Toward a semantic account of *that*-deletion in English*

DANIEL DOR

Abstract

The phenomenon of that*-deletion in English has traditionally been taken to be one of the prime examples of syntactic free variation: it has implicitly been assumed that the language arbitrarily allows predicates to embed their propositional complements with the complementizer* that *(as* that*-clauses), or without it (as bare clauses). Different researchers, however, have demonstrated that the complementizer* CANNOT *be deleted under certain classes of predicates. In this article, I show that the distribution of* that*-deletion across the different predicate classes is determined by a single semantic property, which I call "truth claim":* that*-deletion is possible under predicates which semantically entail that a cognitive agent (most often their subject) has made an epistemic claim concerning the truth of the proposition denoted by the embedded clause. In some cases,* that*-deletion is also possible under predicates which pragmatically approximate the meanings of real truth claim predicates. This pattern is explained on the basis of semantic compatibility. I suggest that* that*-clauses and bare clauses have different meanings:* that*-clauses denote propositions, but bare clauses denote what I call "asserted propositions." Only truth claim predicates, and those predicates which can be pragmatically used as such, are compatible with asserted propositions. Consequently, only they are capable of embedding bare clauses. To account for the impossibility of* that*-deletion in non-complement positions, the above semantic theory is supplemented with a complexity-based, adjacency-type constraint on the pronunciation of the bare clause.*

1. Introduction

The phenomenon of *that*-deletion in English has traditionally been taken to be one of the prime examples of syntactic free variation. It has implicitly been assumed in the literature that the language arbitrarily allows

Linguistics 43–2 (2005), 345–382

0024–3949/05/0043–0345

sentence-embedding predicates, such as the ones in (1), to embed their sentential complements with or without the complementizer *that*:[1]

(1) John { said / told me / believed / knew } { that Bill was completely out of his mind. / Bill was completely out of his mind. }

The fact that things are not as arbitrary as they seem to be has been demonstrated by several writers, most notably by Dwight Bolinger in his classic 1972 *That's That*. Bolinger and the others (cf. Thompson and Mulac [1991], and references therein) have demonstrated that the deletion of the complementizer is impossible, or at least unnatural, in certain semantic, pragmatic, or conversational contexts. The most famous example is the fact that *that*-deletion is much less natural under manner-of-speaking predicates:

(2) John { murmured / mumbled / grunted / moaned } { that this guy was completely out of his mind. / *this guy was completely out of his mind. }

Bolinger tried to capture his empirical findings on the basis of a complex cluster of semantic, pragmatic, functional, statistical, and conversational factors. As Bolinger himself admitted, this was not an attempt to suggest an explicit explanation of the facts — an explanation of the type which he was very skeptical about in the first place — but a partial description of the different factors and contingencies which might have a causal role in the determination of the distributional facts.

Other scholars have concentrated on the fact that *that*-deletion in English seems to be restricted to complement position. It has been observed, for example, that *that*-deletion seems to be impossible in subject position (3a) and under derived nominals (3b), and less natural, at least in some cases, when the clause is shifted to the right (3c):

(3) a. { That you should say such a thing / *You should say such a thing } surprised me.

b. I heard Bill's claim { that the sun is a star. / *the sun is a star. }

c. Mary said in a very loud voice { that she would not eat the cookies. / ??she would not eat the cookies. }

Facts of this type have inspired a series of ECP-based analyses of *that*-deletion (Stowell 1981; Kayne 1981; and Snyder and Rothstein 1992), and, more recently, at least two optimality-based analyses (Pesetsky 1998; Grimshaw 1997).

In this article, I will mostly be interested in the Bolinger-type question, that is, the relationship between the possibility of *that*-deletion (in complement position) and the semantics of the embedding predicate. As we shall see, the facts involved in this relationship are much more intricate, and more interesting, than Bolinger had envisaged. In Section 6.3 of the article, however, I will discuss three possible approaches to the syntactic facts in (3), and suggest that we may need to supplement our semantically-based analysis with an adjacency-type constraint on the pronunciation of bare clauses, in the general spirit of Pesetsky (1998).[2] This descriptive constraint, in its turn, may be accounted for in terms of processing complexity, along the lines of Rohdenburg (1996) and Hawkins (2001).

