

INITIAL *I THINK*: MAIN OR COMMENT CLAUSE?

Gunther Kaltenböck

Abstract

This paper explores the use of 200 occurrences of clause-initial *I think* in a corpus of spoken English with a view to establishing whether they are best classified as main or comment clauses. It investigates two formal cues for signalling prominence of *I think* and hence a possible hierarchical difference between the two clauses: (i) the presence or absence of the *that*-complementizer as an explicit marker of syntactic subordination and (ii) prosodic prominence. The corpus data show that a difference on the structural level, i.e. *that* vs. zero, does not correspond with different prosodic behaviour. Both constructional types exhibit a similar distribution of the three prosodic patterns identified: they are both most frequently realised as heads, less frequently as pre-heads, and only rarely with a separate nuclear accent. From a cognitive-functional perspective, which associates superordinate status with relative prominence, initial *I think* therefore only rarely qualifies for main clause status. Moreover, the corpus data suggest that in spoken language the *that*-complementizer is not so much used as a marker of subordination but rather as a filler used to give weight to *I think* or for rhythmical purposes.

Key words

prosody, parenthetical, comment clause, main clause, corpus analysis, spoken language, *that*-complementizer

1 Introduction

Clause-initial *I think*, *I believe*, *I suspect* and similar complement-taking predicates have been a matter of some discussion because of their unclear syntactic status. Are they main clauses, which syntactically govern a complement clause, or comment clauses, i.e. parentheticals, which are in a syntactically supplementary relationship to the following clause? This question is particularly difficult to answer for cases where the *that*-complementizer has been omitted. Compare example (1).

(1) *I think (that) John has come back from London*

Various views have been expressed on the status of such initial clauses with and without *that* complementizer. They are either taken to be parenthetical (e.g. Kärkkäinen 2003, Krusinga 1932: 486, Ross 1973, Thompson 2002, Thompson & Mulac 1991), matrix clauses (e.g. Peterson 1999: 236, Stenström 1995: e.g.

293, 296, Svensson 1976: 375), or ambiguous, i.e. allowing interpretation as both matrix clause and parenthetical clause depending on context and type of ‘matrix’ predicate (e.g. Aijmer 1972: 46, Biber et al. 1999: 197, Huddleston & Pullum 2002: 896, Quirk et al. 1985: 1113, Urmson 1952: 481).

The present paper tries to shed some light on this question by examining 200 occurrences of initial *I think* in a corpus of spoken English. The advantage of spoken material is that it provides an important cue to the relative prominence of such initial clauses, viz. prosody. The other formal signal for a hierarchical difference between the two clauses is a syntactic one, viz. the presence or absence of the *that*-complementizer as an explicit marker of syntactic subordination.¹ In the present study I will take a closer look at this syntax-prosody interface and investigate to what extent these two formal signals provide an indication of the prominence of initial *I think* in relation to the following clause.

From a cognitive-functional perspective, relative prominence is the key to distinguishing between main and subordinate clauses. According to Langacker (1991: 436ff) a subordinate clause is defined as “one whose profile is overridden by that of the main clause ... *I know she left* designates the process of knowing, not of leaving”, with “profile” referring to the “relative prominence accorded to various substructures” (ibid.: 4). In spoken language this prominence can be expected to be expressed by prosody, which in turn reflects the informational salience in a given context: new or salient information will normally be made more prominent than given or presupposed information.

The corpus investigated is the spoken part of ICE-GB, the British component of the *International Corpus of English* (Nelson et al. 2002), which yields a total of 1,138 instances of *I think*, as detailed in Table 1.

Text type (number of words)	- <i>that</i>	+ <i>that</i>	Total
Private dialogue s1a (205,627)	22.7 (466)	1.2 (25)	23.9 (491)
Public dialogue s1b (171,062)	25.4 (434)	3.0 (52)	28.4 (486)
Public monologue s2a (152,829)	5.2 (80)	0.7 (11)	5.9 (91)
Scripted speech s2b (108,164)	5.3 (57)	1.2 (13)	6.5 (70)
Total	16.2 (1036)	1.6 (102)	17.8 (1138)

Table 1: Clause-initial *I think* followed by *that*- and zero *that*-clause in ICE-GB normalized per 10.000 words (raw figures in brackets)

For the present study only subsection Public dialogue (s1b) was taken into account, which has a sufficiently large number of *that*-clauses. The prosodic analysis is based on all 52 occurrences of *I think* + *that*-clause and 148 random instances of *I think* + zero *that*-clause.

Section 2 first concentrates on the analysis of *I think* + zero *that*-clause. It identifies essentially three different prosodic patterns for *I think* and discusses to what extent they can be used for identifying its syntactic status. It is argued that a simple correlation of prosodic prominence with matrix clause status is problematic, as it ignores the linear dimension of prosody. Section 3 focuses on *I think* + *that*-clause. Here the corpus analysis shows a distribution of prosodic patterns parallel to that of *I think* + zero, which can be taken as an indication of similar uses of both constructions. A closer investigation of the corpus data reveals that the *that*-complementizer is not so much used as an explicit marker of subordination but rather as a filler, which gives extra weight to *I think* or is used for rhythmical reasons.

2 Prosody of *I think* + zero

In spoken language, prosody is a prime indicator of prominence, with prosodic prominence iconically reflecting communicative salience of a linguistic element. In addition, however, prosody also has an important cohesive (linking) function in the sense that prosodic phrasing can signal which linguistic items belong together. In the case of comment clauses prosody may therefore be an important cue for establishing the ‘scope’ of the comment clause. In other words, it can indicate the host construction over which it operates (cf. Kaltenböck 2008 for a detailed discussion). Prosody may therefore be crucial for identifying a comment clause as initial or final, as in example (2), which, in its written form, would allow for *I think* to be associated either with the preceding material (*nine is report too*) or the following (*an interesting document*). Its prosody, however, clearly marks it as initial comment clause.

