CAN YOU SAY “THAT” AGAIN?
The status of direct and indirect speech as grammatical categories

David Nathan

ABSTRACT
Standard accounts of English reported speech (RS) in terms of a binary typology -
direct versus indirect speech - fail when freely produced RS is analysed. Assumptions of binary syntactic encoding (based on clause structure) and binary semantic value (based on lexical identity or nonidentity) are inadequate. This is the first study to compare multiple subjects’ freely produced narrative RS with controlled antecedent utterances. Particular attention is given to elicitation methodology. The study finds that in formulating RS, speakers make choices among verbs, discourse particles, and deictic elements to achieve narrative effect.

1. Introduction
This paper is about reported speech - utterances which purport to bear a relationship to
some other utterance.¹ My aim is to show two things:

(i) that the actual relationship must be investigated by comparing pairwise
occurring utterances; and

(ii) that there are not just two such relationships, those conventionally called
direct and indirect speech.

The actual relationship between reporting and reported utterances has been ignored by
linguists who have taken reporting utterances and merely projected (that is, invented)
antecedent reported utterances from them. This kind of circular analysis has allowed the
description of reported speech to be treated as a priori a syntactic matter.

Analysis of controlled reported speech data shows firstly that direct and indirect speech
are not linguistic categories determined by the form of their matrix clause, and secondly
that their speech complements are too variable, too short of identifying features, and
sometimes too contradictory in features for the purpose of distinguishing an utterance as
direct or indirect speech.

I intend to argue that direct and indirect speech are not formal linguistic categories.
Neither are they categories at a discourse level of description. Instead, they represent the
polar extremes among patterns of choices made by speakers to create the information
flow and structure of their discourse. These choices are among linguistic subsystems of
deixis, lexis, tense, and ordering.

Two methodological issues influenced the work which led to this paper. For reported
speech, it is crucial that spoken data be used to support analysis wherever possible. This
is so because written data is pre-analysed in its use of quotation marks (present for

¹ This paper draws from my revised MA. Dissertation (Nathan 1988). The currency of these issues is
highlighted by the repetition in a recent paper by Clark & Gerrig (1990) of the view that direct and
indirect speech (recast as depiction and description) are grammatical categories. I have benefited from
discussion on these issues with Gerald Gazdar, although both the general approach and the particulars
of analysis are my responsibility. I am also indebted to many individuals who provided the elicited and
spontaneous material analysed here.
direct speech, absent for indirect). Quotation marks are orthographic, not linguistic, devices (Nunberg 1990), and they derive from, while appearing to support, the standard direct/indirect account.

The second issue is that “explanatory” semantic accounts, in terms of whether a speaker uses the exact words of the antecedent utterance (direct speech) or paraphrases them (indirect), are entirely vacuous unless they either systematically compare real reporting and reported utterances, or at least provide a discourse-level account of the effects of a speaker appearing to be verbatim, or merely paraphrasing. Several authors have taken the latter course (eg. Wierzbicka 1974, Coulmas 1986b); this is the first paper, to my knowledge, which has taken the former.  

Note that reported speech refers here only to utterances which claim to represent an utterance (or potential utterance) which is not the speaker’s current utterance - it is not extended, for example, to the report of thought or intentions (cf. Banfield 1982, Leech & Short 1981).  

The first part of the paper is about data and methodological issues. It describes the course of experimental design and the finding that eliciting reported speech data involves more than asking subjects to recall a test dialogue. Emerging from the experiments - and also from other sources - is a significant corpus of reported speech data.

Section 3 of the paper takes a descriptive, atheoretical tour of some linguistic features which bear on the putative distinction between indirect and direct speech, giving a brief standard account, but drawing out some hidden implications. For example, the distinction between direct and indirect speech is almost never marked in reporting matrix clauses. Neither does the lexical content of “quoted” clauses allow us to distinguish direct and indirect speech in traditional terms: turning to these speech complements, I show that the distinction between direct and indirect speech is frequently blurred.

In section 4, examples from the controlled elicitation corpus are used to discuss claims about matrix verb types, the “juncture” between matrix clauses and speech complements, the use of deixis, the inappropriateness of the direct/indirect distinction, and the degree of “veracity” with which reports are given.

2.  Eliciting reported speech data

2.1  The corpus

The data presented and analysed here come from several sources. The main source is a collection of tape-recorded and transcribed stories spoken by experimental subjects derived from a cartoon stimulus. Examples in this paper from this corpus are tagged with the label “E”, together with a number identifying the subject, eg. [E1 3]. Typical examples of the stories are found in Appendix 2. One transcription was made from a pilot experiment, and examples from this corpus are labelled “B”.

Clark & Gerrig (1990:796) report one piece of unpublished research along these lines, but the nature of their elicitation and data are not described.

Neither is it extended to the case of orthographic “scare quotes”, unlike the confusion exhibited by Clark & Gerrig (1990:771) who include in their analysis examples like: These are not ‘I really should’ radishes; these are …
The second source is a seven-page transcription of a tape-recorded spontaneous dialogue. This corpus is labelled “C”. The third source is a set of opportunistically collected utterances which occurred in ordinary conversation. These examples are tagged with “MS”, and a few written examples drawn from newspapers etc. are tagged “N”.

The corpus is too large to be included in an Appendix here, but is entirely transcribed and indexed. Where possible, examples from the corpus, rather than made up ones, are used. Corpus examples are identified via their labels (see above).

There are few special transcription conventions. Although some authors have stressed the importance of prosody as a feature of direct speech, prosodic modulation is conspicuously absent in all my recordings, and so denser coding of such information has not been required. Examples are transcribed using standard English orthography. The only special symbols are “/” to signify a speaker’s repair, “[...]” to frame comments, labels etc., “..” a pause, and “...” for ellipsis of recorded material. The only significant problem is that the use of quotation marks, while necessary for clarity, presupposes an analysis which is often arguable.

2.2 Description of experiments

2.2.1 Pilot experiments

Several pilot experiments preceded the final design. In each case, written or spoken stimulus materials containing antecedent “utterances” were prepared, subjects were given time to listen to or study the stimulus, and were then asked to speak on the basis of their experience of the stimulus. Responses were tape recorded and transcribed.

A first pilot used a stimulus consisting of a written dialogue, a dyadic telephone conversation between a music teacher and his pupil. It was written according to parameters which were felt to reflect ordinary conversation, in particular using heavy reliance on pragmatic meaning and negotiation of interpretation across turn-taking to present the issues to be resolved in the conversation. A subject was asked to simply read and then recount the conversation. The result was an extensive corpus containing many reported speech structures (labelled in this essay “B”). However the subject’s understanding of the stimulus material was so limited that her recounting had no narrative structure at all but was rather a disconnected meta-commentary on her experience of reading and trying to understand the dialogue.

Several conclusions emerged from this pilot. A written stimulus reflecting ordinary conversational principles where pragmatics play a large part in constructing interpretation, may be rather opaque. While a telephone conversation is neat way of presenting a dialogue for subsequent reporting, it also subverts a subject’s comprehension, since normally telephone conversations do not have third parties. The written telephone dialogue stimulus did not sufficiently involve the subject.

