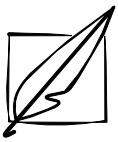

Jan Kochanowski University Press

This is a contribution from *Token*: *A Journal of English Linguistics*
Volume 1/2012.
Edited by Sylwester Łodej and John G. Newman.

Inferentials: Fixed or not?*

Andreea S. Calude and Gerald Delahunty
University of Reading and Colorado State University

ABSTRACT

In this paper, we present arguments for analysing inferentials (which we class as a subtype of *it*-clefts) as partially formulaic. By exploring excerpts of spoken (New Zealand) English from the Wellington Corpus of Spoken New Zealand English, we establish that inferentials have formulaic tendencies: they are lexically limited, situationally bound, and relatively frequent (compared to other lexical bundles), and they serve a specific discourse function. However, they are not (perhaps, yet) fully established "fixed formulae" since they are semantically transparent, compositional, and non-fluent.

1. What are inferentials?

This paper is about the sentence types exemplified in bold:

(1) **It's not that I'm so smart. It's just that I stay with problems longer**. (Albert Einstein)

(2) Jake, come here buddy. Sit down. Look, **it's not that I don't care what you want. It's just that you're a kid, and what you want doesn't matter**. (*Two and a Half Men*, American sitcom)

(3) **It's not that I'm afraid to die**. I just don't want to be there when it happens. (Woody Allen)

* We are grateful to the participants in the Perpignan 2010 conference on "Fixed Phrases in English" for their insightful comments on an earlier draft of this paper. Any errors that remain are ours alone.

Various theories have been put forward about the analysis of these sentence types. A common view is that they are a type of cleft (Declerck 1992; Delahunty 2001; Koops 2007). Heggie (1998) sees them as copular constructions, but not of the cleft variety. Collins (1991) and Schmid (2009) propose that they are extrapositives. Finally, Pusch (2006, forthcoming) and Fraser (1999) refer to them as discourse markers.

We have argued elsewhere (Calude – Delahunty 2011) in favour of a cleft analysis for the inferential construction. The full debate and precise arguments which we take to support the cleft view are beyond the scope of this paper. Here, we briefly summarize our understanding of how inferentials are structured because their elements are relevant to the discussion of whether or not they are fixed.

2. A cleft analysis of inferentials

Cleft constructions privilege a clause constituent by placing it in focus position with the help of a copula and a few other elements that depend on the specific cleft type. The archetypal cleft is the *it*-cleft, but in languages rich in cleft types, such as English, there can be many others, e.g., *wh*-clefts, reversed *wh*-clefts, demonstrative clefts, *since*-clefts, *all*-clefts, and so on. The *it*-cleft is exemplified in (4).

(4a) It is **depression** that I fear most.

(4b) It is **the cookie jar** that I fear most.

(4c) It is **running out of gas** that I fear most.

Examples (4a) through (4c) show that *it*-clefts can focus noun phrases of various levels of complexity: from simple phrases as in (4a) to more complex ones as in (4c).

Noun phrases are not the only types of phrases which can be focused in *it*-clefts. As shown in (5) and (6) respectively, prepositional phrases and adverb phrases can also be focused.

(5) It is **in December** that we hope to go over to Spain.

(6) It is **intelligently, not speedily** that we want our employees to work.

Though they allow variants that we discuss below, *it*-clefts basically consist of an expletive *it*, a form of the copula, a focused expression, and a clause containing a gap of the same syntactic type as the focus. We will refer to the *it* + copula part as "the matrix," the post-copular constituent as "the focus," and the gapped clause as "the clause."

Sentences (1)-(3) and (7) show that a finite clause may occur in the focus position, and it is for this and other reasons that we view inferentials as a subtype of *it*-cleft.

(7) It is **that inferentials are too difficult** that I fear most.

However, many inferentials (particularly those in speech, see Koops 2007), occur without the gapped clause, as in (8).

(8) It is **that inferentials are too difficult.**

In itself, this is not problematic because other *it*-clefts exhibit a similar pattern, discussed by Hedberg (2000), Declerck (1988) and Huddleston – Pullum (2002), under the label "truncated *it*-cleft." Hedberg's classic example is reproduced below.

(9) Who ate the last cookie? It wasn't **me**.

Examples like (9) are *it*-clefts because the ellipted clause is generally recoverable from context, as in (10).

(10) It wasn't me **that ate the last cookie**.

This reasoning applies to the truncated inferential in (8): the ellipted clause may be recovered from context, as in (11).

(11) It is that inferentials are too difficult **that most concerns us.**

However, it is not always possible to recover the missing clause of a truncated inferential, e.g., it is not at all clear what the ellipted clause of the inferential bolded in (12) might be. (This is from the Wellington Corpus of Spoken New Zealand English, henceforth WSC. See section 3 for details, Holmes et al. (1998) for discussion, and the Appendix for a list of annotations used in the excerpts.)

(12) WSC DPC290

FR: [tells a story] … like i was very you know when <laughs> when we were kids we were always taught you gotta lock the car before you leave so i locked all the doors and <"> everything and so the um yeah so i locked his keys in there and so i told him he needed a spare key in the place sec spare set but he didn't you know and <quickly> then the other night </quickly> i know what happened to me

MQ: mm yeah yeah <laughs> he does **it's just he doesn't trust you** that's all cos he knows that you'll get a hold of it and <latch>

FR: <laughs> oh yeah

MQ: take it for a burn

FR: you reckon

Even though an appropriate clause may not be recoverable, we are justified in categorizing inferentials as *it*-clefts because of their lexical and syntactic commonalities and the fact that *it*-clefts and inferentials function very similarly in context, as we show below.