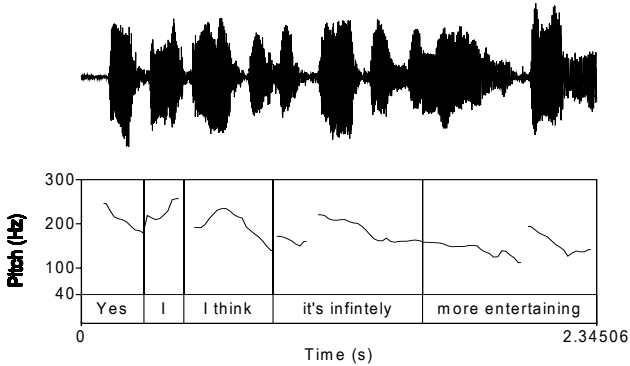
(2) *Nine is report too (I think an interesting document which uhm Professor Greenbaum initiated and which I hope everybody uh will have had a chance to digest (s1b-075-128)*

In a previous study I have distinguished between prosodically bound and independent comment clauses. For comment clauses in initial position the patterns identified were ‘right-bound’, i.e. integrated as head or pre-head into the following tone unit, and ‘independent’, i.e. with a separate tone unit and nucleus (Kaltenböck 2008). All in all, there are thus three different prosodic patterns available for initial *I think*: (i) separate tone unit, (ii) part of the head, or (iii) part of the pre-head. Typical examples of each type are given below.

(i) *I think* with an independent tone unit is illustrated in example (3), which has

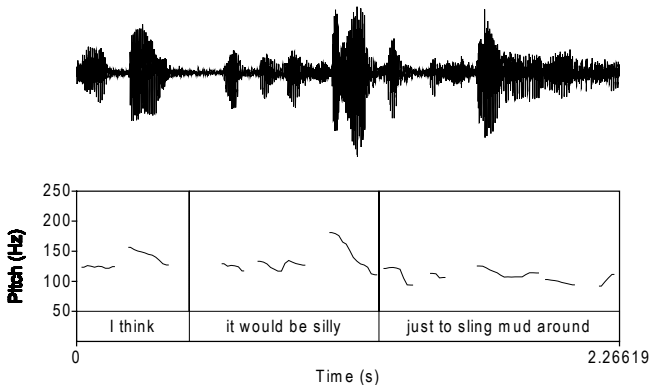
its own nuclear accent on *think* and is followed by a tone unit boundary, indicated by a change in pitch level (cf. Cruttenden 1997: 35 on boundary markers). As a possible alternative the nucleus may also be on the pronoun *I* rather than on the predicate *think* (cf. Simon-Vandenberg 2000: 50; Kaltenböck 2009 forthc. for the function of such uses).

(3) *Yes I I think it's infinitely more entertaining* (s1b-024-12)



(ii) An example of *I think* integrated into the head is given in (4), where *think* represents the first accented syllable in the tone unit, the so-called onset (e.g. Wells 2006: 207) but is less prominent than the nuclear accent on *silly* (cf. Cruttenden 1997: 54 for a definition of head).

(4) *I think it would be silly just to sling mud around* (s1b-022-19)

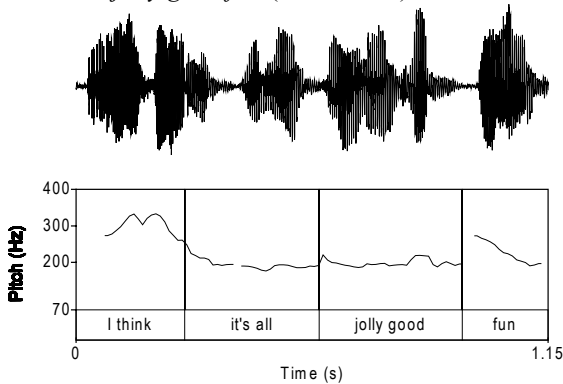


Typically in such cases the accent will be on the predicate *think*, as in example (4) above. As unstressed element, *I* represents the pre-head but may be suppressed altogether as in (5).

(5) ***Think*** the tutorials are helpful (s1b-015-4)

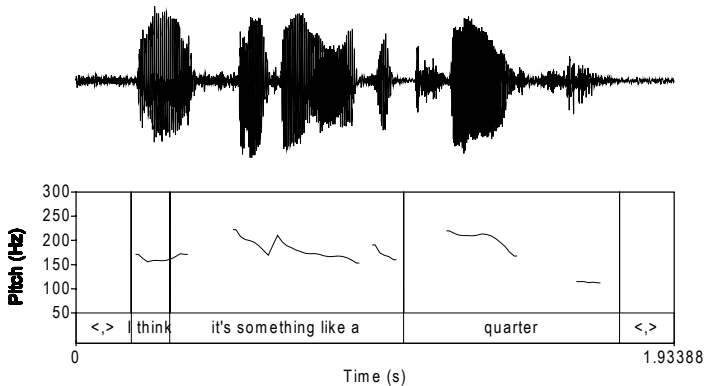
Occasionally, however, the accent occurs on the *I* (rather than on *think*), which then starts the head and gives the *I* an implicit contrastive interpretation (I as opposed to someone else), as in example (6).

(6) ***I think*** it's all jolly good fun (s1b-024-28)



(iii) The third prosodic pattern is that of integration in the form of a pre-head, i.e. an unaccented (typically unstressed and anacrustic) syllable preceding the head (cf. Wells 2006: 214-15).² This pattern is exemplified in (7), where *I think it's* forms the pre-head, followed by an accented syllable *some*, which starts the head, and the nucleus on *quarter*.

(7) ***I think*** it's something like a quarter (s1b-030-29)



As is to be expected with natural speech data, the differences between the three patterns are not always clear-cut and we have to allow for fuzzy boundaries between them. For instance, the difference between pattern (i) (nucleus) and pattern (ii) (head) is essentially one of different degrees of prominence with the onset of the head being the first accented syllable in a tone unit and therefore by definition less prominent than the nuclear syllable. There are, however, some cases where the onset of the head receives so much pitch prominence that it might qualify for classification as separate tone unit. In such cases the following criteria were applied to distinguish between heads and nuclei:

(a) Onset syllables are generally on a higher pitch level than the nucleus owing to declination within a tone unit, i.e. the fact that pitch tends to be lower at the end of a tone unit than at the beginning (e.g. Couper-Kuhlen 1986: 82-83, Wichmann 2000: 103-105).

(b) If at the beginning of a tone unit, i.e. not preceded by a pre-head, the onset will often be anacrustic, i.e. produced with greater speed (cf. Cruttenden 1997: 32).

(c) Only in the case of a separate nucleus is *I think* followed by a tone unit boundary, as indicated by features such as anacrusis, final syllable lengthening, change of pitch level or pitch direction of unaccented syllables (cf. Cruttenden 1997: 35).

(d) Onsets are less prominent than nuclear accents, which is reflected phonetically in a smaller range of pitch movement and/or weaker energy pulses.