The second experimental design used as stimulus a short, spontaneous, recorded, dyadic conversation. It was selected on the basis of a checklist of features, presence of which were predicted to have significance for elicited reported speech. The checklist included explicit speech acts (including questions), verbs of cognition, reported speech itself, swearing, and variation of tense. The results were ambivalent. Some reported speech structures were elicited, but data again tended to favour a meta-discourse centering on

---

4 Extracted from the recordings collected for an unpublished essay.
interpreting the experience of listening to the tape, rather than simply saying what happened. Subjects said that the material was “boring”. So a similar design was used, but with the substitution of the authentic stimulus by a simulated “gossipy” conversation about a marital issue within the University residences from which the subjects were drawn. Since it had also become apparent that eliciting reported speech entails a subject reporting to someone, I played the segment to the subjects, who were then recorded recounting the conversation to another person. The subjects for this pilot did become more involved with the content of the stimulus, and their responses were original and narratively structured versions of the original conversation. Unfortunately, though, few reported speech structures were elicited.

These pilots indicated that for subjects to freely use reported speech structures, they must have personal involvement not merely with the content of the stimulus conversation, but also in its production. They must be interested parties to the original conversation, whether as speakers, addressees (in conversation, or as a radio listener etc.), or overhearers (cf. Clark & Carlson 1982).

2.2.2 The cartoon elicitation

The subjects for the final experiment were all adult, native English speakers, although they came from diverse English-speaking communities. The procedure was designed to elicit reported speech structures within a narrative format. A specially constructed cartoon (see Appendix 1) was used as a stimulus. There were 19 subjects. Each subject was first told the procedure: they were told that the experiment was not aimed at testing their memories, but at finding out how they told stories. They would look at a cartoon (for about a minute, or as long as they wished), and then, after it was withdrawn, they would be asked to say what had happened. Their responses to my question “What happened?” were tape-recorded. I remained facing the subjects, providing a potential addressee for them, and gave appropriate “back channel” cues (Brown & Yule 1983:92) where appropriate. Later the tapes were transcribed and analysed.

2.2.3 Discussion

The aim was to collect several reported speech-containing narratives recalling a single original conversation. It is not easy to get speakers to be linguistically involved across two texts. The solution was to use a stimulus of greatly reduced explicitness and content. A frame-by-frame cartoon (see Appendix 2) was created, representing an episode in the life of a character named “Bill”. Speech is represented using conventional “balloons”, and designed to provide the basis for a coherent event as well as on the basis of the feature checklist (see above).

The procedure described in 2.2.2 was successful in eliciting a comprehensive range of reported speech structures within narrative structures. Virtually none of the subjects commented on the form of the stimulus. The reason for its success is that the intertextual problem was dissolved; subjects were creatively involved in constructing an original interpretation for the narratively underspecified cartoon during their reading of it.

There are other factors contributing to the success of this procedure. Paradoxically, it lacks the contrived flavour of more “naturalistic” stimuli - while these are more linguistically authentic, this is subverted by the experimental context. Texts may be authentic, but the subjects’ experience of them can turn out otherwise. The cartoon, however, proclaims its own fictional status. Further, cartoon format demands that subjects use their own knowledge of the world to interpret it, and so narratives are
elicited which interpolate the motives and personality of characters in terms of subjects’ personal and cultural frameworks.

It needs to be clear, however, that the elicited narratives are not merely artifacts of the cartoon itself. While in 4.4.2 it is shown that several subjects used remarkably similar linguistic devices in framing one particular conversational move of their narratives, generally they made the facts and sequence of their stories suit schemas which relate to their own narrative style and knowledge in the domain of the cartoon. The cartoon is not a narrative - it is an antecedent to a constructed interpretation.

Subjects produced many features not present in the cartoon. The most significant linguistic features are the narrative cohesion produced through the choice of appropriate referring terms for characters mentioned only by pronouns in the cartoon, and also the synthesis of information found in disparate cartoon frames. Subjects supplied roles to participants (e.g. employer/employee), embroidered the basic script (cf. Schank & Abelson 1977) of their stories (with mentions of oversleeping, getting in and out of cars etc.), and added personality profiles of the main character. They also routinely failed to mention depictive aspects of the cartoon not important for their narratives.

3. Syntax and reported speech

3.1 Definitions

The standard view of reported speech in English is that there are two semantic/syntactic modes: indirect speech, in which the gist of a previous utterance is repeated; and direct speech, where an original utterance is reproduced verbatim (Leech & Short 1981:318, Palmer 1986:134). In the following pans, I describe grammatical features of reported speech, and the extent to which they reflect this binary semantic classification, more-or-less along orthodox lines. Problems in basic assumptions emerge: these are addressed later in section 4 which also takes a closer look at the experimental data.

An instance of reported speech involves the tacit indexing of another (usually previous) utterance. The only evidence for which original utterance is involved is the content of the reporting utterance. Since utterances are not indexed, the difference between gist and verbatim does not necessarily involve formal, linguistic criteria and is fundamentally a semantic one. The difficulties in determining what counts as “saying that X” go beyond establishing ‘criteria for type-token identity between different instances of X’ (Lyons 1977:117). For example, verification of an utterance like “Nixon said that eating muesli is un-American” depends not only on finding some way of determining agreement about the paraphrase relationship between this proposition about patriotism and the breakfast menu and what was originally said, but also depends on the identification of the original speaker and the time, place etc. of utterance.

Generally a speaker’s self-attributed part of a reporting utterance consists of the main, or matrix, clause, and I refer to this part as the framing (cf. also McGregor, to appear). The reported part, to avoid confusion with the term “reported speech”, will be called speech content. In some cases, we need to distinguish different classes of speech content. The actual or putative antecedent utterance - the one being reported - is original speech, spoken by the original speaker. (By analogy we can also refer to the addressee and the original addressee.)

Thus, for the sentence above:

| speaker: | current author |
| framing: | Nixon said on the 13th November 1972 that |
| speech content: | eating muesli is un-American |
| original speaker: | Richard Nixon |
For English (and many other languages), there appear to be grammatical correlates of the gist/verbatim dichotomy:

(1) Tom said that he was disgusted with Will’s behaviour
(2) Tom said “I am disgusted with Will’s behaviour”
(3) I’m disgusted with Will’s behaviour

The sentence in (3) is original speech, the sentence type in (2) is an example of direct speech, and (1) is indirect speech.

3.2 Framing clauses and complement types

3.2.1 Speech verbs

Reported speech generally consists of a framing clause, and speech content. Framing clauses give the report status of the speech content, and may include information about the identity of the original speaker, place and manner of original speech, identity of original addressee(s), and perhaps something of the nature of the original utterance itself - whether, for example, it was a question, an order, or an opinion.

The core of the framing clause is a speech verb, or verbum dicendi (Banfield 1982:23; Munro 1982:304 calls all such verbs “say’ verbs”). In general, the presence of a speech verb is obligatory (cf. 4.2.3): the most commonly occurring are say and tell

(4) She said “I’ve written you a letter” [C2]
(5) He told the boss he’d be in early [E7]

English has a large number of speech verbs; examples in my corpus include insist, relate, go, remark, promise, ask, decide, and also order, suggest, reply, show, opine, write.5

Most speech verbs can be used to frame either direct or indirect speech, and the highest frequency verbs are certainly found in both cases (see 4.2.1). However, there are some which are restricted. Hear can frame a “hearsay” report, or in formal legal reporting, and only occurs with indirect speech:

(6) I heard you’re going to Greece. [MS27]
(7) A US Navy sailor killed his wife by throwing her into shark infested waters off the Indian Ocean, a court-martial heard. [N18]

Manner-descriptive speech verbs such as stammer and blurt occur only with direct speech (cf. Fonagy 1986:264). The only commonly occurring speech verb (see 4.1.1) which is restricted is go. Go can only be used to frame direct speech (9), although without speech content go refers to an act of speaking:6

(8) Please go more slowly. [Request to speak slower, MS23]
(9) I can imagine E. going “Oh this bloody C. woman!” [MS6]

Go is often used to frame speech content which is expressive (cf. 9), unusual, poorly-understood, in a different language, or even non-linguistic:

(10) He went something like “Who likes wheat germ anyway?”