3. Data

The data for this paper consist of 55 inferentials taken from excerpts in the Wellington Corpus of Spoken New Zealand English (WSC), which contains approximately 250,000 words of spontaneous conversation. New Zealand English is probably the newest variety of English; its origins date back to the 19th century flow of Australian and British immigration. The most distinctive features of the New Zealand English variety pertain to its phonology (centralised vowels and various mergers), lexis (loanwords from the indigenous te reo Māori) and the widespread use of the final pragmatic particle *eh*. As far as grammar is concerned, New Zealand English does not have any completely unique variants, but rather exhibits a unique combination of features found in other varieties (mostly British English and its non-standard varieties), e.g., the use of *youse/yous* for the second person plural, the use of *she* in inanimate contexts (*she'll be right*), unmarked plurals (*that'll be fifty cent*), double comparisons (*more better*), demonstrative *them* (*them things, them people*), no clear distinction between *shall* and *will*, variability in past participle forms such as *proved* and *proven*, confusion between *bought* and *brought*, and so on (see Bauer 2007 for a summary). As far as we

are aware, no variation has been documented with regard to inferentials specific to New Zealand English.

Our WSC data were identified manually: because the constituent elements of inferentials occur in a broad range of constructions it was impossible to automate the search. Any corpus search for inferentials quickly turns up expressions that share many elements with the canonical inferential and thus raise the question of what to include in the data set. We found instances of expressions that are expectable grammatical variants of the canonical inferential: instances in which the copula is in the past tense; instances from which the conjunction *that* was missing; instances in which the clause was modified by adverbs such as *just*, *only*, *simply*, *actually*, and *not*; as well as instances with a modal, *could* or *may*, in the matrix. Because these are variants of the basic inferential licensed by the grammar and because they function contextually like inferentials but with modulations predictable from these grammatical modifications, we include them in our count.

We also discovered examples that seemed somewhat more distant from the canonical inferential, viz., those with the conjunctions *as if* and *like* in place of *that*. Because their meanings and discourse properties are similar to those of canonical inferentials and their differences from canonical inferentials are predictable from the meanings of *as if* and *like* we include these in our count also. (Calude and Delahunty 2011 provide an expanded discussion of this issue.) Table 1 lays out the variants.

Table 1. Inferentials in spoken New Zealand English

Types		Example	Spoken NZ English (WSC corpus)
Modified			
	negation	*it's not that/as if*	9
	just	*it's just that*	39
	epistemic modals	*it could/may be that*	1
	discourse markers	*it is well/you know/ i mean that*	2
Unmodified			
	plain unmodified	*it is that …*	4
TOTALS			55

4. The interpretations of inferentials

Positive cleft sentences are widely regarded as foregrounding the focus expression relative to some background, which, in the case of full canonical clefts, is the clause: the clause represents an open proposition and the focus represents the value of its variable (Delahunty 1984). In the case of truncated clefts, the focus is foregrounded relative to some contextually recoverable open proposition. Positive clefts assert and negative clefts deny that the focus is the relevant value for the open proposition. Canonical inferentials function similarly: the positive ones assert the relevance of the clause against a pragmatically determined local context; negative ones deny its relevance against such a context. (See Delahunty 2001 and Calude – Delahunty 2011 for more developed discussions of these claims.) We can see these interpretations in the following extracts.

4.1 That-*inferentials*

We assume that inferentials with *that* as introducer of the focal clause are the basic type of inferential and that inferentials with matrix negation or without a clause introducer are variants of this type. Delahunty (2001) is concerned solely with *that*-inferentials and argues that these inferentials are to be interpreted either as an aspect of the context in which some other expression is to be interpreted, or as an interpretation or reformulation of an immediately prior piece of text (an interpretation that does not seem relevant to any of our current examples). It is also argued there that the contextualized interpretations of *that*-inferentials are derivable from the interaction between the lexico-grammatical characteristics of the inferential, the principles of relevance theory, and the local context, with no need to stipulate any construction-specific procedural or idiomatic interpretation. The proposition represented by the inferential clause may function as either a premise or a conclusion relative to another contextually determined proposition, though the former seems to be more frequent, both in English and cross-linguistically (Delahunty – Gatzkiewicz 2000; Delahunty 2001). Inferentials interpreted as premises may be more richly interpreted as explanations, causes, reasons, and the like; inferentials interpreted as conclusions may be more richly interpreted as results or consequences (Delahunty 2001: 536-537). As an illustration, consider LL's negative *that*-inferential in (13):

(13) WSC DPC007
RR: you so you don't want to have to look over your shoulder to make sure you don't
→ LL: oh **it's not that i don't want to have to look over my shoulder** NOTHING should HAPPEN should OCCUR during those procedures ANY PART of it that's ALL formal the WHOLE lot right from the time the people come onto the marae (Maori: "meeting house") until the time everyone's LEFT the marae all right

RR's comment that LL would not want to look over his/her shoulder is prologue to what appears to be a result clause, *to make sure …*, and so functions as a premise. LL recasts this proposition as a negative inferential, thereby denying its relevance. The proposition represented by the inferential functions as a conclusion which follows from the remainder of his/her utterance: s/he won't want to look over her shoulder because NOTHING should HAPPEN, etc.

Because the basic inferential can be interpreted without resorting to any construction-specific semantic or pragmatic stipulations, we believe that the other inferential types may also be so interpreted, and that their interpretational differences from the basic type are due entirely to their lexico-grammatical differences from that type. In the following sections, we aim to show that this assumption is well-founded.

4.2 *Negative inferentials*

Negative inferentials work as their linguistic characteristics would predict – the inferential form triggers the interpretation that the focused clause is to be interpreted as a contextually marked assumption whose relevance is denied by the speaker.