For the corpus analysis a total of 148 instances of *I think* + zero in the text category Public dialogue (s1b) were prosodically analysed both auditorily and instrumentally with the help of the acoustical analysis programme PRAAT (Boersma & Weenink 2008). Of these 148 instances 13 are cases where *I think* is separated from the following clause by a pause or some intervening material such as a filler or hesitation sound.³ Such cases, which are generally accepted as uncontroversial instances of comment clauses in the literature (e.g. Peterson 1999: 236, Biber et al. 1999: 197) were singled out in a previous study (Kaltenböck 2008) for prosodic analysis in an attempt to exclude potential main clauses or ambiguous cases. The present study includes these together with other (i.e. non-separated) instances since the two sets of data have been found to show no difference prosodically. The results of the prosodic analysis for *I think* followed by a zero *that*-clause are summarised in Table 2 below.

	n	%
Prosodically independent	7 (of which nucleus on <i>I</i> : 3)	4.7%
Right-bound: part of head	112 (of which accent on <i>I</i> : 9)	75.7%
Right-bound: part of pre-head	29	19.6%
Total	148	100.0%

Table 2: Prosodic patterns of initial *I think* in Public dialogue (s1b) followed by a zero *that*-clause

We can see that the dominating pattern is that of *I think* being realised as part of the head (75.7%), followed by its realisation as pre-head (19.6%). An independent tone unit for *I think* is extremely rare (4.7%). This lack of nuclear prominence is, however, not really surprising, as *I think* is clearly the most grammaticalized of all comment clauses and has therefore been subject to a high degree of semantic bleaching (cf. Mindt 2003, Kaltenböck 2008). This semantic reduction makes *I think* an unlikely candidate for nuclear highlighting. Nonetheless, *I think* does receive some prosodic prominence, as can be seen from the high number of heads as opposed to pre-heads. This can be explained by its initial position, where it has important signposting function, introducing a new thought or indicating a new turn.

In a previous study I have identified various functions of comment clauses and *I think* in particular (Kaltenböck 2008, 2009 forthc.), showing that comment clauses can be further grammaticalized from epistemic markers into pleonastic structuring devices. These uses tend to be phonetically reduced and lack prosodic prominence. Initial *I think* realised as pre-head can therefore safely be equated with this structural function.

In a wider perspective, such pleonastic filler uses can be seen as the final stage in a long process of grammaticalization with *I think* starting out as matrix clause (cf. Thompson & Mulac 1991) and developing via epistemic marker (comment clause) into a discourse marker with filler function (cf. Mindt 2003, Kaltenböck 2007, 2008). Although the exact origin of comment clauses is disputed, with Brinton (1996) and Fischer (2007) dismissing Thompson and Mulac's matrix clause hypothesis and suggesting a derivation from adverbial clauses (cf. *as I think*), it is generally agreed that the starting point must have been a fully lexical item. In the case of *I think* this full lexical meaning can be paraphrased as 'cogitation' (Aijmer 1997). This process of grammaticalization is of course far from finished with the different stages existing side by side.

In an attempt to systematise the development of such complement-taking predicates (CTPs, e.g. *I think*) Boye and Harder (2007) propose the following

three stages: (1) primary lexical CTPs, (2) secondary lexical CTPs, and (3) secondary grammatical CTPs. This classification takes into account both the structural and the usage status, each of which is described by a binary set of values, viz. lexical-grammatical structural status and primary-secondary usage status. While the first stage is easily identifiable as matrix clause and the last stage as (clause internal and final) comment clause, the second stage is a hybrid category, which exhibits a discrepancy between usage status and structural status and as such is descriptively ambiguous (cf. Boye & Harder 2007: 586). *I think* in clause-initial position seems to qualify for precisely this intermediate stage: its morphosyntactic form is that of a lexical clause (Boye & Harder 2007: 591) and its syntactic position that of a prototypical matrix clause. However, in terms of its discourse function initial *I think* is typically secondary. There are hardly any cases in the corpus where contextual use would suggest communicative salience in the sense that *I think* expresses the main point of the utterances (cf. e.g. Boye & Harder 2007: 575).

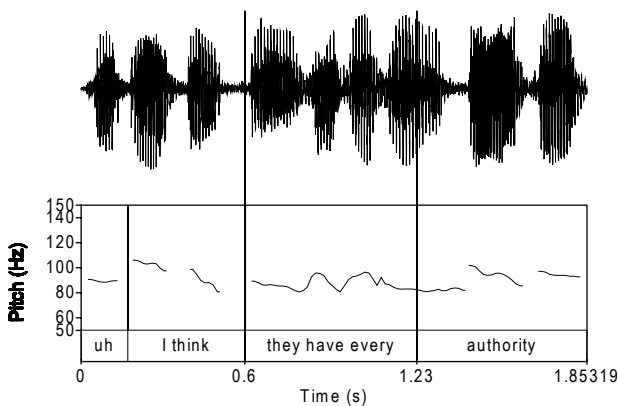
In a similar vein, Nuyts (2000: 122ff) identifies epistemic modal expressions (e.g. *I think*) as a ‘battleground’ where two conflicting functional forces are at work: an information structural force and an iconic (or conceptual semantic) force. In cases of epistemic evaluation these forces pull in different directions. From the perspective of iconicity the status of the epistemic evaluation is that of an operator (i.e. a meta-representational element) over a state of affairs and therefore favours main clause status for the epistemic expression “since it directly reflects the meta-status of the qualification relative to the state of affairs” (Nuyts 2000: 123). In terms of information structure, on the other hand, the epistemic qualification is clearly backgrounded and the state of affairs foregrounded, carrying the focal information. The information structural force therefore works against a main clause interpretation for the epistemic expression, since main clauses prototypically carry foregrounded information and embedded clauses backgrounded information (cf. Givón 1984, Mackenzie 1984, Sadock 1984, Tomlin 1985). With initial *I think*, however, positioning suggests main clause status.

The result of these conflicting forces seems to be one of neutralisation and indeterminacy. Initial epistemic markers such as *I think* can therefore be thought of as ‘undecided battles’, to take up Nuyts’ metaphor, where the different forces outbalance each other and allow for different interpretations of the status of *I think*. In spoken language, of course, an additional force enters the ‘battleground’ and may ‘tip the scales’: prosody. As an iconic reflection of prominence, prosodic signals may be seen as decisive factor for the interpretation of the syntactic status of *I think* and it is tempting to correlate prosodic prominence in the form

of a separate nuclear tone (tone unit) with matrix clause status and reduced prominence in the form of a head with comment clause status. Pre-heads, as pointed out above, can safely be associated with fully grammaticalized filler function. Such a correlation of syntactic status with prosodic prominence would also correspond with the intuitive notion that initial *I think* only rarely acts a matrix clause (cf. also Kärkkäinen 2003), as is evidenced by the low number of nuclear tones on *I think*.