---

5 The second group is not in the spoken corpus but found amongst written materials. Verbs are given in non- finite form but may appear in other forms - tensed, or as gerundives or nominalisations.

6 Names have been suppressed in corpus examples where politic
Brian went “Ngayalu palyanyinanyi.”

Puss went “miaow-w-w”.

My Hoover goes “clonk clonk clonk”.

Note the decreasing possibility of substituting other speech verbs for go in moving from (10) to (13).

3.2.2 Framing syntax

The syntax of reported speech framing clauses is generally identical to the grammar of main clauses (cf. Leech & Short 1981:3 18, Silverstein 1985:153).

I told John the story.

I told the story to John.

*I told the story John.

I told John that Nixon hated muesli.

*I told that Nixon hated muesli to John.

The conditions which govern the acceptability of these sentences are paralleled by the widely described dative alternation (14-16) and “heavy NP shift” (17-18). Examples (14-16) obey the same grammatical conditions on the appearance of direct and indirect objects as benefactive or transfer verbs such as give, send, sell, as long as the speech content is regarded as direct object and the original addressee as an indirect object.

Example 17 illustrates a structure similar to the result of “heavy NP shift”, where a ‘relatively long and complex’ object noun phrase appears at the end of a sentence (Winograd 1983:510). However, for reported speech the criteria for “heaviness” are fully grammaticalised: sentences such as (17-18) which contain speech content never allow the indirect object (here, John) to be separated from the rest of the matrix clause. This constraint also applies to speech verbs such as say which only optionally have an “addressee” noun phrase - in the speech corpus used for this study there are no examples of the type I said ... to her. Such sentences are always expressed:

She said to Jan “It might still work out.” [C40]

She said to Jan that she knew I was really angry. [C56]

It thus seems that (i) the framing clause of reported speech must be contiguous, that is, not interrupted by speech content; (ii) speech content has a grammatical role similar to direct object noun phrases (cf. Coulmas 1986b:20, Clark & Gerrig 1990:77 1, contra: Munro 1982); and (iii) these observations about framing clauses are independent of the direct/indirect speech dichotomy.

3.2.3 Speech complement types

While the direct/indirect typology is barely reflected in the framing of natural, spoken reported speech, there do appear to be constraints on framing clause syntax dependent on which of three types of generalised speech complement appears.

The three types are:

SUPPRESSED DESCRIPTIVE

SPEECH ACT COMPLEMENT

DIRECT OR INDIRECT SPEECH
The first type is shown in (14-15) above, where the report content is ‘suppressed’ (Banfield 1982:42), and reduced to a simple descriptive noun phrase. Such sentences have fairly general syntactic properties, including the possibility of clefting (21) and passivisation (22):

(21) It was a story that John told.
(22) That story was told to me by my mother.

The second type consists of reported speech sentences where the report content is reduced not to a simple noun phrase, but to the simplest expressions of certain speech acts - agreement, denial, greeting and leave-taking. This type appears to be intermediate between the suppressed descriptive structures and full direct/indirect speech content structures:

(23) ?I told her yes/good-bye.
(24) *I told yes to her.
(25) ?I said to her yes/good-bye.
(26) I said yes/good-bye (to her).
(27) We said our good-byes. (cf. Coulmas 1986c:165).

Although (23) and (25) could be written with quotation marks, or uttered with appropriate intonation, they would be heard as verbatim reports, i.e. direct speech. In the other examples, yes and good-bye are not necessarily the (only) original words spoken, but unlike the suppressed descriptive case above, express possible, indeed typical, words used to express agreement, assent, leave-taking etc. (Silverstein 1985:139).

Sometimes there are collocational restrictions between the speech verb and its “speech act complement”. Note the cases illustrated in (28-31), and the parallel between the “speech act complement” examples and simple speech verbs in (32-33):

(28) We said (the) prayers.
(29) *We told the prayers.
(30) *We remarked the prayers. . .etc..
(31) We intoned the prayers.
(32) That bureaucrat always says no/yes to our requests.
(33) That bureaucrat always refuses/agrees to our requests.

The third type, in (19-20) above, consists of reported speech with true “speech content”, either a direct speech report of an original speaker’s words, or an indirect report paraphrasing the original proposition. Here, speech content is less like a direct object, and has been described as a loosely bound complement (Banfield 1973, Munro 1982). In the case of tell, it is the addressee argument which functions as a direct object, since (i) only the addressee is obligatorily expressed (cf. 34-35); and (ii) its speech content (except suppressed-descriptive) cannot be the subject of a passive construction (cf. 36.37). While some passive structures are possible with say and tell, these contrast with those with descriptive noun phrase and simple speech act speech complements, in that only the addressee noun phrase or an extraposed it - and not the speech content - can be the subject of the passive sentence (cf. 38-39):

(34) Mary was angry. Later, she told Bill.
(35) Bill was visiting. *Mary told that she was angry.
(36) *That corn flakes were nutritious was said/claimed.
“Wow! Let’s eat corn flakes!” was said across the land.
People were told to eat corn flakes.
It was said that corn flakes were nutritious.

Thus the syntax of framing clauses depends more on the choice among the three speech complement types than according to whether the speech content is interpreted as direct or indirect speech.

3.3 Complementisers
Traditionally, the presence of the complementiser that is held to be a marker of indirect speech. Here I include a larger set, including that, and the “Wh-” words what, where, whether, how etc., as well as if, and later (4.3) enlarge the set under the more general notion of juncture.

Nixon said that eating muesli was un-American.
Nixon said (*that) “Wow, how I love corn flakes!”
Nixon told us what we should eat.
*Nixon told us what should we eat.
Nixon asked if we could all eat corn flakes.
*Nixon asked if could we all eat corn flakes.

Wh- words can also occur at the opening of speech content in direct speech. In fact, Wh-speech content reports are a paradigm case where the form of the speech content is the basis for deciding whether a reporting utterance is direct or indirect speech. It is auxiliary inversion in the speech content which marks the utterance as direct speech. So in (43) and (45), what and if cannot be complementisers but must be part of the speech content of direct speech sentences. (43) is a report where Nixon is heard to ask our preference: it is not a report of his edict (but unacceptable because of the choice of the speech verb tell7). (45) is ungrammatical because if could we is not a grammatical sequence beginning the speech content.

Thus, while complementiser distribution is related to the direct/indirect distinction, the decision-making process for distinguishing (42-45) as direct or indirect speech sentences must begin with the properties of the speech content, not the framing clause.