In (14), KT describes the Māori classes she chose and her reasons for choosing them. Her initial utterance potentially implicates that she did not remain in the bilingual class beyond her first year, an interpretation consistent with her inferential. KT's inferential functions as a premise, which moots and rejects the proposition that she wanted to leave the bilingual class and go to the main stream as an explanation for her not continuing in the bilingual class. The sentence following the inferential (underlined) explicitly expresses her actual reason for leaving the bilingual class.

(14) WSC DPC240

KT: you know like i was in the bilingual class in my first year you know cos she was she we were just having a chat and um she said have you got a piece of maori in you are you part maori and i said yeah my dad's just under half and um <,> she said oh yeah you know there was a few teachers that were wondering about that some of the parents and stuff and <,> and i said yeah i was in the bilingual class in my first year **it wasn't actually that i wanted to leave the bilingual class to go to main stream in my second year** it was because i wanted to take a an advanced maori paper at <,> the varsity because i wanted to you know nurture <drawls> my language

4.3 Just-*inferentials*

Just-inferentials are the most frequent type in our data, and we assume that those without *that* are elliptical versions of *that*-inferentials modified by *just*. In (15), the speakers are discussing the possibility of increasing orders for their business, which they run from home.

(15) WSC DPC293

MK: okay <next utterance directed to person with tape recorder> okay just pause it <"> can you handle like two kits

→ FY: oh forgot about that i suppose we could **it's just i ca i haven't seen the books** so thanks to your sister's fantastic way of cleaning her room

MK: because

SS: <laughs>

MK: because er <latch>

FY: suppose we could

MK: yeah because i think you know with the er <,,> with <drawls> er kate

FY: kate and carmen and see the thing is there's n if carmen and mike can pay for it this week

FY's *i suppose we could* is a dispreferred response to MK's request *can you handle like two kits*; compare it to more positive alternatives such as *Sure*. FY neither accedes nor rejects MK's request, ostensibly because FY is not sure that they can handle two kits. The inferential presents the proposition

i haven't seen the books as a contextual premise from which it would follow that FY would not know whether they could handle two kits or not.

According to Nevalainen (1997), *just* is a focusing particle meaning 'merely' and 'only', and 'exactly P' where "P" is not of great importance. (See also Quirk et al. 1985: 604.) According to Aijmer (2002: 158), the core function of *just* as 'exactly' and 'only' is a procedural marker indicating an "indexical relation to the speaker's attitudes or emotion towards a discourse event", so *just* always carries evaluative overtones. Consequently, the inferential in (15) indicates that FY's not having seen the books is the only reason why they might not be able to handle two kits, and it also suggests that this reason is of no great importance, and that MK can reasonably expect a positive response once FY has seen the books.

We note that because the clause of an inferential is interpreted as special in its context, it can be used to counter contextually possible assumptions or interpretations, which may arise from the prior discourse, as in (15). (The segment understood as the inferential trigger is underlined.)

4.4 (Not) as if-*inferentials*

Our hypothesis regarding the interpretation of inferentials with *like*, *not like*, *as if*, and *not as if* is that their interpretation is the same as that of *that*-inferentials except in so far as (*not*) *like/as if* differ from *that*. We begin with a discussion of *(not) as if* inferentials and then deal with the *(not) like* variant.

There is relatively little research on either of these inferentials, though Huddleston and Pullum (2002) include some suggestive remarks. According to Huddleston – Pullum (2002: 1146), *as* denotes comparison, and *if* "is primarily conditional" and thus has a "close relation" with *though*, "which is primarily concessive" (2002: 737). They claim that *as if* may function as a "single compound preposition," which, we believe, denotes a sense of hypothetical comparison (2002: 1151). This sense may be quite "attenuated" (2002: 1151) in certain contexts, and consequently, in some instances, *as if* may be replaced by *that* or its zero alternant without change of meaning, and so may be interpreted as merely suggesting the truth of a proposition rather than (strongly) asserting it. This is especially the case after *appear, feel, seem, sound,* and *be*, which may induce a "medium strength epistemic modality."

Biber et al. (1999: 840-841) claim that with non-finite clauses, *as if* and *as though* indicate that the "adverbial clause is showing similarity but is

not to be taken factually." This is consistent with the analysis we developed above.

We found only two *as if* inferentials in WSC and both were negative so we use a positive *as if* inferential from a written corpus of New Zealand English to begin our discussion of *(not) as if* inferentials.

(16) WWC SECTION F, F42 186-194
The Sunday News used to be the main proponent of the idea of celebrity in New Zealand. It was in that tabloid beloved of life's losers that we first read about Graeme Thorne's perm and much other such trivia. **It was as if successive editors had a list of so-called personalities from which they never really deviated.** It is probably still pasted up in the news-room, slowly yellowing under the harsh fluorescent lights. My guess is that it includes the old names Ray, Bob, Max, Marilyn, the other Ray and Howard. You should know the surnames. They've been around for years.

This inferential can be interpreted as a hypothetical premise from which the celebrity of Graeme et al. would follow, viz., the editors acted as if they drew the names of the celebrities from a list from which they never deviated. Thus this *as if* inferential functions as we predicted.

The inferential in (17), from WSC, shows that *not as if* inferentials also function as we predict.

(17) WSC DPC032
AW: well there's only there's only five or six in the race **it's not as if they're racing going up three wide ra round fields of eighteen** they're only going round fields of six they race sort of there'll be one in the front one <long pause> on the e one on th the trail and one on the outer behind the horse on the on the trail not facing the breeze

In the conversation from which (17) is taken, two people are discussing horse racing and specifically the differences between real racing and practicing. AW's description contrasts *going round fields of six* [horses] with *going round in fields of eighteen* [horses], and rejects the conclusion that in the former the horses are racing. This interpretation is supported in the utterance immediately following the inferential when AW characterizes going round fields of six as only "sort of'' racing. This is consistent with Huddleston

and Pullum's remark that "*It's not as if he wasn't trying* … is used to deny a proposition that might otherwise have been deduced (perhaps he didn't perform as well as expected)" (2002: 1152, fn. 36).