However, while a simple correlation of prosodic prominence with the syntactic or functional status of *I think* may be intuitively appealing and may have indeed some theoretical value, it falls short of providing a complete explanation for the corpus data. Simply correlating hierarchical status, i.e. main vs. comment clause, with degrees of prosodic prominence ignores the fact that prosody not only has ‘vertical’ function in the sense of foregrounding/backgrounding or *mise en relief*, but may also have linear or ‘horizontal’ function by linking and rhythmically structuring elements of speech. A closer look at the corpus data shows that there are indeed cases where prosodic prominence seems to have been prompted by rhythmic considerations. Compare, for instance, example (8), where the separate chunking of *uh I think* as an independent tone unit with nucleus on *I* may have been influenced by an implicit desire to conform to a rhythmical pattern which involves chunks of roughly six milliseconds: / *uh I think* / *they have every* / *authority* /

- (8) *Uh I think they have every authority both from their governments and from the UN resolutions to do that* (s1b-027-103)

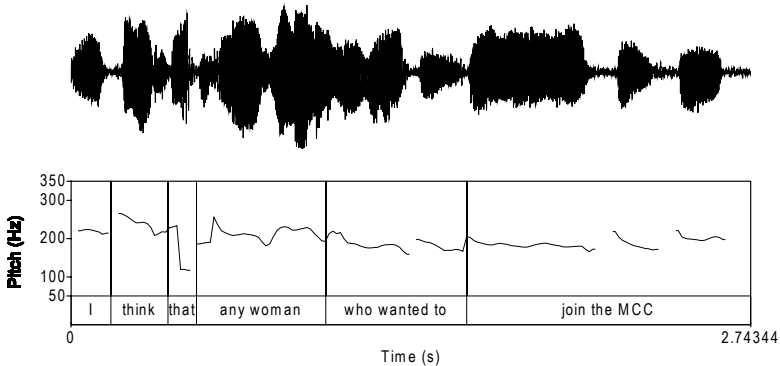


This temporal or linear aspect may be more prominent where *I think* is used in a hesitation phase as a staller, whose function is to ‘buy time’. Giving *I think* more prominence (e.g. a nuclear rather than an onset accent) may allow the speaker to do precisely that. It may also be assumed that rhythmic considerations come more into play in public speaking with experienced speakers (i.e. the text category under investigation). Note also that the insertion of a *that*-complementizer in the above example would disrupt the regularity of the rhythm. I will discuss this issue in more detail in the following section.

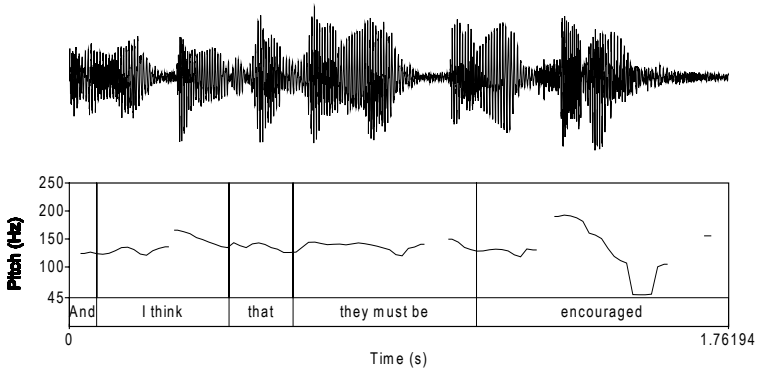
3 Prosody of *I think* + *that*

Let us now turn to a closer investigation of *I think* followed by a *that*-complementizer, which is, with only 52 instances, the marked option (cf. Table 1 above). These cases are particularly interesting, since *that* is generally seen as explicit marker of subordination and as such would identify *I think* as main clause. The prosodic analysis of all 52 instance of *I think* + *that* in Public dialogue (s1b) shows the same three patterns as identified above for zero clauses, viz. (i) nuclear accent, (ii) accented syllable in the head, and (iii) pre-head. These three patterns are illustrated by the examples in (9), (10) and (11) respectively.

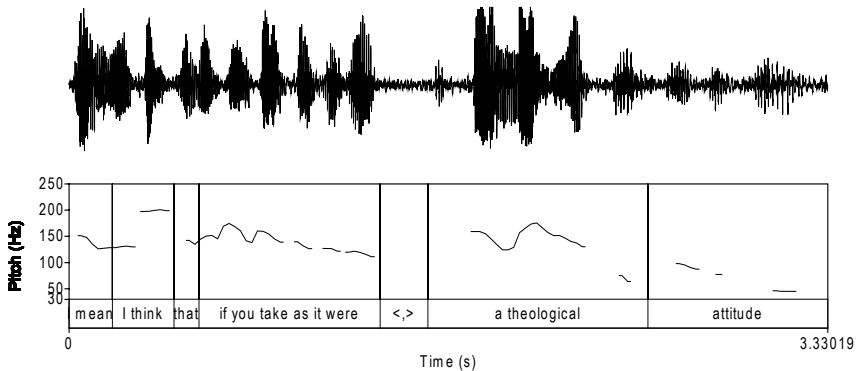
(9) *I think that / any woman who wanted to join the MCC* (s1b-021-26)



- (10) *And I think that they must be encouraged* (s1b-036-72)



- (11) *I mean I think that if you take as it were a theological attitude* (s1b-039-93)



In example (9) *think* takes a nuclear tone with a tone unit boundary after the complementizer, as indicated by the pitch change on *any*. In example (10), on the other hand, *think* represents the onset of the head, which leads up to (and includes) the initial syllable of *encouraged*. *Think* is preceded by the unstressed syllables *and* + *I*, which represent the pre-head. In example (11) the pre-head includes both *I mean* and *I think*, with the head starting on *that*. As noted for zero *that*-clauses, the accent (both nuclear and non-nuclear) may shift away from *think* to the pronoun *I*, as for instance in example (12) below.