While some (low occurrence) speech verbs such as remark obligatorily require a that complementiser when not postposed, often the complementiser seems to be missing, or ‘suppressed’ (Coulmas 1986c:163). There appears to be, in Banfield’s (1982:71) words, a ‘superficial’ difference in framing between the following:

He told them he’d run out of petrol. [E13]
He told him that he’d got stuck. [E20]

This apparent optionality of the complementiser is part of the same general phenomenon of complementiser use in general English described by Thompson & Mulac (n.d.). Thompson & Mulac analysed a spoken corpus and described the conditions under which that was present or absent. They found that the factors in main clauses (ie. in reporting utterances, framing clauses) which condition complementiser appearance are elaborations of content beyond simple, bleached “epistemic phrases”.

---
7 Tell presupposes a proposition has been understood by the addressee in an original speech event; cf. 4.2.2.
More importantly for reported speech, where the bulk of spoken examples have semantically unelaborated main clause propositions, Thompson & Mulac found a feature of complement clauses (here, speech content) which is a significant determinant of the presence of *that*. If the complement clause subject is a pronoun, then the complementiser that is only half as likely to appear than if the subject appears as a full NP. Their finding not only reinforces the claim that it is the speech content which drives the direct/indirect distinction, but also correlates with the analysis of my corpus and the claims for juncture in 4.3.

3.4 Speech content: direct and indirect speech

3.4.1 Non-deictic distinctions

In the previous section we saw that despite, the existence of lexical and syntactic constraints on *framing* clauses, these have little role in distinguishing direct and indirect speech. The features which distinguish direct and indirect speech lie in speech content. A widely cited marker of *direct* speech is that of speech content ‘in the form of a performative speech act such as a command or a question’ (Li 1986:37) - see the discussion of (41-45) above. The same restriction applies to speech content which is non-propositional, expressive, incomplete, or presented as a dialectal or non-standard variation (see also examples illustrating *go* above):8

(48) He says “Oh, ta mate!” [E6]
(49) *He says that oh, ta mate.
(50) David Wharton walked over to where Nehemiah was standing and said, “Now you are going to work, you understand. You are going to pick four hundred pounds of cotton today.” “Wal, Massa, dat’s aw right,” answered Nehemiah, “but ef Ah meks yuh *laff*, won’ yuh lemme off fo’ terday?”

With *indirect* speech, however, we find general conditions on the form of speech content resembling those applying to other complement constructions (Munro 1982:303). So where speech content is not performative or expressive, we have to look elsewhere for clues to distinguishing direct and indirect speech.

3.4.2 Deictic distinctions

3.4.2.1 Tense

Deictic elements encode point of view and orientation and thereby affect the interpretation of speech content as direct or indirect. They are analysed here along two axes: (i) an intrasentential axis, where there are constraints between framing clause and speech content, and (ii) an intersentential axis, where there are constraints on speech content according to the relation between the current speech situation and the original speech situation and content.

---

8 (50) is from Brewer 1968:90. Although Leech & Short (1981:321) believe that such renderings are ‘more faithful representation[s]’, and Clark & Gerrig (1990:783ff) treat them as basically depictive, they are really not phonetically analytical but rather encode a speakers or writers pejorative connotations. It is instructive to try to find parallel non-standard orthographic rendering of NON-stigmatised dialects or registers (see also Harris 1986 who contests the notion that alphabetic writing enables phonetic rendering of language).
Traditionally, a “shifting of tense” rule has been postulated as a sentential rule for indirect speech in English (Jespersen 1924:293). The rule is that the tense of speech content must, if necessary, be shifted so that the speech content is interpreted according to the time deixis of the framing speech verb. Because reported speech typically reports past events, speech content tense is back-shifted.

The effect of the rule is to constrain the possibilities for expressing tense in indirect speech. While it is claimed that tense shift is required to make the interpretation of the speech content subject to the deictic datum-point established by the tense of the framing clause (Coulmas 1986b:17), Comrie has stated a simpler formal rule for English whereby the tense of the supposed original speech is retained in indirect speech, except where the framing verb is past, in which case the speech content is backshifted (Comrie 1985:114). Thus, if John were to say, while driving a car, “I am speeding”, an indirect report will show a tense shift only if the reporting sentence has a past tense frame:

(51) John said that he was speeding.
(52) *John said that he is speeding.
(53) John will say that he is speeding.
(54) *John will say that he will be speeding. (cf. Comrie 1985:114)

Thus a sentence like:

(55) I told my boss I was going to be there on time. [E14]

does not necessarily mean that the original promise of punctuality was made in the past tense.

While the tense “rules” are thought of as an intrasentential aspect of indirect speech (main verb tense constrains speech complement tense), they are formulated in terms of two relationships: firstly between framing clause and speech content (intrasentential), and secondly between reporting utterance and original speech (intersentential). The latter relationship is circular where the original speech is merely projected from the reporting utterance (or the pairs are simply invented), undetermined in grammatical terms even if there is a real antecedent utterance (see 3.1), and is a relationship which has never before been empirically investigated.

3.4.2.2 Verb and adverb deixis

Various other deictic expressions such as time and place adverbials, and motion verbs, are “constrained” by both intrasentential and intersentential principles. There is lexical modification of the putative original utterance so that deictic expressions can be interpreted relative to the deictic datum-point established by the reporting utterance. For example an original utterance “I don’t like the decor here” could be reported as:

(56) He said he didn’t like the decor there.

when reported elsewhere than the original (distasteful) location (here would only be retained if the report takes place at that original location). The same principle operates for motion verbs and time adverbials. The true direct speech purported to be reported by:

(57) I’m going to phone and say I’m going to go tomorrow night. [MS5]

could be “I’m going to come tonight” if spoken to a person at the destination on the morning of travel, “I’m going to come now” spoken immediately before departure, or “I’m going to go tomorrow night” spoken to a person who is not at the destination. There are several other possibilities. Note that unlike tense shift, this form of “deictic
adjustment” involves changes of lexical material rather than grammatical elements, and thus offers much more potential for expressing resolution of time and space.

3.4.2.3 Pronominalisation

Pronominalisation establishing speaker/addressee deixis is the least ambiguous distinguisher of direct and indirect speech. Direct speech repeats the pronominal or nominal forms of the original utterance, so that in:

(58) She said “I’ve written you a letter” [C2]

the interpretation of you is insulated with respect to the current speech situation but conveys the putative form used by an original speaker. Assignments of referents is with respect to participants specified (or understood as specified) in the framing clause. Thus I is, as in true direct speech (see 3.1), the original speaker, predicated in the framing clause as being the referent of she. Similarly, you, the canonical addressee, is the person originally addressed by the referent of she. You may be coreferent with the utterance addressee, but this is undetermined and is a matter of contextual interpretation.

Thus in direct speech, speech content has as its deictic datum-point the original speech situation, which is in turn indicated by the framing clause. It is in this sense that direct speech is said to involve its utterer in ‘playing the role of the reported/original speaker’ (Li 1986:38).

By contrast, consider the following indirect speech sentence:

(59) She said I hadn’t told her if I was definitely coming or not. [C12]

Here the speech content pronouns are interpreted with reference to the current utterance situation; that is, the speech content expresses the proposition “speaker hadn’t told her ... .” But the proposition itself is attributed to the original speaker (the referent of she).

Indirect speech draws out an asymmetry in the pronominalisation of conversational roles (speaker/hearer/other). One result is that third person pronouns are commonly ambiguous; however such cases ‘follow the general rule of pronominalisation’ in complement clauses (Li 1986:32):

(60) Jan just said that she’s very vague. [C11]

(61) Jan just decided that she’s very beautiful.

In both cases, she may refer either to Jan, or to another unspecified person, who may be the original, but not the utterance, addressee.