4.5 (Not) like-*inferentials*

Even though our negative occurrences of *like* inferentials outnumber our positive ones, we begin here also with the latter as it is more basic than the former:

(18) WSC DPC326

JI: and she was the legal advisor for er ronnie burch

AL: right <latch>

JI: you know <drawls> when when he was yeah race relations yeah <with creaky voice> mm but um

AL: yeah race relations oh <drawls> good so she didn't have a problem getting a job i suppose when her when er <unclear word>

→ JI: no no but sh she found that particular job very stressful **it's like she doesn't think she'd like to go back into it** you know cos she was always dealing with problems

AL: <drawls> mm right

Like, whose basic meaning reflects similarity, is being grammaticalized as a marker of reported speech and thought (Romaine and Lange 1991). Our analysis is consistent with this, but we suggest that by virtue of its basic meaning and its grammaticalization, *like* in inferentials is a marker of "interpretive use'' (Sperber and Wilson 1995: 224-231). That is, the proposition represented by the clause introduced by *like* is to be interpreted as, to one degree or another, resembling a proposition from which relevant contextual effects would follow. Crucially, the proposition represented by the inferential clause is not a proposition assumed by the speaker, but it merely resembles some such proposition. In (18), the proposition represented by the inferential clause is presented as similar to a proposition which is to function as a conclusion that would follow from the proposition represented by the immediately prior sentence, *she found that particular job very stressful*. This cause and effect relationship is made explicit by the conjunction "cos" which introduces the sentence that follows the inferential, *she was always dealing with problems*.

Not like inferentials reject the relevance of the proposition represented by the inferential clause as a more or less faithful interpretation of the proposition entertained by the speaker. Thus the negative inferential in (19) rejects the potential characterization of the situation as "they NEED someone". Because it is presented in inferential form, this proposition functions as a contextual proposition, in this context, most likely as an explanation, that is, as a proposition from which *They wouldn't want Thomas* would follow. This is consistent with the positive inferential that follows the *not like* one, *It's just they're just doing it as a favor because Susannah's a mate*. We interpret this as a premise from which it would follow that the Wilkins would have AC come *over there … every week*.'

(19) WSC DPC059

AC: well missus wilkins said i could do it over there i mean every week but they they wanted me to do it every week
BS: yeah pity in some ways isn't it because it's quite good money
AC: mm
BS: do you think thomas would do it
AC: they wouldn't they wouldn't want thomas
BS: mm
→ AC: **it's not like they NEED someone** it's just they're just doing it as a favour because susannah's a mate
BS: <drawls> yeah <">
AC: not like they need anyone
BS: i've OCCASIONALLY thought that you could actually do some work for kelvin and sharon but i'm not sure
AC: but they i should just do it for them for free

It is our intuition that in the examples above, *(not) like* may replace *(not) as if*, with only a stylistic shift; we find the *(not) as if* variants to be somewhat more formal than the *(not) like* variants. For example, (16) is from the written portion of the Wellington Corpus of New Zealand English and so we must assume that the author (and perhaps editors) chose their words carefully when they wrote *as if* instead of *like*. But given that substituting *like* for *as if* seems to have no effect on the interpretation of the text, we might assume that the choice merely reflects a different stylistic level. We believe that *like* indexes an informal context because of its origins in casual conversational use by teenagers (though they are not the sole users, see Miller 2009). As a result we find this use of *like* in informal, unplanned discourse and in

representations of such discourse, for example, in fictional renditions of speech. *As if*, on the other hand, cues a higher stylistic level, so we find it in more formal contexts, such as academic prose.

5. What is "fixed" and how might we identify it?

The notion of fixed or formulaic language has come to be used in opposition to that of novel or creative language, that is, expressions which are generated by the syntax and lexis of a language.

In contrast, fixed phrases are not produced with the aid of these rules, they are recurring expressions which are thought to be stored whole in memory, behaving much like single words. Because fixed phrases have come under scrutiny in various fields of linguistics and psycholinguistics, the phenomena appear in the literature under a variety of different labels (47 according to Wray 2002), some of which are given below:

frozen phrases	*idioms*
formulaic expressions/language	*collocations*
fixed expressions	*ready-made chunks*
lexical bundles/phrases	*composites*
epistemic phrases	*recurring utterances*
prefabricated patterns	*conversational routines*

As summarised by Edmonds (2010), two major trajectories have been explored in the understanding of fixed phrases. One of these is a functionally and pragmatically motivated path which focuses on the recurring and fixed nature of these expressions, encompassing work by Altenberg (1998), Bardovi-Harlig (2009, 2010), Coulmas (1979), Kuiper et al. (2007), Pawley (2008), and others (see Edmonds 2010: 14). Within the functional-pragmatic approach, a fixed phrase is an invariable or minimally variable expression which conventionally uses a particular turn of phrase in a specific situation, community, or culture (or combination of these). For example, *fish and chips* refers to a specific type of meal, where the potatoes are cut into strips of a given size (which may vary across countries) and then deep-fried, and the fish is usually a piece of cod which is battered and served (ideally) wrapped in a newspaper sheet, different and distinct from *chips and fish* (which could be any kind of chips together with any kind of fish cooked in any way).