If we compare the distribution of the three prosodic patterns for *I think* + *that*-clause with that of *I think* + zero as discussed in the previous section, we find that they closely correspond. Table 3 shows that the most frequent pattern by far is again that of heads (75%), followed by pre-heads (13.5%) and independent tone units (11.5%).

	n	%
Prosodically independent	6 (of which nucleus on <i>I</i> : 2)	11.5%
Right-bound: part of head	39 (of which accent on <i>I</i> : 5)	75.0%
Right-bound: part of pre-head	7	13.5%
Total	52	100.0%

Table 3: Prosodic patterns of initial *I think* in Public dialogue (s1b) followed by a *that*-clause

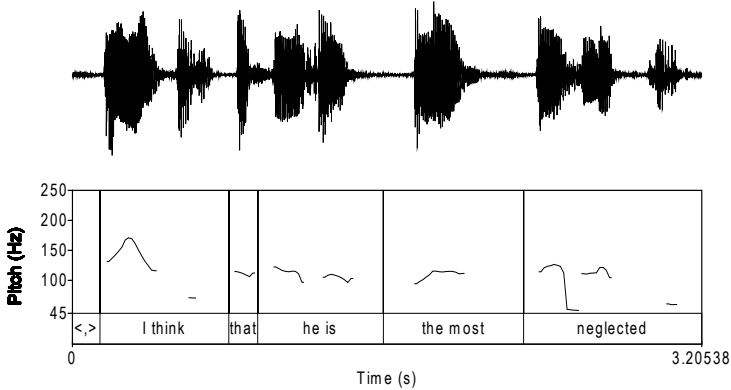
If we take nuclear prominence in the form of a separate nuclear tone as a possible cue for main clause status (as discussed in Section 2), we have to conclude that the prosodic realisation of *I think* provides no indication for a general main clause status of *I think* + *that*, only for a small minority of instances (11.5%). This may be somewhat surprising as *that* is a marker of subordination and therefore points at main clause status of *I think* on the structural level (cf. Boye & Harder 2007). On the usage level, however, *I think* + *that* generally has secondary status, as indicated by its preferred prosodic realisation: its reduced prosodic prominence signals that *I think* is not normally the main asserted content of the utterance. Moreover, the similar distributional pattern of the three prosodic types for *that* and zero clauses suggests there is no fundamental difference in usage between the two constructional types. This confirms an assumption that has already been variously expressed in the literature, for instance by Kärkkäinen (2003, 2009 forthc.) or Nuyts (2000: 129 note 13).⁴

On the level of usage, therefore, overt marking of subordination by a *that*-complementizer does not make a significant difference in terms of prosodic foregrounding of *I think*. The only difference is that with zero clauses we find a somewhat higher percentage of pre-heads and lower percentage of nuclear accents, which can be taken as an indication of *I think* + zero having moved even further down the path of grammaticalization. Overall, however, the distributional pattern is similar, which in turn raises the question whether the main function of *that* is really that of a marker of hierarchical difference between the two clauses. I will discuss this point in more detail in the following.

If we take a closer look at the prosodic realisation of the *that*-complementizer, we notice that it can be intonationally grouped either with *I think* or the following clause. This difference in association is most obvious in cases where *I think* carries

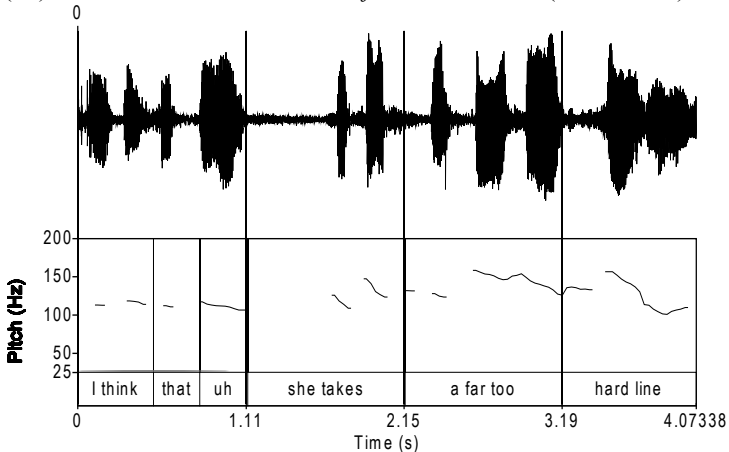
its own nuclear tone and is therefore followed by a tone unit boundary which may either include *that*, as in example (9) above, or exclude it, as in example (12).

- (12) *I think / that he is the most neglected of that uh number of composers around the turn of the century* (s1b-032-103)



In head and pre-head realisations of *I think* the complementizer is typically integrated into the larger pitch contour but may occasionally also show signs of association or dissociation with *I think*, albeit less markedly so. Compare for instance example (11) above where *that* is part of the head with example (13) below, where it is part of the pre-head.

- (13) *I think that uh she takes a far too hard line* (s1b-035-20)

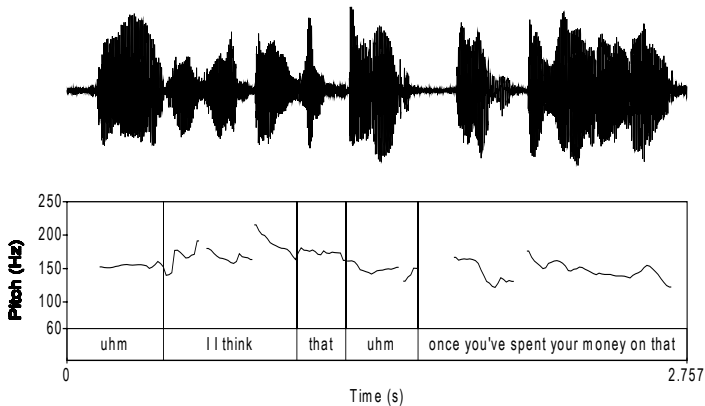


Prosodic realisation of *that* therefore does not necessarily reflect the syntactic analysis of the construction, which identifies the complementizer as part of the subordinate clause.⁵ Such a mismatch between syntax and prosody is hardly surprising and has been noted by previous studies (e.g. Brazil 1997, Wichmann 2001, Dehé 2007). It is interesting, however, that there is a tendency for *that* to be prosodically grouped with *I think* rather than the following clause.⁶

Associating the complementizer on the usage level with *I think* (and indeed inserting it in the first place) seems to result from a desire on the part of the speaker to add weight to *I think* in the form of an extra syllable. The reason for this may be twofold:

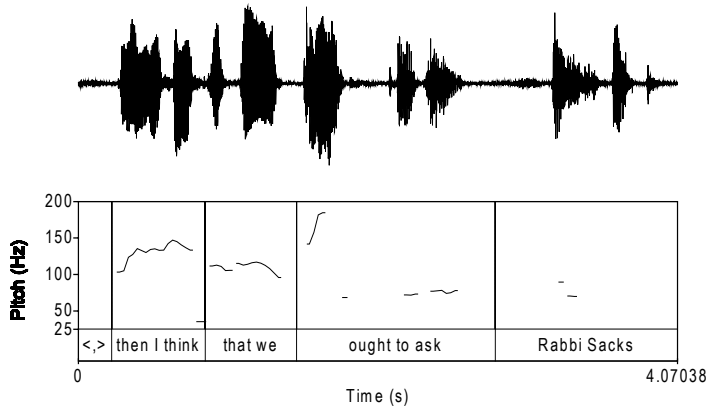
First, adding an extra syllable to the initial clause makes it longer and therefore more effective as clause initial staller used for bridging a hesitation phase. Compare, for instance, the following example where *that* has a staller function similar to that of *uhm* and *and* (cf. also example 13 above).