However, where an indirect reporting sentence is itself framed as direct speech (unlike (59) where an indirect report is indirectly framed), the speech content pronouns are interpreted according to the speech situation encoded in the direct speech framing:

(62) His boss said” You said you’d be here before nine.” [E12]

Here the interpretation of you (2) is with reference to the outermost direct speech framing clause, that is, the boss’ addressee. This interpretation is independent of the intervening pronoun you (1), since substituting I, she, or Harry for you (1) does not alter the referent of you (2). The following examples (63-64) show how a change from canonical direct to indirect speech shifts pronoun interpretation from the original to the utterance situation. However note the similar patterning between examples (63) and (65). In (65) the pronouns are interpreted according to an “original situation” which is
not ambiguous, despite the fact that there are two speech verbs and ultimately three speech situations:

(63) The boss said “I should dismiss you.”

\[ ^1 \quad ^2 \]

(64) The boss said that I should dismiss you.

\[ ^3 \quad ^4 \]

(65) The boss said “Harry said that I should dismiss you.”

\[ ^5 \quad ^6 \]

Pronominal interpretation in direct speech, even where the pronoun occurs within “embedded” indirect speech content, is thus fully dependent on the identification of the original participants specified in the framing clause. Indirect speech is, however, ‘transparent’ (Munro 1982:303) to pronominal reference, and is distinguished by its ability to separate reference made in speech content propositions from the attribution of those propositions to particular speakers. It is in this separation of reference from propositions that the propositional nature of indirect speech noted by so many authors has its origin.\(^9\)

### 4. Results and discussion from data analysis

#### 4.1 Introduction

This section describes aspects of the elicited data. It presents empirical results, with quantitative results given in relation to specific corpuses, and also discusses real examples of the complex selections among lexical and grammatical elements that speakers bring to their formulation of reported speech to serve ongoing narrative sequences.

The first part describes framing clauses - especially their verbs - and there is a discussion of the way that *tell’s* presuppositional properties say far more about its occurrence in discourse than its grammatical properties could ever hint at. Later there is description of a tentative juncture analysis, which transcends grammatical analysis by offering a discourse-oriented pattern which is part of a compositional account of reported speech.

#### 4.2 Framing

##### 4.2.1 Speech verbs

Occurrences for framing verbs are given in Table 1. In some instances there is no associated speech content (eg. *thank*, cf. 3.2), and in other cases there is speech content but no framing verb (labelled as ‘0’ in the table). Citation forms are shown, ignoring tense, aspect etc. Only those verbs with at least two occurrences are listed (the table accounts for over 90% of corpus speech verbs; a few others, such as *relate, mention* have only single instances).

---

\(^9\) See also McGregor, to appear, for a similar analysis from a different perspective.
The verbs are overwhelmingly skewed toward *say*, with over 70% of examples. *Tell* also occurs with regularity (11% of examples), and, except for the zero-predicate strategy (see 4.2.3), other speech verbs are numerically insignificant.\(^\text{10}\)

While *tell* may seem to have a secondary prevalence because it (unlike *say*) can take a non-finite speech complement (as in orders: “I told him to ...”), this turns out not to be the case - in fact only three such examples occur. Similarly, *tell*’s ditransitivity - its obligatory addressee expression as well as speech content (see 3.2) - might seem relevant; however there are 16 examples of *say* with an overt addressee (amounting to 15% of all *say* verbs, matching the total instances of *tell*):

(66) She said to Jan last night that she was feeling really bad. [C3]

*Say* and *tell* have different occurrences in direct and indirect speech. While *say* to <addressee> clauses have “normal” distribution for the whole corpus (approximately two-thirds direct and one-third indirect speech, although *say* is so numerous that it defines “normal” here), *tell* occurs overwhelmingly with indirect speech. These results suggest that *say* is the generic, unmarked speech framing verb in English. The intermediate status of *tell* is discussed in the following section.

### 4.2.2 On *Tell*

Fonagy (1986) has distinguished between speech verbs which are deictic and those which are descriptive. Deictic verbs are those which frame direct speech by point[ing] to the reported utterance’ (ibid:262, also Partee 1973:416f). Speech content is treated as a referent, as if the framing clauses always contain an implicit demonstrative direct object:

(67) Nixon said this: “Muesli is un-American” (cf. 3.1)

so that in direct speech the speech content is not grammatically a direct object but is integrated anaphorically (actually, cataphorically), and is mentioned rather than used (Gorovitz & Williams 1965:116).

*Descriptive* speech verbs are used to describe ‘the circumstances of the speech act’, and Fonagy classifies them as a secondary lexis (Fonagy 1986:264f). Examples are *cried*, and *sighed*; these verbs serve as speech verbs by assimilating to the grammar of the

---

\(^\text{10}\) The occurrence of *promise* is thought to derive from the specific content of the experiments stimulus dialogue.
primary speech verbs - the verb of saying is deleted, and the verb referring to a simultaneous event absorbs its meaning (ibid:272-3).

Adopting this analysis here would involve labelling say as a primary speech verb, and those such as insist, remark as secondary, with tell, being barely descriptive, falling in between. This status for tell reflects a bleached reference to an act of “making sense of”:

(68) She can tell the time.
(69) She will tell your fortune.
(70) She can tell if the lines converge.

None of these examples entail an utterance. Addition of addressee arguments introduce such entailments, although strictly, any form of communication might do, as long as “sense is made”:

(71) I sort of said “Well what!- why can’t you tell me?” She said “I’ve written you a letter.” [C1-2]

Here tell means cause to know”: C’s interlocutor will tell C something (by writing a letter), but she can’t say it. In fact, later in the same text, the speaker herself provides an analysis, saying “you might as well just sort of say it to me” ([C24]).

Tell with an addressee expression presupposes transmission a proposition. If someone phones me and gives me a message to deliver to her husband, I confirm my intention to cooperate by replying “Yes, I’ll tell him (that)”. However if I say “Yes, I’ll say that (to him)”, I am not acknowledging that my understanding of the message is what is at stake. By apparently stating that I am a conduit for words, rather than propositions, I am implicating (cf. Grice 1975) that I have either not understood the message, or that I do not have regard for its truthfulness or value.

Kartunnen (1973) has drawn attention to a semantic analysis where speech verbs are classed as “plugs”, preventing the presuppositions of their speech complements applying to the whole sentence. This is not unequivocal for tell, as demonstrated in negating or questioning framing clauses. In most circumstances, (72) implies Harry’s success; (73) does not:

(72) Harry didn’t say (to her) that he managed to win.
(73) Harry didn’t tell her that he managed to win.

In 3.4.2.3 we found that indirect speech allows separation of reporting speakers’ attribution of propositions from their acts of referring. With our analysis of tell, we can tentatively turn the formulation around, and see that its presuppositional behaviour could be the source of its occurrence as an indirect speech framing verb (see Table 1). Since tell involves the current speaker presenting a presupposition about the original speech situation, the current discourse relevance of tell’s speech content is highlighted, and reference made in tell’s speech content becomes an act of the current speaker.