A further example is the use of the phrase *to be someone's shout* which in New Zealand English refers to someone's turn to pay for the following round of (typically) drinks. This phrase is used conventionally in this way only by members in (or familiar with) the New Zealand community, and standing up on the way to the bar in a British pub uttering the phrase *it's my shout* would cause some confusion.

A second strand of research concerning fixed phrases comes from psycholinguistics, developed in work by Biber – Conrad – Cortes (2004), Weinert (1995), Wray (2002), Wray – Perkins (2000), and others. These studies privilege the stored aspect of fixed phrases, with the main characteristic being the claim that they are retrieved whole from memory and thus behave like single words. Prime examples include *by and large*, whose internal grammar is frozen and no longer in line with the grammar of modern English, sayings such as *nice guys finish last* and *life is like a box of chocolates*, or greetings and endings like *how do you do*, *may I help you*, and *yours sincerely*.

The large and growing level of interest in fixed phrases is reflected in the existence of the *Yearbook of Phraseology* series. Despite the vast body of work to date (only some of the works have been mentioned in the brief summary above), fixed phrases remain slippery and elusive in that there is arguably no widely-agreed upon definition or term (though the label *formula* is perhaps gaining the most support), or a definitive set of criteria for identifying these recurring expressions.

In spite of the lack of a consensus, some criteria have been proposed as generally characteristic of fixed phrases and are thus useful in identifying them. Like many other notions in linguistics, fixedness is not a binary feature, but rather encompasses a continuum between something which is more or less formulaic (or more or less novel, depending on how one looks at it). Hence these criteria are not a necessary and sufficient set, but rather a set of heuristics which help identify those expressions that occupy the more fixed end of the continuum. The criteria are listed below in (A) – (K), and are based on the detailed summary of the literature presented in Edmonds (2010: 19-31):

(A) Multiword/multimorpheme
(B) Invariability
(C) Frequency
(D) Community-wide use
(E) Situational boundedness
(F) Syntactic coherence

(G) Semantic opacity
(H) Noncompositionality
(I) Discourse planning and greater fluency
(J) Complexity
(K) Overextension

Criterion (A) has to do with whether or not a single word may count as a fixed expression. Opinions diverge on this issue. The debate is not relevant to the present study, as the inferential matrix is a multiword expression anyway. Criteria (J) and (K) relate specifically to learner uses. And because the data analysed here are exclusively produced by adult native speakers, these two properties cannot be investigated in this study, and will not be discussed further. Therefore, the relevant criteria to our work on inferentials are those given in bold, from (B)-(I).

Table 2 below lists these eight criteria, exemplifying what an idealised fixed phrase might be like, relative to each of these properties.

Table 2. Fixed phrase criteria relevant to the current study

Criterion	If expression X is fixed then…	Example
1	2	3
Structural and/or lexical invariability	X is largely invariant, though some open slots are permitted	*What's this A doing in my B*, where A and B can refer to any NP which fulfills the pragmatic requirements
Frequency	X occurs with relatively higher frequency than expected (e.g., more frequently than it could have occurred given the topic, context, etc. at hand)	*I don't know what* (Biber et al. 1999: 996)
Community-wide use	X is known to an entire speech community, not just to restricted sets of speakers	On the other hand
Situational boundedness	X is associated with a particular situational context, such as a social or pragmatic situation	*How do you do*? is associated with greetings

1	2	3
Syntactic coherence	X does not cross constituent boundaries and applies to full phrases, NPs, VPs, PPs, full clauses, etc.	All things considered
Semantic opacity	The meaning of X is not clear from its component parts	*Kick the bucket* has two meanings, one of which (the idiomatic one) is not transparent from its component words
Noncompositionality	The grammar of X is frozen and may be at odds with current patterns of the grammar of the language it belongs to	*By and large* is no longer in line with grammatical patterns of modern English
Discourse planning and greater fluency	Because X is fixed, it functions much like a single word, it is uttered more quickly than other phrases might be, it is not interrupted by pauses, discourse markers, false starts, etc.	*How do you do* is typically uttered without any pauses, discourse markers or interruptions, and faster than other four word phrases

The first observation to be made is that some of the criteria in Table 2 are not easily operationalisable. In particular, the frequency criterion is notoriously difficult to apply as it is rare that anyone can accurately ascertain just how many times a given phrase could have occurred in a text. In practice, this criterion often becomes a question of collocationality instead (e.g., *kick* and *bucket* more often occur together than not).

Second, these criteria cannot be used in a vacuum, but must be applied to an expression whose fixedness we want to ascertain within a given corpus or data set. We cannot know whether a particular expression has greater fluency or whether it is situationally bound unless we examine how it is actually used. The full range of possible uses of an expression may not always be fully borne out in actual use. Complementarily, real linguistic documentation may surprise us in that the use of a given expression may be wider than previously realised.

Third, it must be remembered that few phrases will exhibit all of these criteria. The majority of expressions with fixed tendencies will exhibit some of these properties but not others. As might be expected, the more of these criteria an expression meets, the more certain we can be that it is fixed, and the more fixed it is likely to be.

6. Are inferentials fixed?

As mentioned in the introduction, we conducted an investigation of inferentials such as those exemplified in (1)-(3) with the aim of establishing their best syntactic analysis. Our findings, discussed in detail in Calude – Delahunty (2011), suggest that the inferential is a type of cleft. However, while exploring the properties of inferentials in WSC, we developed the hypothesis that inferentials might be fixed, at least in spoken English. (We cannot say anything about inferentials in written English.) The reasons for this hypothesis are threefold.