The major hypothesis which I would like to present in this article is the following: The distribution of *that*-deletion (in complement position) in English is determined by a single semantic property, which I will call "truth claim": The predicates which can embed the bare clause, without the complementizer, are those which entail that a cognitive agent (in the majority of cases, their subject) has made an epistemic claim concerning the truth of the proposition denoted by the embedded clause. Generally speaking, cognitive agents can make the truth claim either by THINKING it or by EXPRESSING it. For two obvious examples, take another look at the examples in (1). In the examples involving the predicates *say* and *tell*, John, a cognitive agent, publicly expresses the truth claim. He says that the proposition that Bill is completely out of his mind is true. In his role as the cognitive subject of the predicates *believe* and *know*, on the other hand, John thinks the truth claim. He thinks that the proposition that Bill is completely out of his mind is true.[3] The predicates which cannot embed the bare clause are those which do not entail (and do not pragmatically implicate) that a cognitive agent (in the majority of cases, their subject) has made an epistemic claim concerning the truth of the proposition denoted by the embedded clause. An example may be the verbal expression *read aloud*. A sentence such as *John read aloud (from the paper) that George was engaged in some illegal activity* does not imply that John himself was committed to the claim that George was breaking the law.[4]

In the next three sections of this article, I will present a semantic survey of a long list of sentence-embedding predicates. I will distinguish three different types of predicates on the basis of the notion of truth claim:

(I) Type I predicates: These predicates semantically ENTAIL that a cognitive agent (most often, their subject), has made an epistemic claim concerning the truth of the proposition denoted by the embedded clause. They can only be used to report the public expression or the mental representation of a truth claim;

(II) Type II predicates: These predicates do NOT entail the notion of truth claim, but allow their meanings to be pragmatically EXTENDED to approximate that of a truth claim predicate;

(III) Type III predicates: These predicates do NOT entail the notion of truth claim, and CANNOT be pragmatically used to report truth claim events.

As I will show, *that*-deletion is virtually always acceptable under Type I predicates, and never acceptable under Type III predicates.[5] With the predicate classes of Type II, which sometimes approximate the meanings of real truth claim predicates, patterns of acceptability are, predictably, much more complex. First, speakers' judgments vary to a considerable extent: some speakers are more permissive than others and allow *that*-deletion to occur in a larger set of lexical contexts. Second, individual speakers find it difficult to come up with clear judgments. In some cases, they find that bare clauses are ALMOST acceptable, although they testify that they always prefer the complement WITH the complementizer. In all these interesting cases, the variability in grammatical judgments actually correlates with the degree of vagueness in the meanings of the embedding predicates. The fact that grammaticality judgments vary in these cases, and that speakers find it difficult to make clear judgments about them, actually strengthens the hypothesis that the syntactic facts are semantically-driven. More on this complex issue will follow in the concluding section of the article.

As a first step towards an explanation of the correlation between the lexical semantics of the embedding predicates and the syntactic distribution of *that*-deletion, I will suggest that *that*-clauses and bare clauses, in English, denote two different types of semantic entities, propositions and asserted propositions respectively (cf. Cushing 1972). A proposition is a description of a state-of-affairs, which may or may not have a truth value. An asserted proposition is a proposition which has a putative truth-value, a proposition which is asserted by a cognitive agent to be true. (The etymology of the word *proposition* is actually helpful here: an asserted proposition is a proposition which is proposed by someone to be true.) I will claim that the bare clause, which denotes the asserted proposition, can only appear on the surface under the semantic scope of a truth claim. In other words, an expression denoting an asserted proposition needs to be

explicitly licensed by a truth claim: it can only appear when a cognitive agent is reported to have an epistemic attitude towards its truth-value. The *that*-clause, on the other hand, can be freely used to refer to propositions, as descriptions of states-of-events, whether they are "proposed" to be true or not. I will suggest that this characterization of the meanings of the two types of complements, together with the lexical analyses of the different embedding predicates, determines the relevant distributional patterns on semantic-pragmatic grounds.

A word about methodology: The grammaticality judgments which form the empirical basis for this account are based on both field research and corpus analysis. Grammatical and semantic judgments were collected from about 50 native speakers of English — 30 undergraduate and graduate students at Stanford University, and about 20 native speakers of English who currently reside in Tel Aviv. The quantitative results of the corpus analysis are presented in Section 5.

In the next section, I will discuss Type I predicates, those which entail truth claims and allow *that*-deletion. In Sections 3 and 4, I will discuss the more interesting classes of predicates of Type II and Type III. My goal in these descriptive sections is to establish the correlation between the semantic facts and the syntactic distributions, not to explain it. The classification of the predicates is partially based on Levin (1993), Searle and Vanderveken (1985), Karttunen (1978), Erterschik-Shir (1977), Bolinger (1972), and Zwicky (1971). In Section 6, I will attempt to explain the observed correlation on the basis of meaning compatibility. In Section 7, I will suggest that two additional explanatory components may play a role in further constraining the distribution of *that*-deletion. First, the bare clause may be associated with informal register. Second, considerations of processing complexity may impose an adjacency constraint on the pronunciation of the bare clause.