- (14) *Uhm I think that uhm once you've spent your money on that the thing to spend your money on is a subscription to the local horticultural society* (s1b-025-133)



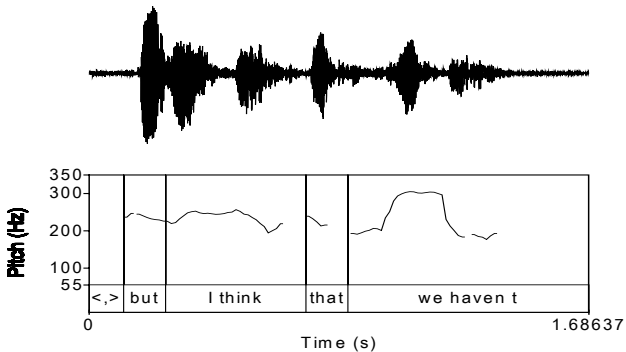
Second, prosodic association or dissociation of *that* with *I think* can be motivated by rhythmic considerations. This is illustrated, for instance, in example (15), where chunking *that* with material following it rather than material preceding it results in two rhythmic chunks of roughly equal length: *then I think* and *that we*.

- (15) *Then I think that we ought to ask Rabbi Sacks t uh uh to uh uh to say more because of course he has said two important things* (s1b-028-63)



Occasionally, rhythmic chunking of *that* is underpinned by parallel intonation contours as in example (16), where association of *that* with the following clause creates two three-syllable chunks (*but I think* and *that we haven't*), each with the same fall-rise-fall intonation contour.

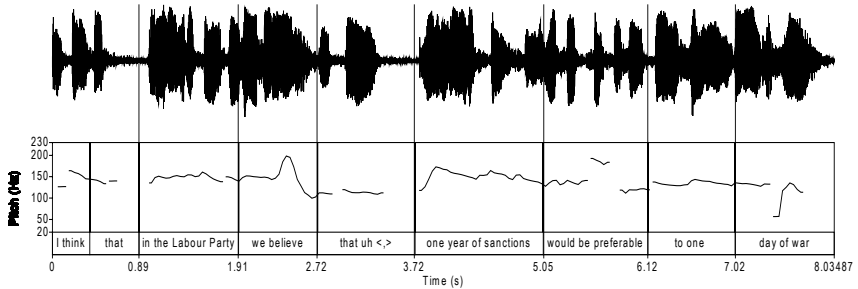
- (16) *But I think that we haven't in the sense that we have just classification still* (s1b-012-104)



The underlying principle for rhythmic chunking seems to be that of rhythmic harmony, viz. a tendency towards rhythmic chunks of roughly equal size

(cf. principle of isochrony). This is illustrated in example (17), where association of *that* with *I think* brings the first rhythmic unit in line with the average length of the following ones, i.e. roughly one millisecond (note incidentally the same length of the second hesitation phase: *that* + *uh* + pause <,>).

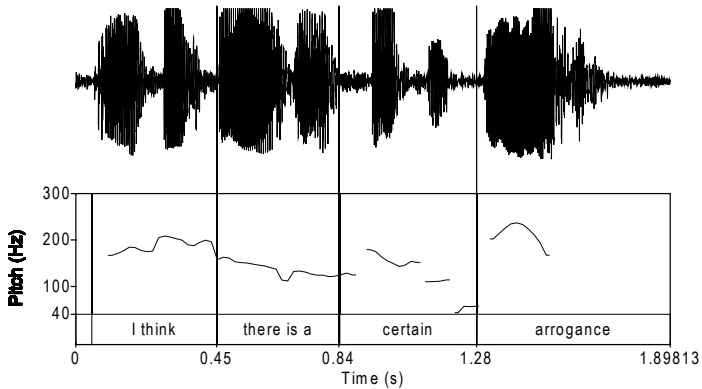
- (17) *I think that in the Labour Party we believe that uh <,> one year of sanctions would be preferable to one day of war* (s1b-035-29)



While it is clear that the principle of rhythmic harmony cannot be pressed too far, it seems that the text type of public conversation is particularly susceptible to it, especially the text categories broadcast discussions and broadcast interviews, which typically involve highly experienced public speakers and incidentally have the highest proportion of *that* in the corpus (6.6 and 3.2 occurrences per 10,000 words respectively).

Moreover, the rhythmic structure of the construction is closely linked to the type of subject in the *that*-clause. Consider, for instance, example (18), where the subject of the second clause consists of an unstressed syllable (*there*), which is followed by another two unstressed syllables (*is a*). The resulting rhythmic pattern of the entire construction *I think there is a certain arrogance* (which has an accent on *I*) is thus: ●●●●●●●●●●. Note that a *that*-complementizer would add an extra unstressed syllable to the row of three unstressed syllables, which is not desirable for rhythmical reasons.

- (18) *I think there is a certain arrogance on the part particularly of the extreme left in Britain on this matter* (s1b-027-136)



A closer analysis of the corpus data shows that unstressed subjects in the form of existential *there* (*there + be*) or the pronoun *it* (*it + be*) strongly prefer omission of *that*: 28.4 per cent (46 instances) of all zero clauses have such unstressed subjects, but only 6.9 per cent (2 instances) of *that*-clauses. This finding ties in with Elsness' (1984) observation that complex subjects correlate with *that*-retention. As a possible explanation for this he notes that “[a]lthough there is no risk of ambiguity in such constructions, one may see the selection of *that* connective as a contribution to greater syntactic clarity” (ibid.: 532). This may be true for written texts. For spoken language, however, it is necessary to take into account rhythmic considerations, viz. unstressed subjects favouring *that*-omission and, as a second factor, memory constraints in online production. Thus, production of a syntactically complex subject, which can also be expected to have high informational value, will normally require extra ‘thinking time’, which is provided for by the *that*-complementizer.