First, our attention was drawn to the fact that the majority of inferentials seemed to have a recurring, relatively invariant structure of the form [*it's (just) that* S]. Consider the following examples from WSC:

(20a) WSC DPC129

CH: hey by the way can i borrow a pair of your earrings <,> **well it's just that** <laughs> <latch>
RG: it depends <,,> which pair
CH: oh well which pair are which pair are you going to wear and i'll borrow another <latch>
RG: the new ones i just bought today
CH: oh well then how about i wear those big ones with the bl with the crystals

(20b) WSC DPC136

AT: why mum was complaining about how GRUESOME the murder was
BD: oh yeah and then they exorcised the demon out of this guy who last episode that was quite fun
AT: i'm sure she loved that <laughs>
BD: and it was really weird um it <,> **it's just that they don't know what happened** but they actually DO know what happened all the sprinklers went on because um you don't know any of the characters <laughs> so it's really hard to explain this to you but but lucy was having this conference with andy and dick who are the possible fathers of her child
AT: sounds like molly dodd <laughs>

(20c) WSC DPC138

AA: oh that's all he said cos it at first i was just saying about how i jim's about the <,,> only one i've had problems with

BC: <drawls> oh right yeah
AA: cos like jim's the only one <,> that won't always <latch>
BC: yeah i know he's just harder to get on with
AA: yeah oh **it's just that i don't think that jim does it on purpose** i think it it's him <latch>
BC: it's just him oh yeah
AA: yeah it's his personality <latch>

Second, similar constructions have been argued to be formulaic (particularly in spoken data). A study of demonstrative clefts in the same corpus argued that this cleft type is fixed, encompassing a formula with a few open but predictable slots, and a specific discourse function (see Calude 2009a, 2009b). The clauses in bold in (21a) and (21b) are typical examples:

(21a) WSC, DPC096 (from Calude 2009a, ex. 9, p. 65)
BG: oh no she is lovely she's gossipy though
AT: mm
BG: very gossipy like bill **that's where Bill get's it from**
AT: <unclear word> oh he is a little gossip talking about Mike Furley

(21b) WSC, DPC214 (from Calude 2009a, ex. 18, p. 69)
BH: the brace helps to hold you upright <,,,>
UV: the only thing for a sore back is bed rest
BH: well **that's what they say** eh
UV: yup
BH: and heat

Similarly, Hopper and Thompson (2008) show that English *wh*-clefts and extraposed clauses as well as German *wenn*-clauses are also fixed. These complex clauses behave more like monoclausal units than like biclausal complexes and should thus be analysed as "single, partly formulaic clauses deployed by speakers in managing interactional discourse" (2008: 99). For example, the pseudoclefts in (22a) and (22b) start off a set of instructions or explanations, and have an invariant recurring pattern:

(22a) (from Hopper – Thompson 2008, ex. 9, p. 6)
So then what you do is, you sprinkle the fifth-graders out evenly.

(22b) (from Hopper – Thompson 2008, ex. 13, p. 7)
What we do, then that's … that's where the ferrier comes in.

Note that the recurrence of *you* in (22a) would render it ungrammatical in a highly edited text, and that there is no grammatical connection between the introductory *wh*-expression and the clause it prefaces, contrary to what would be expected under a biclausal analysis. These types of clefts were also noted in the Map-Task Dialogue corpus of spoken Scottish English analysed by Miller – Weinert (2009), and are so widespread in spoken language that they have now entered textbooks of English grammar (see Miller 2000 and 2011).

Third, as the body of work just mentioned reiterates, spoken language in general (and spontaneous spoken language in particular) appears to make systematic use of formulaicity and the recycling of structures and phrases. In this medium, many constructions are simplified and invariant, consisting of a set of predictable patterns associated with specific discourse-related properties and interactional characteristics. The conditions under which spoken language is produced and parsed, and the functions for which it is used have a substantial effect on its character. Large scale analyses (see for example Miller – Weinert 2009 and Biber et al. 1999) document a reduction in the variety of forms and structural integratedness in spoken language in comparison with written language. The most affected expressions are complex constructions, particularly those involving subordination. These are 'reduced' to only a few possible patterns, in part due to the decreased cognitive loading required for their encoding and parsing, and in part, to their acquiring specialised interactional functions.

Given these observations, we might expect our analysis of the inferential cleft in spontaneous spoken conversation excerpts from the WSC to reveal that it is at least partially formulaic, thus categorizing it with pseudo-clefts, demonstrative clefts, extraposed clauses, and German *wenn*-clauses. We test our hypothesis using the criteria we listed above.

As we noted, the great majority of inferentials can be partially described by the formula [*it's (just) that S*]. We say "partially" here because such linear expressions omit whatever hierarchical organization might be present, and the specification that *it* is expletive. However, it is noteworthy that *it* and *is* are contracted in every example in our data where they may be, which we take as support for our hypothesis.

This formula may be adapted to include the remaining inferential types:

(23) It (modal) (not) BE (adv) (conjunction) S

implying that expletive *it*, the copula, and the complement clause are necessary and sufficient elements. That we can describe them with formulae of this sort suggests that inferentials may be fixed, though we must add that

it and *be* are the minimum required to focus an expression, which is the purpose of the sentence type.

Frequency. Our data show that inferentials occur 55 times in approximately 250,000 words, which is approximately 220 times per million words of conversational English. We also know that they occur in other registers, though we do not know how frequently. Biber et al. (1999) identify a lexical bundle as "a recurring sequence of three or more words" and that four word bundles must occur at least ten times per million words, though longer bundles occur less frequently, and must occur in at least five different texts. As inferentials occur far more frequently than Biber et al's. criterial frequency, we might conclude that they are at least as fixed as lexical bundles. However, Wray (2002) cautions against relying on frequency as a fixity criterion.

Community-wide use. Our data show that inferentials occur in spoken New Zealand English and Koops (2007) shows that they occur in conversational US English, so we can reasonably conclude that they are incommunity-wide use. However, as the grammar and vocabulary of English are in community-wide use, we would have to conclude that any expression generatable by the grammar would be fixed, a patently absurd conclusion.

