D. Joseph & Richard D. Janda (eds.). *The handbook of historical linguistics*. Malden, MA and Oxford: Blackwell, 575-601.

Himmelmann, Nikolaus P. 2004. Lexicalization and grammaticalization: opposite or orthogonal? In Walter Bisang, Nikolaus P. Himmelmann & Björn Wiemer (eds.). *What makes grammaticalization? A look from its fringes and its components*. Berlin and New York: Mouton de Gruyter, 21-42.

Hooper, Joan B. 1975. On assertive predicates. In John P. Kimball (ed.). *Syntax and Semantics*, vol. 4. New York: Academic Press, 91-124.

Hopper, Paul J. 1991. On some principles of grammaticalization. In Elizabeth Closs Traugott & Bernd Heine (eds.). *Approaches to grammaticalization*, vol. 1. Amsterdam: Benjamins, 17-35.

Ifantidou, Elly 2001. *Evidentials and relevance.* Amsterdam: Benjamins.

Kaltenböck, Gunther. 2008. Prosody and function of English comment clauses. *Folia Linguistica* 42 (1), 83-134.

Kaltenböck, Gunther. 2009. English comment clauses: position, prosody, and scope. *Arbeiten aus Anglistik und Amerikanistik* 34 (1), 49-75.

Kaltenböck, Gunther. 2010. Pragmatic functions of parenthetical *I think*. In Gunther Kaltenböck, Wiltrud Mihatsch & Stefan Schneider (eds.). *New approaches to hedging*. Amsterdam: Elsevier, 243-272.

Kaltenböck, Gunther. 2011 forthc. Linguistic structure and use: explaining diverging evidence. The case of clause-initial *I think*. In Doris Schönefeld (ed.). *Converging evidence – discussing and extending the methodological tool-kit of the linguist*.

Kärkkäinen, Elise. 2003. *Epistemic stance in English conversation: A description of its interactional functions, with a focus on* I think. Amsterdam: Benjamins.

Kärkkäinen, Elise. 2007. The role of *I guess* in conversational stance taking. In Robert Englebretson (ed.). *Stancetaking in discourse: Subjectivity, evaluation, interaction*. Amsterdam: Benjamins, 183-219.

Kärkkäinen, Elise. 2010. Position and scope of epistemic phrases in planned and unplanned American English. In Gunther Kaltenböck, Wiltrud Mihatsch & Stefan Schneider (eds.). *New approaches to hedging*. Amsterdam: Elsevier, 207-241.

Kearns, Kate. 2007. Epistemic verbs and zero complementizer. *English Language and Linguistics* 11, 475-505.

Leech, Geoffrey, Marianne Hundt, Christian Mair & Nicolas Smith. 2009. *Change in contemporary English. A grammatical study*. Cambridge: Cambridge University Press.

Lehmann, Christian. 1995 [1982]. *Thoughts on grammaticalization* [revised and expanded version]. München and Newcastle: Lincom Europa.

Mackenzie, J. Lachlan. 1984 Communicative functions of subordination. In J.L. Mackenzie & H. Wekker (eds.). *English language research: The Dutch contribution I*. Amsterdam: Free University Press, 67-84.

Mair, Christian. 2006. *Twentieth-century English: history, variation, and standardization*. Cambridge: Cambridge University Press.

Mindt, Ilka. 2003. Is *I think* a discourse marker. In Ewald Mengel, Hans-Jörg Schmidt, Michael Steppat (eds.). *Proceedings of the Conference of the German Association of University Teachers of English* 24. Trier: Wissenschaftlicher Verlag, 473-83.

Östman, Jan-Ola & Mirjam Fried (eds.). 2005. *Construction grammars: Cognitive grounding and theoretical extensions*. Amsterdam: Benjamins.

Palander-Collin, Minna. 1999. *Grammaticalization and social embedding. I think and methinks in Middle and Early Modern English*. Helsinki: Société Néophilologique.

Posner, Roland. 1973. Redekommentierung. *Funk-Kolleg Sprache*, vol. 2. Frankfurt: Fischer, 124-133.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.

Rissanen, Matti. 1991. On the history of *that*/zero as object clause links in English. In Karin Aijmer & Bengt Altenberg (eds.) *English corpus linguistics*. London: Longman, 272-89.

Sadock, Jerrold M. 1984. The pragmatics of subordination. In W. de Geest & Y. Putseys (eds.). *Sentential Complementation*. Dordrecht: Foris. 205-213.

Schneider, Stefan. 2007. *Reduced parenthetical clauses as mitigators: A corpus study of spoken French, Italian and Spanish*. Amsterdam: Benjamins.

Simon-Vandenbergen, Anne-Marie. 2000. The functions of *I think* in political discourse. *Journal of Applied Linguistics* 10, 41-63.

Stefanowitsch, Anatol. 2007. Konstruktionsgrammatik und Korpuslinguistik. In Kerstin Fischer & Anatol Stefanowitsch (eds.). *Konstruktionsgrammatik. Von der Anwendung zur Theorie*. Tübingen: Stauffenburg, 151-176.