4.2.3 Zero framing predicates
Although Banfield (1982:23) claims that a sentence’s status as reported speech is conferred by its speech verb, there are several examples in my corpus of reported speech content unframed by verba dicendi. A few are examples of the type known as free indirect speech, where the illocutionary content is expressed directly but ‘tense and
pronoun selection are those associated with indirect speech’ (Leech & Short 1981:325, see also Jespersen 1924:292, Banfield 1982, Clark & Gerrig 1990:786):\footnote{Banfield (1982:68) claims that free indirect speech occurs only in written, not spoken, form.}

(74) Looks at his petrol gauge .. got any juice left?.. it’s the last thing he needs. [E10]

Direct speech is more common, occurring either unframed (75-78), or framed simply by a noun phrase referring to the original speaker (79):

(75) Put it in, change-over alternator, you know, “fitted seventy five dollars.” [D9]

(76) the petrol starts to run out. “Oh shit!” So he gets out.. [E14]

(77) he looked at the clock. “Oh god, I’m going to be late.” [E15]

(78) I was feeling all jolly.. and “let’s put the kids to bed early” [MS 13]

(79) Jan said that she really knows best.. which I could believe she probably is ..but Jan, you know “She’s just like!- she’s a very soft person as well ..“ [C25-26]

Most interesting are two examples where direct speech occurs framed by a non-linguistic representation of turn taking. The first involves turning the head (the first speech content is attributed to the speaker’s daughter, the second is the speaker’s self-reported response):

(80) ‘But I love you!’ [turns head 900] “Go away!” [MS2]

In the following example, there is an implicit report which, paradoxically, consists only of the “frame” interpreted from the rules of speaker sequence derived from speech act constraints on turn-taking (cf. Sacks, Schegloff & Jefferson 1974; Labov & Fanshel 1977:70). The speaker’s father is reported as having asked a question (for which, in the current speech situation, the addressee could anticipate what the father’s response was going to be, or rather how it was going to be reported). Then, after a short pause (approximately one second), the father’s reaction to the (unspoken) reply is reported. The current addressee interpolates the existence and content of the reply:

(81) and my father said “Oh yeah, how far away from America is that?’..[1 second] “Okay!” [MS1]

4.3 juncture

Many writers, discussing reported speech, have not distinguished between written and spoken data (eg. Partee 1973, Banfield 1982). Some have treated only written (eg. Fonagy 1986), or spoken (eg. Silverstein 1985) data, while others have clearly separated them (eg. Leech & Short 1981, Clark & Gerrig 1990). However even for spoken data, it is often the categories and distinctions of written material which provide the parameters for analysis (Coulmas 1986b:10). This may be due to written forms themselves being seen as the canonical mode of reporting, being fixed so re-enactable, and content-oriented (Tannen 1982:14).

Both analysts and everyday hearers of reported speech need to decide not only the deictic centre and reference of speech content, but also its boundaries. Earlier in this paper, as elsewhere, boundaries were treated as unproblematic since clear examples, some with quotation marks representing \textit{a priori} analysis, were used.

Analysis of spoken reported speech leads to the notion of \textit{juncture} - a system of lexical and grammatical features which marks the transition between framing clause and speech content.
(82) below gives the flavour of the juncture system by summing up actual reporting structures used by the main speaker in corpus “B”. Notice the different speech complement types (cf. discussion in 3.2): “quote” corresponding to direct speech, “5” to indirect speech, and the element hello which is an instance of the simple speech act complement. Note also the elements of the first disjunction (headed you know), which are further discussed below.

\[
(82) \quad \text{...<speech verb> \{ (to __) (umm) \{ you know \text{ well that if \} err sort of okay \} \} \{ Quote \} } \]

hello

What does the juncture system include? Elements include the complementisers discussed in 3.3. There are also the discourse particles (Levinson 1983:162, Schiffrin 1987) well, you know, oh and look; and the lexico-grammatical elements that, if and the presence of Wh- or an inverted auxiliary marking a question (labelled “÷Q” in the table below). Juncture markers form a system whose elements have complementary distribution - we do not find examples of the form “She said THAT ‘YOU KNOW...’”, “I said ‘YOU KNOW DO 1...’” etc..

Table 2 elaborates on (82) and gives the frequency for all juncture elements in corpus “C”, together with a judgement about whether the instances are direct or indirect speech.\(^{12}\)

<table>
<thead>
<tr>
<th></th>
<th>Direct</th>
<th>Indirect</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>well</td>
<td>10</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>0</td>
<td>6</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>that</td>
<td>-</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>you know</td>
<td>-</td>
<td>2</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 3: Direct Reported Speech Junctures, Clark & Gerrig (1990)

<table>
<thead>
<tr>
<th></th>
<th>Direct</th>
<th>Indirect</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>well</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+Q</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>oh</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>you know</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>no</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>egads</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gee whizz</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It is not always possible to assign the non-syntactic juncture elements (well, you know etc.) to one or other of the framing clause or speech content. The elements which can occur in this position are discourse oriented: they correspond well to the set which

\(^{12}\) By applying deictic criteria, especially pronominalisation, discussed in 3.4.2.
Schiffrin (1987) has described as discourse markers: *oh, well, I mean, you know* and others. These do not occur in grammatically governed positions, but at locations which guide the interpretation of pragmatically defined units of discourse (Stubbs 1983:68, Schiffrin 1987:310. Discussion of units of discourse is beyond the scope of this paper; however note that analyses which identify the system used by speakers to manage conversation with the juncture elements described here have a natural, iconic extension to reported speech since both direct and indirect speech structures represent, to differing degrees, the discourse notion of *change speaker*, the fundamental feature of conversation management (Sacks, Schegloff & Jefferson 1974).

The claim here is not that boundaries are located simply by observing the position of the juncture elements. The elements *signal* the presence of a boundary, but may themselves be best analysed as part of the framing clause (eg. *that, you know*), as syntactic or discourse elements of the speech content (eg. inverted question auxiliaries and *look*, respectively), or wholly undecidable (eg. *well*).

Turning specifically to the juncture *well*, we find it is widely used in marking the beginning of conversational turns (Sacks, Schegloff & Jefferson 1974). The *same* conversational competence can be exploited by speakers/hearers to signal/determine the *same* change in deictic datum-point in direct speech utterances; and it is *well* and the other discourse particle juncture elements which occur preferentially with direct speech. Schiffrin (1987:125) and Levinson (1983:162) sum up studies showing that speakers use *well* to signal topic divergence. For reported speech it must be the current speaker who comments on sequential relationships between reported topics, because it is only the speaker who selects and orders the material she reports (see 4.4.2, Labov 1972). The following example shows *well* used to mark topic divergence: note that the divergence is with respect to the exposition of the current speaker’s narrative, but the original speeches actually occurred in two separate conversations!:

(83) I said ... “I’ll work down here and I’ll pay the rent on two flats” ... and then she rang me, she said “Well, save me having to do that ..”  [C52-54]

Like the use of the final anaphoric *that* in (83), *well* helps create narrative coherence. It signals a relationship between conversational moves. But its use here is the *current speaker’s* device rather than partial verbatim reproduction of the original conversation(s).13

4.4 Speech content

4.4.1 Direct or indirect speech?

While the distribution of juncture elements corresponds fairly well with the direct vs. indirect speech distinction, the parameters by which the distinction was defined earlier in this paper often fail to distinguish which of the two types a given utterance should be classed as.