Situational boundedness. An expression is situationally bound if it is consistently used for a particular social or pragmatic purpose. Inferentials perform a single pragmatic function, viz., asserting or denying the special relevance of the proposition represented by the clause. When we combine this criterion with the relative invariability of inferentials, we conclude that they are partially fixed.

Syntactic coherence. Inferentials are generatable by the grammar and lexis of English and so are syntactically coherent.

Semantic opaqueness. The semantics of inferentials is a function of their lexis and syntax and is therefore transparent.

Noncompositionality. Because the semantics of inferentials is transparent, it is compositional. This criterion, the opacity criterion, and the syntactic coherence criterion together indicate that the inferential forms are generatable by the grammar and lexis of English and that their meanings are a compositional function of those forms, and therefore indicate that inferentials need not be fixed.

Fluency. We have found several instances in which the inferential matrix is interrupted by fillers, e.g., *it's um that i mean public transport …*; *it's just you know <,> it was like <leave me alone>*, suggesting that the matrix is generated or at least generatable analytically and so not stored and produced holistically as a single word-like unit.

Discourse planning. Because fixed expressions are likely to be relatively frequent and therefore to come readily to mind, we should reasonably expect them to be produced at points in discourse where speakers need space to plan what they are about to say. Given that, we should also expect that the utterance that follows the production of a fixed expression need not be grammatically integrated with it, as Hopper and Thompson (2008) discovered with English *wh*-clefts and extrapositives and German *wenn*-clauses. Because inferentials have a specific pragmatic or discourse function and are composed of elements that are frequent and easily processed, they should be readily exploitable in this way. However, inferential matrixes do not have the intonational contour Hopper and Thompson observed with the expression types they studied and which "projects" further talk by the producer, and we have so far not found examples that display the lack of connectedness between matrix and complement that Hopper and Thompson found. This too suggests that inferential matrixes may not be fixed.

How are we to interpret the fact that our application of these fixity criteria to inferentials gives us inconsistent results? Because inferentials are describable with a partially fixed formula with very limited possibilities in its variable slots, and they are in frequent, situationally bound, community-wide use, they appear to be formulaic. However, their matrixes are syntactically coherent, are semantically compositional and transparent, are motivated by the need to focus a clause, may be interrupted by fillers, and do not display either the kind of intonational separation of matrix from complement, intonational projection of the complement, or grammatical independence of the complement from the matrix that Hopper and Thompson identified in their target expressions, which we use as a benchmark for formulaicity, suggesting that they may be constructed as needed rather than stored and produced whole.

It seems to us that being describable as a formula, relative frequency, situational boundedness, and community-wide use do not entail fixity. Rather they are conditions that are consistent with fixity and which perhaps predispose expressions to fixity. We think that English inferentials, at this point in their history, are not fixed, though they may be on the cusp of becoming so, as the consistency with which *it* and the copula are contracted suggests. This development is also indicated by the fact that their matrixes are created out of the kinds of linguistic items – *it* and the copula – that would lend themselves to fixation and eventually to reduction and univerbation. That is, items which occur frequently, are readily accessed and processed, but are non-salient (Hudson 1998).

7. Conclusion

Spoken inferentials fail several of the tests for formulaicity: they are generatable by the current grammar of English; they are semantically transparent and compositional; their pragmatic and discourse effects are predictable from their forms and meanings; they are not always produced fluently and without interruptions. However, though predictable from their grammar, inferential matrixes are lexically very limited, allowing only a very few possibilities. All inferential variants have an expletive *it* subject, a form of *be*, and a complement, which are essential for their interpretation and function; they allow limited tense variability on the copula, the possibility of modals (though only two show up in our data), and a very restricted set of adverbs and clause introducers. They are relatively frequent and are situationally bound because, however we define or delimit them, they have a specific discourse purpose. To resolve these inconsistencies we suggest that they may be at the beginning stages of fixation.

REFERENCES

Sources

Bauer, Laurie
1993 *Manual of Information to Accompany the Wellington Corpus of Written New Zealand English*. Wellington: Department of Linguistics, Victoria University of Wellington. (http://khnt.aksis.uib.no/icame/manuals/wellman/INDEX.HTM).

Holmes, Janet, Bernadette Vine, and Gary Johnson
1998 *Guide to the Wellington Corpus of Spoken New Zealand English*. Wellington: School of Linguistics and Applied Language Studies, Victoria University of Wellington. (http://khnt.aksis.uib.no/icame/manuals/wsc/INDEX.HTM).

Special studies

Aijmer, Karin
2002 *English Discourse Particles*: *Evidence from a Corpus*. (Studies in Corpus Linguistics 10.) Amsterdam: John Benjamins.

Altenberg, Bengt
1998 "On the phraseology of spoken English: The evidence of recurrent word-combinations". In: A. P. Cowie (ed.) *Phraseology: Theory, Analysis and Applications.* Oxford: Clarendon, 101-122.