Stenström, Anna-Brita. 1995. Some remarks on comment clauses. In Bas Aarts & Charles F. Meyer (eds.). *The verb in contemporary English*. Cambridge: Cambridge University Press, 290-299.

Stenström, Anna-Brita, Gisle Andersen & Ingrid Kristine Hasund. 2002. *Trends in Teenage Language. Corpus compilation, analysis and findings*. Amsterdam: John Benjamins.

Tabor, Whitney & Elizabeth Closs Traugott. 1998. Structural scope expansion and grammaticalization. In Ramat Giacolone & Paul J. Hopper (eds.). *The limits of grammaticalization*. Amsterdam, Philadelphia: John Benjamins, 229-272.

Thompson, Sandra A. & Anthony Mulac. 1991. A quantitative perspective on the grammaticalization of epistemic parentheticals in English. In Elizabeth Closs Traugott & Bernd Heine (eds.). *Approaches to grammaticalization*, vol. 2. Amsterdam: Benjamins, 313-39.

Thompson, Sandra A. 2002. 'Object complements' and conversation: Towards a realistic account. *Studies in Language* 26, 125-64.

Torres Cacoullos, Rena & James A. Walker. 2009. On the persistence of grammar in discourse formulas: A variationist study of *that*. *Linguistics* 47, 1-43.

Traugott, Elizabeth Closs. 1982. From propositional to textual and expressive meanings: Some semantic-pragmatic aspects of grammaticalization. In Winfred P. Lehmann & Yakov Malkiel (eds.). *Perspectives on historical linguistics*. Amsterdam: Benjamins, 245-71.

Traugott, Elizabeth Closs. 1988. Pragmatic strengthening and grammaticalization. *Berkeley Linguistics Society* 14, 406-16.

Traugott, Elizabeth Closs. 1995a. The role of the development of discourse markers in a theory of grammaticalization (www.stanford.edu/~traugott/ect-papersonline.html).

Traugott, Elizabeth Closs. 1995b. Subjectification in grammaticalization. In Dieter Stein & Susan Wright (eds.). *Subjectivity and subjectivisation: Linguistic perspectives*. Cambridge: Cambridge University Press, 31-54.

Traugott, Elizabeth Closs. 2007. The concepts of constructional mismatch and type-shifting from the perspective of grammaticalization. *Cognitive Linguistics* 18, 523-57.

Traugott, Elizabeth Closs. 2008. Grammaticalization, constructions and the incremental development of language: Suggestions from the development of degree modifiers in English. In Regine Eckhardt, Gerhard Jäger & Tonjes Veenstra (eds.). *Variation, selection, development: Probing the evolutionary model of language change*. Berlin: Mouton de Gruyter, 219-50.

Traugott, Elizabeth Closs & Richard B. Dasher. 2002. *Regularity in semantic change*. Cambridge: Cambridge University Press.

Traugott, Elizabeth Closs & Graeme Trousdale. 2010. Gradience, gradualness and grammaticalization: How do they intersect? In Elizabeth Closs Traugott & Graeme Trousdale (eds.). *Gradience, gradualness and grammaticalization*. Amsterdam: Benjamins, 19-44.

Trousdale, Graeme. 2008. Words and constructions in grammaticalization: The end of the English impersonal construction. In Donka Minkova & Susan Fitzmaurice (eds.). *Studies in the history of the English language IV: Empirical and analytical advances in the study of English language change*. Berlin: Mouton de Gruyter, 301-26.

Trousdale, Graeme & Nikolas Gisborne (eds.). 2008. *Constructional approaches to English grammar*. Berlin: Mouton de Gruyter.

Urmson, James O. 1952. Parenthetical verbs. *Mind* 61, 480-96.

Van Bogaert, Julie. 2006. *I guess, I suppose* and *I believe* as pragmatic markers: Grammaticalization and functions. *Belgian Journal of English Language and Literatures (New Series)* 4, 129-49.

Van Bogaert, Julie. 2009. The grammar of complement-taking mental predicate constructions in present-day spoken British English. Doctoral dissertation University of Gent.

Van Bogaert, Julie. 2010a. 'I'm guessing it looks probably like a sea monkey right now': The recent emergence of comment clauses in the progressive. Paper held at the ICAME 31, 26-30 May 2010, Giessen.

Van Bogaert, Julie. 2010b. A constructional taxonomy of *I think* and related expressions: accounting for the variability of complement-taking mental predicates. *English Language and Linguistics*.

Wallis, Sean. 2009. Binominal distributions, probabililty and Wilson's confidence interval. London: Survey of English Usage. www.ucl.ac.uk/english-usage/staff/sean/resources/binominalpoisson.pdf

Wichmann, Anne. 2001. Spoken parentheticals. In Karin Aijmer (ed.). *A wealth of English. Studies in honour of Göran Kjellmer*. Göteborg: Acta Universitatis Gothoburgiensis, 177-193.

Wierzbicka, Anna. 2006. *English: Meaning and culture*. Oxford: Oxford University Press.

Ziv, Yael. 2002. This, I believe, is a processing instruction: discourse linking via parentheticals. http://micro5.mscc.huji.ac.il/~english/IATL/18/Ziv.pdf (Dec. 2006).