And a final note: As its title indicates, the article deals exclusively with English. Obviously, other languages seem to manifest different distributional patterns with respect to *that*-deletion. The theory developed in this article, however, allows for what seems to me to be a realistic approach to this question of variability, along the following lines. To the extent that the phenomenon of *that*-deletion is determined in essence by considerations of semantic compatibility, and to a lesser extent by issues of register and adjacency, variability on each of these levels should result in observable variabilities on the syntactic level. First, it goes without saying that lexical items, including verbs of the types this article deals with, may differ in their meanings from language to language, and as we shall see, grammaticality judgments rely heavily on the lexical semantics of these verbs. Second, as some researchers (e.g. Ramchand 1998) indicate,

languages may differ, to a significant extent, with respect to semantic categorization. This, to be sure, is the major insight behind the new wave of research on linguistic relativism (Gumperz and Levinson 1996; Lucy 1992; and references therein). Such semantic variability may strike some as counterintuitive, as semantics is usually thought of in universalistic terms. As far as I am concerned, however, the fact that we find remarkable degrees of variability in all the other levels of linguistic representation indicates that we should expect at least some variability in the domain of semantics. Another insight coming out of the literature on linguistic relativism is that patterns of semantic categorization, whether or not they are strictly shared by different languages, need not be correlated in the different languages with the same syntactic features. This, too, may result in observable variability. Moreover, variability on the semantics–pragmatics interface (as we shall see below, semantic compatibility interacts in interesting ways with pragmatic compatibility), and in the manifestations of register and adjacency, may play a complementary role in the determination of syntactic variability. All this suggest what seems to me to be a very reasonable hypothesis, namely, that the attested variability in grammatical patterns in the different languages is correlated with a parallel degree of variability on the lexical-semantic, semantic, pragmatic, and morphological level. An interesting recent discussion of *that*-deletion and related phenomena in Italian (Giorgi and Pianesi 2000, forthcoming), indicates that this position may be on the right track. As Giorgi and Pianesi show, the complementizer can only be deleted in Italian under *belief*-predicates, whereas *communication*-predicates, such as *say*, disallow *that*-deletion. This pattern differs from the English one on the basis of lexical semantics, and is related to the different semantics of propositional complements in the two languages: Italian, as opposed to English, productively distinguishes between indicative and subjunctive complements. In this sense, identifying the complex set of determinants of *that*-deletion in English should be thought of as the first step towards a typology of parallel correlations in a wider set of languages. I take the exploration of this position to be the major goal for my future research on this topic.[6]

2. Type I predicates

2.1. *Speech act predicates*

Say, *tell*, *claim*, *assert*, *affirm*, *state*, *argue*, *inform*, *remind*, *predict*, *suggest*, *disclose*, *reveal*, *indicate*, *intimate*, *insist*, *hypothesize*, *admit*, *confess*, *divulge*, *show*, *demonstrate*, *make clear*, *point out*, *contend*, *remark*, *note*.

Each of the above predicates is obviously different from all the others, but they all share one central component of meaning: they denote ASSERTIVE illocutionary acts. Searle and Vanderveken (1985: 37) define the ILLOCUTIONARY POINT of assertive speech acts as the attempt "to say how things are ... (I)n utterances with the assertive point the speaker presents a proposition as representing an actual state of affairs in the world of utterance." Another way to state the same intuition — again, in the terms suggested by John Searle and formalized by Searle and Vanderveken — is to say that in a successful and nondefective assertive speech act, the speaker informs the interlocutor of a fit between the propositional content of his or her utterance and the way things are in the world (a "word-to-world" direction of fit). This is the essence of the notion of TRUTH CLAIM: describing cognitive agents as performing the speech acts denoted by the above predicates entails that they made the epistemic claim that the proposition expressed by their utterance is true in the world. Thus, the sentences in (4) entail the sentences in (5):

(4) John { claimed / admitted } that Bill tried to steal the money from the bank.

(5) John { claimed / admitted } that it was true that Bill tried to steal the money from the bank.