Bardovi-Harlig, Kathleen
2009 "Conventional expressions as a pragmalinguistic resource: Recognition and production of conventional expressions in L2 pragmatics", *Language Learning 59*, 755-795.
2010 "Recognition of formulaic sequences in L2 pragmatics". In: G. Kasper et al. (eds.) *Pragmatics and Language Learning*. Vol. 12. Honolulu: University of Hawai'i, National Foreign Language Resource Center, 139-160.
Bauer, Laurie
2007 "Some grammatical features of New Zealand English", *New Zealand English Journal 21*, 1-25.
Biber, Douglas et al.
1999 *Longman Grammar of Spoken and Written English*. Essex: Longman.
Biber, Douglas – Susan Conrad – Viviana Cortes
2004 "If you look at…: Lexical bundles in university teaching and textbooks", *Applied Linguistics 25*, 371-405.
Calude, Andreea S.
2009a "Formulaic tendencies of demonstrative clefts in spoken English". In: R. Corrigan et al. (eds.) *Formulaic Language*. Vol. 1: *Distribution and Historical Change*. Philadelphia: John Benjamins, 55-76.
2009b *Cleft Constructions in Spoken English*. Berlin: VDM Verlag.
Calude, Andreea S. – Gerald P. Delahunty
2011 "Inferentials in spoken English", *Pragmatics 21,* 307-340.
Collins, Peter
1991 *Cleft and Pseudo-cleft Constructions in English*. London: Routledge.
Coulmas, Florian
1979 "On the sociolinguistic relevance of routine formulae", *Journal of Pragmatics 3*, 239-266.
Declerck, Renaat
1988 *Studies on Copular Sentences, Clefts and Pseudo-clefts*. Leuven: Foris.
1992 "The inferential *it is that*-construction and its congeners", *Lingua 87*, 203-230.
Delahunty, Gerald P.
1984 "The analysis of English cleft sentences", *Linguistic Analysis 13*, 63-113.
2001 "Discourse functions of inferential sentences", *Linguistics 39*, 517-545.
Delahunty, Gerald – Laura Gatzkiewicz
2000 "On the Spanish inferential construction *ser que", Pragmatics 10*, 301-322.
Edmonds, Amanda
2010 *On the Representation of Conventional Expressions in L1-English and L2-French*. Ph.D. dissertation, Departments of French and Italian and Linguistics, Indiana University.
Fraser, Bruce
1999 "What are discourse markers?", *Journal of Pragmatics 31*, 931-952.
Hedberg, Nancy
2000 "The referential status of clefts", *Language 76*, 891–920.

Heggie, L.A.
1998 *The Syntax of Copular Structures*. Ph.D. dissertation, Department of Linguistics, University of Southern California.
Hopper, Paul – Sandra Thompson
2008 "Projectability and clause combining in interaction". In: Ritva Laury (ed.)
Crosslinguistic Studies of Clause Combining: The Multifunctionality of *Conjunctions*. Amsterdam: Benjamins, 99-123.
Huddleston, Rodney – Geoffrey Pullum
2002 *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
Hudson, Jean
1998 *Perspectives on Fixedness: Applied and Theoretical*. Lund: Lund University Press.
Koops, Christian
2007 "Constraints on inferential constructions". In: G. Radden et al. (eds.) *Aspects of Meaning Construction.* Amsterdam: John Benjamins, 207-224.
Kuiper, Koenraad et al.
2007 "Slipping on superlemmas: Multi-word lexical items in speech production", *The Mental Lexicon* 2, 313-357.
Miller, James
2009 "LIKE and other discourse markers". In: P. Peters et al. (eds.) *Comparative Studies in Australian and New Zealand English: Grammar and Beyond*. Amsterdam: Benjamins, 317-338.
2000 *An Introduction to English Syntax*. (Edinburgh Textbooks on English Language.) Edinburgh: Edinburgh University Press.
2011 *A Critical Introduction to Syntax*. (Continuum Critical Introductions to Linguistics.) London: Continuum.
Miller, James – Regina Weinert
2009 *Spontaneous Spoken Language: Syntax and Discourse*. Oxford: Oxford University Press.
Nevalainen, Terttu
1997 "Recycling inversion: the case of initial adverbs and negators in Early Modern English", *Studia Anglica Poznaniensia* 31, 203-214.
Pawley, Andrew
2008 "Developments in the study of formulaic language since 1970: A personal view". In: P. Skandera (ed.) *Phraseology and Culture in English*. Berlin: Mouton de Gruyter, 3-48.
Pusch, Claus
2006 "Marqueurs discursifs et subordination syntaxique: La construction inférentielle en français et dans d'autres langues romanes". In: M. Drescher – B. Frank-Job (eds.) *Les marqueurs discursifs dans les langues Romanes: Approches théoriques et méthodologiques*. Frankfurt: Peter Lang, 173-188.

In preparation "Pragmatic markers involving subordination in Romance: Do they structure discourse or comment on it?"
Quirk, Randolph et al.
1985 *A Comprehensive Grammar of the English Language*. London: Longman.
Romaine, Suzanne – Deborah Lange
1991 "The use of LIKE as a marker of reported speech and thought: A case of grammaticalization in progress", *American Speech* 66, 227-279.
Schmid, Hans-Jorg
2009 "Rare but contextually entrenched: The English *not-that* construction". Paper presented at the *International Conference on the Linguistics of Contemporary English*, University of London.
Sperber, Dan – Deirdre Wilson
1995 *Relevance: Communication and Cognition*. Oxford: Blackwell.
Weinert, Regina
1995 "The role of formulaic language in second language acquisition: A review", *Applied Linguistics 16*, 180-205.
Wray, Alison
2002 *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.
Wray, Alison – Michael Perkins
2000 "The functions of formulaic language: An integrated model", *Language and Communication 20*, 1-28.

APPENDIX

WSC annotations

<drawls>	speaker drawls
<latch>	overlapping speech
<laughs>	speaker laughs
<„>	longer than 1 second pause
<,>	1 second pause
<reads> … </reads>	portion given between tags was read by the speaker
<quickly>	speech portion is uttered quickly
<quietly>	speech portion is uttered in a quiet voice
<softly>	speech portion is uttered softly
<unclear word>	speech is inaudible or incomprehensible
<with creaky voice>	speech portion is uttered with a creaky voice