The *that*-complementizer, in other words, has an important temporal function, like typical fillers, which allow the speaker to buy time. This in turn can help alleviate production difficulties, as noted for instance by Jaeger (2005) (cf. also Clark 2004). Close analysis of the corpus data shows that there is indeed a trade-off between the use of *that* and production difficulties, with insertion of *that* correlating with fewer instances of repetition and/or restarts immediately preceding or following *I think that*. More precisely, with *that* omission we find such disfluencies in 16.7 per cent (27 instances) of all cases, such as example

(19). With *that* insertion, on the other hand, such disfluencies occur in only 3.4 per cent (1 instance) of all cases.

- (19) *III think there wh some of us are in great difficulty here*
(s1b-028-101)

To sum up, the corpus results suggest that the *that*-complementizer following initial *I think* acts like a typical filler. Its function in spoken language is therefore primarily a linear one, i.e. on the temporal plane, not so much a hierarchical one, i.e. marking syntactic subordination and backgrounding. Such syntactic backgrounding, incidentally, would run counter the typical pragmatic use and information structure of these constructions, where the *that*-clause presents the main point of the message. *I think* only has secondary, qualifying function, which typically reduces the speaker's commitment to the proposition of the *that*-clause. In fact, it is precisely this hedging or distancing function of *I think* that makes the use of *that* as a marker of subordination redundant. As argued elsewhere (Kaltenböck 2006), the (hierarchical) function of a *that*-subordinator is essentially also that of distancing the speaker from the proposition it introduces. With initial *I think* this distancing function is already taken care of and results in omission of *that*,⁷ except where it is needed for linear purposes, i.e. as a filler.

4 Conclusion

The analysis of the corpus data has shown that difference in syntactic realisation, i.e. presence or absence of the *that*-complementizer, does not correspond with different prosodic behaviour. Although *I think* + *that* shows a slightly higher propensity to occur with a separate nuclear tone than *I think* + zero (which conversely shows a slightly higher preference for pre-head realisation) both constructional types exhibit a similar distribution of the three prosodic patterns identified: they are both most frequently realised as heads, less frequently as pre-heads, and only rarely with a separate nuclear accent. This means that the two formal signals available for indicating relative prominence of *I think*, prosody and explicit marker of subordination, do not match. If we take relative pitch prominence as an indication of matrix clause status, we have to conclude that both constructional variants may qualify for main clause status but at the same time only very rarely do. This equivalence in actual use of the two syntactic types casts doubt on the subordinator function of the *that*-complementizer in spoken language, which is corroborated by prosodic evidence suggesting that

the *that*-complementizer is mainly used as a filler to give weight to *I think* or for rhythmic purposes.

Prosodic evidence thus supports the view that initial *I think* generally has secondary status as a qualifier of the proposition in the following clause rather than being the main assertion itself. Whether this is enough for classifying initial *I think* as a comment clause rather than a main clause depends on one's definition of a main clause. In a cognitive-functional view an analysis of initial *I think* as comment clause would seem to be justified (cf. Langacker 1991, Thomson 2002), although much depends on its actual use in context. It is clear, however, that Thompson and Mulac's (1991) assumption of the presence or absence of the *that*-complementizer marking the difference between main and comment clause is not tenable from a usage point of view. A more useful framework for the description of initial *I think* is Boye and Harder's (2007) distinction of two different levels of analysis, a structural and a usage level. It allows us to capture the mismatch between actual usage and structural representation with the latter lagging behind the former as a result of diachronic change, viz. the semantic bleaching and grammaticalization of *I think* into a discourse marker. Spoken language with its use of prosodic cues is clearly at the forefront of this development. To what extent initial *I think* can still be analysed as matrix clause (as suggested by Kearns 2007) or has already been reanalysed as comment clause (or unitary epistemic phrase, cf. Thompson & Mulac 1991) depends on one's theoretical framework. From a cognitive-functional perspective the indeterminate status of initial *I think* seems to suggest that this process of reanalysis has not only begun, but is already well under way.

Notes

¹ The different status of sentences with *that* and without is generally supported by syntactic tests such as the tag-question test (e.g. Aijmer 1972: 52, 1997: 8; Hand 1993: 501), the question test (e.g. Huddleston & Pullum 2002: 896), or the negation test (e.g. Erteschik-Shir & Lappin 1979: 56).

² The term 'stress' is used here as rhythmically stressed, while 'accent' refers to a syllable made prominent by rhythmic stress and pitch prominence, i.e. by a change in pitch, movement in pitch, or the start of a pitch movement (cf. Wells 2006: 93).

³ The total number is 14 but one occurrence did not have a sound file.

⁴ This view is also implied, although not overtly expressed, in Thompson (2002) and Boye and Harder (2007).

⁵ Cf. however the semantic analyses by Davidson (2001), Lepore and Loewer (1989) and Hand (1993) for a different view.

⁶ This is also reflected in the fact that in the corpus *I think* and *that* are never separated by any intervening material (e.g. hesitation sound, filler), whereas *that* is frequently separated from the clause it is head of. The level of performance therefore seems to suggest a closer association of *that* with the main clause rather than the subordinate clause.

⁷ A similar view has recently also been expressed by Kearns (2007), who argues that “[t]he modifier sense of an epistemic verb and its subject in matrix position promotes zero in the complement clause” (ibid.: 501).