The tense-shift “rule” for indirect speech (3.4.2.1) is not triggered in the many reported speech utterances framed in the historical present (“she says ..”), . In other cases, there is no tense shift where the reporting speaker believes that the speech content proposition is

---

13 It is thus inadequate to handle analysis of *well* by transcribing it within quotation marks, so deeming it part of speech content (cf. a number of examples in Labov and Fanshel 1977).
Currently valid (Comrie 1985:114). This occurs where reported speech is used to display a plan or obligation (example 57; and below), or for statives such as know (86):

(84) She was saying that you do a first year but you don’t choose your main subjects until you’ve finished .. [B 10.3]
(85) The woman said they’ll be here by nine-thirty. [MS26]
(86) Claire said to me.. that she doesn’t know how she’s going to afford the room [C9]

While pronominalisation has a more central role in distinguishing direct and indirect speech, it occasionally does fail to establish the appropriate person-deictic centre. There are several examples of this in my (natural, conversational) corpus, but none, apparently, was misunderstood by the addressee(s):

(87) He was sort of saying that if I see her father ... [MS24]
(88) It wasn’t fair because “you knew I was coming and you should’ve known.” [MS4]
(89) You said what dress am I going to wear? [MS 16]
(90) .. she said to me do I mind if she put this girl in [C47]

In (87), the juncture that would indicate indirect speech, yet I is coreferent with the framing he, so that there is a change of person deixis as in direct speech. Context indicated that the original addressee was the current speaker, and that her father was actually the current speaker’s father, (ie. the original addressee’s father). We might have expected to see your father, or my father. The speaker’s selection of her is motivated by the narrative context: the original speaker was “going wild”, making a very public display. with a weapon and also with his loud words. The change of pronoun deixis (he to I), suggesting direct speech, casts current hearers into the interpretive frame of the original hearer(s). But the choice of her impersonalises the hearer role, and conveys the public nature of the original speech situation by creating a wider, unspecified deictic centre. The described event is made vivid for the current hearers because they are momentarily cast in the role of the original audience.

In (88), the speech content is not explicitly framed, and the deictic shift interpreted by the hearer (indicated, incidentally, by some perhaps ad hoc quotation marks here as elsewhere) is informed by the lexical content following because understood from the narrative context.

In (89), the surface form would indicate a directly quoted question, but there is no change of deictic centre since the speaker is standing in full view before her wardrobe, and the current hearer is aware of her own recent antecedent question about the speaker’s sartorial intentions. (89) fuses the immediacy of direct speech, indicated by the inverted auxiliary, with a retention of deictic centre which in this case presented no ambiguity since all participants were present.

(90) is similar, with mixed pronominal deixis, cast into typical direct speech form.

4.4.2 On determining the veracity of reported speech.

In discussing reported speech some have argued (eg. Labov & Fanshel 1977:339), or taken for granted (cf. Polyani 1982) that the report is faithful to the original words. Expressed in the standard distinction between direct and indirect speech are two glosses for faithful: “verbatim reproduction” and “presenting the gist” (see 3.1).

While it has been argued that veracity is not an important level of analysis for reported speech due to the speaker’s primary exploitation of reported speech structures as a
narrative ‘presentational mode’ (Bauman 1986:65, also Polyani 1982:163, Clark & Gerrig 1990:795ff), studies of the actual mapping relationship between original and reported speech have not been carried out.\textsuperscript{14}

The experiment described in section 2 was designed to supply data which could be used to make an initial investigation of this relationship. Before presenting results, some issues need to be mentioned. The intertextual problem in designing a procedure which could provide the data for comparing original and reporting speech was discussed in 2.2. But even having the data at hand leaves many questions to be answered. What sort of decisions should be made about labelling reported speech as direct or indirect so as to decide whether correspondence of words or sense is required, especially in the light of the earlier finding that the distinction is not really categorical? What should be the directionality of comparison? Must utterances be reported in their original order? How \textit{completely} must an utterance be reported to be considered faithful? And how should we interpret the veracity or otherwise of changes in deictic point of view? How do we interpret \textit{degrees} of synonymy or non-correspondence? What shall we make of non-correspondence anyway?

Only some of these questions can be answered at this stage. Order is relatively unimportant - speakers have a narrative task and their ordering of material is aimed at making the story ‘tellable’ (Labov 1972). Where speakers make explicit claims to order, such as “then he said ..“, the sequentiality issue is a general one not specially relevant to reported speech. Nevertheless, “surface” orderings do have effects on speech content, as indicated by the appearance of the anaphoric \textit{that} (whose referent is a previous reported utterance) in (83).

In order to make sense of the veracity issue, a liberal approach to anaphora and deixis is needed. There are two reasons for this: firstly we must not introduce circularity by reading off from deictic shifts a distinction between direct and indirect speech and then imposing the verbatim/gist criterion for veracity. This would be treating pronominalisation, tense etc. as systems of evidentiality marking. Secondly, anaphoric and deictic elements provide narrative or conversational coherence, partly by marking shifts of deixis to enable reconstruction of antecedent points of view. But such point-of-view elements are not generally at stake where the truth-values of propositions are concerned.

The following discussion of the relationship between original and reported speech, using the experimental results to compare the pairs of utterances, is thus discursive rather than quantitative. What is found is that original speech can appear in a narrative report in a number of guises. They may simply be reported verbatim:\textsuperscript{15}

\begin{enumerate}
\item \begin{enumerate}
\item “Oh shit! That’s the last thing I need.” [cartoon]
\item ... as he said “Oh shit, that’s the last thing I need.” [E19]
\end{enumerate}
\item \begin{enumerate}
\item “You said you’d be here before nine.” [cartoon]
\item His boss said “You said you’d be here before nine.” [E12]
\end{enumerate}
\end{enumerate}

or they may be cast into standard indirect form (93b):\textsuperscript{16,17}

\textsuperscript{14} Refer to footnote 2.
\textsuperscript{15} In this section, original speech is labelled [a], and various subjects reported forms [b], [c] etc.. See Appendix 1 for the original cartoon stimuli showing all original speech content.
\textsuperscript{16} See also 74 in 4.2.3 for an example cast in free indirect speech.
(93) a. “You said you’d be here before nine.” [cartoon, = (92a)]
b. ... because he said he’d be there by nine. [E19]

The following example shows that subjects do not necessarily see indirect speech as giving licence to paraphrase:

(94) a. “... run out of petrol .. near the Bristol turnoff” [cartoon]
b. ... says that he’s broken down .. near the Bristol turnoff. no, says that he’s run out of petrol [E12]

Here the gaps in the original have been filled, but all the original words are reproduced verbatim. On the other hand, there is direct speech which varies considerably from the original:

(95) a. “I told him I’d be in early, LOOK AT THE TIME!” [cartoon]
b. “I’m going to be late, I told my boss I was going to be there on time,” he says. [E14]

What the subject expresses here is to be understood from its narrative context: “I’m going to be late...” follows on from mentioning that the original speaker “Bill” had been frequently looking at his watch. “The boss” has been inserted on the basis of producing a coherent narrative structure - no-one is actually indicated as a boss in the stimulus material but the roles are suggested in the later cartoon frames (see 2.2). “On time” is clearly a paraphrase of “early”, interpreted by the speaker in the light of the boss’s later mention of “nine” (see 92a).