References

- Aijmer, K. (1972) *Some Aspects of Psychological Predicates in English*. Stockholm: Almqvist & Wiksell.
- Aijmer, K. (1997) ‘*I think* – an English modal particle.’ In: Swan, T., Westvik, O. J. (eds) *Modality in Germanic Languages. Historical and Comparative Perspectives*. Berlin: Mouton de Gruyter. 1-47.
- Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E. (1999) *Longman Grammar of Spoken and Written English*. Harlow: Longman.
- Boersma, P., Weenink, D. (2008). ‘Praat: doing phonetics by computer.’ (Version 4.4.33) [Computer programme] Retrieved December 2005 from <http://www.praat.org>
- Boye, K., Harder, P. (2007) ‘Complement-taking predicates: Usage and linguistic structure.’ *Studies in Language* 31/3, 569-606.
- Brazil, D. (1997) *The Communicative Value of Intonation in English*. Cambridge: Cambridge University Press.
- Brinton, L. J. (1996) *Pragmatic Markers in English. Grammaticalization and Discourse Functions*. Berlin: Mouton de Gruyter.
- Clark, H. (2004) ‘Pragmatics of language performance.’ In: Horn, L. R., Ward, G. (eds) *Handbook of Pragmatics*. Oxford: Blackwell. 365-382.
- Couper-Kuhlen, E. (1986) *An Introduction to English Prosody*. London: Edward Arnold.
- Cruttenden, A. (1997) *Intonation*. 2nd ed. Cambridge: Cambridge University Press.
- Davidson, D. (2001). *Inquiries into Truth and Interpretation*. 2nd ed. Oxford: Clarendon.
- Dehé, N. (2007) ‘The relation between syntactic and prosodic parenthesis.’ In: Dehé, N., Kavalova, Y. (eds) *Parentheticals*. Amsterdam: John Benjamins. 261-284.
- Elsness, J. (1984) ‘*That* or zero? A look at the choice of object clause connective in a corpus of American English.’ *English Studies* 65, 519-33.
- Erteschik-Shir, N., Lappin, S. (1979) ‘Dominance and the functional explanation of island phenomena.’ *Theoretical Linguistics* 6/1, 41-86.
- Fischer, O. (2007) ‘The development of English parentheticals: A case of grammaticalization?’ In: Smit, U. Dollinger, S. Hüttner, J., Kaltenböck, G., Lutzky, U. (eds) *Tracing English Through Time*. Wien: Braumüller. 99-114.
- Givón, T. (1984) *Syntax. A Functional-Typological Introduction*. Vol. 1. Amsterdam: John Benjamins.
- Hand, M. (1993) ‘Parataxis and parentheticals.’ *Linguistics and Philosophy* 16, 495-507.
- Huddleston, R., Pullum, G. K. (2002) *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Jaeger, F. T. (2005) ‘Optional *that* indicates production difficulty: Evidence from disfluencies.’ *Proceedings of DiSS’05 Disfluency in Spontaneous Speech Workshop*. 10-12 September 2005, Aix-en-Provence, France. 103-109.
- Kaltenböck, G. (2006) ‘... *That* is the question’: Complementizer omission in extraposed *that*-clauses.’ *English Language and Linguistics* 10/2, 371-396.
- Kaltenböck, G. (2007) ‘Position, prosody, and scope: The case of English comment clauses.’ *Vienna English Working Papers (VIEWS)* 16, 3-38.
- Kaltenböck, G. (2008) ‘Prosody and function of English comment clauses.’ *Folia Linguistica* 42/1, 83-134.

- Kaltenböck, G. (2009 forthc.) 'Pragmatic functions of parenthetical *I think*.' In: Kaltenböck, G., Mihatsch, W., Schneider, S. (eds) *New Approaches to Hedging*. Amsterdam: Elsevier.
- Kärkkäinen, E. (2003) *Epistemic Stance in English Conversation*. Amsterdam: John Benjamins.
- Kärkkäinen, E. (2009 forthc.) 'Position and scope of epistemic phrases in planned and unplanned American English.' In: Kaltenböck, G., Mihatsch, W., Schneider, S. (eds) *New Approaches to Hedging*. Amsterdam: Elsevier.
- Kearns, K. (2007) 'Epistemic verbs and zero complementizer.' *English Language and Linguistics* 11/3, 475-505.
- Kruisinga, E. (1932) *A Handbook of Present-Day English*. Part II. Groningen: Noordhoff.
- Lepore, E., Loewer, B. (1989) 'You can say *that* again.' In: French, P., Uehling, T. E., Wettstein, H. K. (eds) *Midwest Studies in Philosophy*. Vol. XIV. Notre Dame, Ind.: University of Notre Dame Press. 338-56.
- Langacker, R. W. (1991) *Foundations of Cognitive Grammar*. Vol II: *Descriptive Applications*. Stanford: Stanford University Press.
- Mackenzie, J. L. (1984) 'Communicative functions of subordination.' In: Mackenzie, J. L., Wekker, H. (eds) *English Language Research: The Dutch Contribution I*. Amsterdam: Free University Press. 67-84.
- Mindt, I. (2003) 'Is *I think* a discourse marker?' In: Mengel, E., Schmid, H.-J., Steppat, M. (eds) *Proceedings Anglistentag 2002 Bayreuth*. Trier: WVT. 473-483.
- Nelson, G., Wallis, S., Aarts, B. (2002) *Exploring Natural Language. Working with the British Component of the International Corpus of English*. Amsterdam: John Benjamins.
- Nuyts, J. (2000) 'Tensions between discourse structure and conceptual semantics: The syntax of epistemic modal expressions.' *Studies in Language* 23/1, 103-135.
- Peterson, P. (1999) 'On the boundaries of syntax: Non-syntagmatic relations.' In: Collins, P., Lee, D. (eds) *The Clause in English*. Amsterdam: John Benjamins. 229-250.
- Quirk, R., Greenbaum, S., Leech, G., Svartvik, J. (1985) *A Comprehensive Grammar of the English Language*. Harlow: Longman.
- Ross, J. R. (1973) 'Slifting.' In: Gross, M., Halle, M., Schützenberger, M.-P. (eds) *The Formal Analysis of Natural Languages*. The Hague: Mouton. 133-169.
- Sadock, J. M. (1984) 'The pragmatics of subordination.' In: Geest, W. de, Putseys, Y. (eds) *Sentential Complementation*. Dordrecht: Foris. 205-213.
- Simon-Vandenberghe, A.-M. (2000) 'The functions of *I think* in political discourse.' *International Journal of Applied Linguistics* 10/1, 41-63.
- Stenström, A.-B. (1995) 'Some remarks on comment clauses.' In: Aarts, B., Meyer, C. F. (eds) *The Verb in Contemporary English*. Cambridge: Cambridge University Press. 290-299.
- Svensson, J. (1976) 'Reportindicators and other parentheticals.' In: Karlsson, F. (ed.) *Papers from the Third Scandinavian Conference of Linguistics*. Turku: Textlinguistics Research Group, Academy of Finland. 369-380.
- Thompson, S. A. (2002) "'Object complements' and conversation. Towards a realistic account.' *Studies in Language* 26/1, 125-164.
- Thompson, S. A., Mulac, A. (1991) 'The discourse conditions for the use of the complementizer *that* in conversational English.' *Journal of Pragmatics* 15, 237-251.

- Tomlin, R. (1985) 'Foreground-background information and the syntax of subordination.' *Text* 5, 85-122.
- Urmson, J. O. (1952) 'Parenthetical verbs.' *Mind* 61, 480-496.
- Wells, J. C. (2006) *English Intonation. An Introduction*. Cambridge: Cambridge University Press.
- Wichmann, A. (2000) *Intonation in Text and Discourse. Beginnings, Middles and Ends*. Harlow: Longman.
- Wichmann, A. (2001) 'Spoken parentheticals.' In: Aijmer, K. (ed.) *A Wealth of English. Studies in Honour of Göran Kjellmer*. Göteborg: Acta Universitatis Gothoburgensis. 177-193.