Some subjects even collapse original utterances from a number of frames to produce suitable narrative structure - (96b) is from one of the shorter stories:

(96) a. “I told him I’d be in early, LOOK AT THE TIME!” “You said you’d be here BEFORE NINE!” [cartoon, frames I and 9]
b. ... he’s saying to himself as he’s going along, “Oh shit, I promised I’d be there by nine o’clock.” [E1 1]

Sometimes, re-use of original speech content presents a different point of view:

(97) a. “You said you’d be here before nine.” [cartoon, = (92a)]
b. ... because he said he’d be there by nine. [E19]
c. ... boss says ... “... you said you’d be here before nine.” [E6]

*Mutatis mutandis* for the deictic forms, (97b) is faithful to the original. Notice, though, that the original utterance (97a) is itself reported speech, so that while the *speech content* in (97b) is faithful to the original, such words may never have been spoken by the purported speaker (the referent of *he*). Casting the same speech content as direct speech, as in (97c), yields a perfect mapping. But this does *not* reflect a change along the verbatim/paraphrase axis, but a change of point of view. The utterer of (97b) speaks from the point of view of “the boss”.

Finally, we find original (stimulus) utterances appearing as non-reporting narrative:

(98) a. “... run out of petrol .. near the Bristol turnoff” [cartoon]
b. ... and then he run out of petrol, just near the Bristol turnoff of.. some road. [E7]

17 I will continue to use the terms direct and indirect despite the discussion in 4.4.1; they are used for descriptive purposes.
These examples show that original speech is manipulated in various ways to suit subjects’ own narrative strategies (cf. Coulmas 1986b:6, Clark & Gerrig 1990:797). A surprising result is that there was actually greater similarity across subjects in framing particular utterances than in formulating speech content. For example in reporting the cartoon character’s first utterance (95a above), nearly half of the subjects used the verb promise.

Even more striking is the framing of the character’s final account. Three-quarters of the subjects frame their account with a clause of characteristic syntactic form (the frequency of the components decreases down the page):

Labov (1972, 1977) has noted the narrative element he calls evaluation, the part of the story which provides the point of its telling. The schema in (102) corresponds with the features of evaluation mentioned by Labov (1977:108) - there is a narrative discontinuity, with a lexically explicit “metacommentary”.

In reporting the cartoon character’s final utterance, subjects have a complex task. Firstly, the original utterance itself has a narrative sequence, which has to be embedded within the story. Secondly, the content of the utterance needs to be reproduced to give the story its “point” - the cartoon character’s strategy of misleading his boss. And finally, the fact that the character did lie must be mentioned - albeit implicitly - so that the “point” is not only made, but also, in Labov’s terms, evaluated. It is worth reflecting on the considerable creative achievement by the subjects in weaving together a number of narrative threads at this stage of their story; and it is an even greater challenge to ask why indeed the subjects’ reports converged at this, perhaps the most complex, part of the story.

5. Conclusion
The main finding of this essay is that there is no grammatical distinction between direct and indirect speech. They are not syntactic types, but tendencies emerging from the compositional effects of speech verb, pronoun, juncture, lexis, and tense selection in narrative/conversational context. Where these factors conspire we find the canonical reported speech structures, but in the complexity of performing both narrative and everyday conversation, speakers can manipulate selection so as to achieve much more dense and rapid modulation of point of view than the standard clausal account suggests.

---

18 Labov actually describes two types of evaluation; the type mentioned here is called ‘external evaluation’ (Labov 1977:147).

19 There are also some residual problems. For example, in the discussion of juncture, the role of that was not fully explored. In the corpus, that occurs more often where the framing clause has only third person participants, suggesting that it may function as kind of evidentiality marker, but the data is too thin to draw any conclusions. Also, the case of zero juncture was conspicuously not discussed; again, there are too few examples to support analysis but there is a strong tendency for zero juncture to be followed directly by the first person pronoun. These observations are strengthened by similar ones in a recent paper by Thompson & Mulac (n.d.) interpreted in terms of pronoun topicality.
The notion that direct and indirect speech allow the speaker to play either (and only) the roles of original or current speaker (Li 1986, Coulmas 1986b) is oversimplistic. Reported speech is not so much theatrical in nature (Wierzbicka 1974), as providing the most fertile theatre for the manipulation of these linguistic features.

Clark & Gerrig (1990) have advanced the understanding of reported speech by recasting direct and indirect speech in terms of demonstration and description, respectively. Their demonstration account of direct speech, where only selected aspects of original utterances are depicted, helps attack the verbatim reproduction myth. However, their account adheres to a clausal analysis and a binary typology of quotation. It is yet unclear in what sense demonstration is a fundamental and distinctive communicative act, and ultimately their demonstration theory sits uneasily with their claim that the verbatim assumption is false. The key to detaching reported speech from the verbatim notion is, as I have shown, that the putative reported speech structures are simply particular patterns among the linguistic devices which speakers have at their disposal for creating discourse. While reporting utterances are treated as primarily “saying again” - or demonstrating etc. - then the subtlety of choices available to speakers will not be uncovered.

What has also been shown is the need for a large corpus of suitable data to support the analysis of reported speech. My emphasis on methodological issues reflects this. Eliciting and analysing reported speech data reminds us that whilst language is located in time and space, that time and space are themselves linguistically established.
REFERENCES


Thompson, S. & A. Mulac. n.d. The discourse conditions for the use of the complementizer *that* in conversational English. ms.


APPENDIX 1
THE CARTOON STIMULUS

Bill is late manner...

Oh shit! That's the first thing I need.

Oh boy! What'll I do?

Exam still going. Must call now.

I'm late! How could I have been so stupid?

Tell mom to wake me in the morning.

If you know there was a terrible accident in the N.Y. and I was writing there... but is the tailback...
APPENDIX 2

THREE SELECTED NARRATIVES

[7] Ah well.. Bill was supposed to er be going in early.. to work .. he told the boss he’d be in early, but er I think he overslept. Anyway he didn’t start off too early. Ah got in his car, started driving to work and then he run out of petrol er just near the Bristol turnoff of some road. And er.. what he did then was he just got out, he had to find a phone, er he had to walk quite a long way to get the phone, he was out of breath by the time he got there .. and .. phoned up the A.A.. They sent a man round with some petrol. Finally he got to work, the boss wasn’t too pleased with him, so Bill made up er a story about er there being a multiple car crash on the way and getting caught in the .. the tailback. And that’s what happened.

[18] This fellow was on his way to work .. he was late .. he ran out of petrol .. on the motorway. He walked to telephone box, phoned the A.A., who come along, give him some petrol. He went to work, got a bollocking off his boss, and then he made up some story about why he was late.

[15] Well this guy was .. going to work, and .. had told his boss presumably that he was gonna be on time .. and urn he was in a rush, and .. he looked at the clock Oh god I’m gonna be late and .. raced along the road .. and urn .. looked at the petrol gauge saw that it was on empty .. and .. freaked out [laugh] got into more of a state.. and.. and the car stopped .. got out, looked around for a phone box .. er eventually found one, rang up the A.A. .. told him his position .. ran out of petrol near the Bristol turnoff.. and .. um waited for them to arrive. When they did .. they filled his car with petrol .. he thanked them, and urn .. off he went to J. A. Zong and Company [laugh]. And he was only ten minutes late, but his boss er remarked on the fact that he .. that he said he was gonna be early. And instead of just.. saying what had happened, he just made up this whole other story .. I suppose because he just couldn’t er face the fact that it was just running out of petrol and his kind-of own .. incompetence or his own inefficiency .. and er made up a story that he!- that there had been this er nine car crash on the A71 [laugh] and that there’d been a long tailback .. and er I suppose he thought that was more acceptable to his boss than saying that he’d just run out of petrol.