The different speech act predicates entail different types of truth claims: the predicate *predict*, for example, entails a future-oriented truth claim, whereas *admit* entails a "reluctant" truth claim.[7] Thus, a negation of one speech act does not imply that the subject has not made any truth claim, but it nevertheless narrows down the set of possible truth claims made by the subject. If (6) is true,

(6) John didn't { claim / admit } that Bill tried to steal the money from the bank.

we are entitled to infer that there is a certain type of truth claim which John did not make (although he may have made a different type of truth claim: he may have *suggested*, or *hypothesized*, that Bill tried to steal the money from the bank). This rather trivial fact will prove to be of importance later on.

As is well known, the great majority of speech act predicates allow *that*-deletion:

(7) John { said / told me / admitted / confessed / suggested / argued } { that George made Bill steal the money. / George made Bill steal the money. }

With others, deleting the complementizer seems to be unnatural, or even unacceptable for most speakers:

(8) John { asserted / hypothesized / affirmed } { that George made Bill steal the money. / ?? George made Bill steal the money. }

2.2. *Predicates of belief, knowledge, and conjecture*

Believe, *think*, *conclude*, *suppose*, *agree*, *maintain*, *fancy*, *presume*, *assume*, *know*, *be aware*, *recall*, *remember*, *forget*, *find out*, *discover*, *notice*, *realize*, *guess*, *conjecture*, *imagine*, *figure*.

The fact that truth claim is a central meaning component of these predicates is rather obvious. They all entail a commitment on the part of their cognitive subjects to the truth of the proposition denoted by the embedded clause. This time, the commitment is not publicly expressed, but mentally represented. The sentences in (9) entail their counterparts in (10).[8]

(9) John { knew / believed / realized / figured out } that Bill was the one who planned the whole thing.

(10) John { knew / believed / realized / figured out } that it was true that Bill was the one who planned the whole thing.

Again, the different belief and knowledge predicates entail different types of truth claims: the predicate *believe*, for example, entails a much stronger truth claim than the predicate *guess*. Thus, a negation of one belief or

knowledge predicate narrows down the set of possible truth claims made by the subject in the mental domain. If (11) is true,

(11) John { think / know } that Bill was involved in the crime.

we are entitled to infer that there is a certain type of truth claim which John did not make (although he may have made a different type of truth claim: he may have *guessed*, or *presumed*, that Bill was involved in the crime).

As far as I can tell, all predicates of belief and knowledge allow *that*-deletion:

(12) John { knew / believed / assumed / realized / found out / remembered } { that Bill was involved in the crime. / Bill was involved in the crime. }

3. Type II predicates

3.1. *Manner-of-speaking predicates*

Babble, *bark*, *bawl*, *crackle*, *chirp*, *cluck*, *coo*, *groan*, *shout*, *scream*, *yell*, *holler*, *bellow*, *whisper*, *shriek*, *wail*, *lisp*, *hoot*, *growl*, *grunt*, *mumble*, *moan*, *howl*, *mutter*, *whine*.

Zwicky (1971) characterizes manner-of-speaking predicates as referring to acts of communication by speech and (describing) physical characteristics of the speech act. It seems, however, that these predicates are better characterized as referring to acts of emission of oral sounds, although they may sometimes be pragmatically interpreted as referring to acts of communication by speech. Some of the most important arguments for this view come from Zwicky's own observations.

First, manner-of-speaking predicates can be used to refer to noncommunicational acts of speaking of all types, including those where the produced signal does not make any communicational or linguistic sense. We can use these predicates to refer specifically to the physical production of the speech signal. This is impossible with the real communication predicates:

(13) Mary { shrieked. / hollered. / hooted. / *said. / *claimed. / *told. }

Second, as Stowell (1981) has observed, the object of manner-of-speaking predicates seems to refer to "the physical noise produced by the act of speaking," as opposed to "the propositional content of the message" (Stowell 1981: 401). (Ungrammatical examples with real communication predicates are quite impossible to construct):

(14) John { shrieked a shriek. / whispered a nearly inaudible whisper. }

So, to begin with, manner-of-speaking predicates denote different acts of producing sounds of different types, not necessarily sounds of speech. However, these sounds may sometimes correspond to linguistically well-formed utterances. In these cases, the predicates are interpreted as denoting the utterance of a meaningful expression. However, even this optional interpretation does not necessarily entail that the expression is produced with a communicative intention. Real speech act predicates — *say*, *reveal*, *claim*, and their likes — can only be used to describe reported declarative speech acts, where a speaker is described as attempting to make some interlocutors believe that a certain claim is true. This is the essence of the truth claim — an obligatory component of their meanings. This is not necessarily the case with manner-of-speaking predicates. Thus, they can occur with some directional adverbials, where they are interpreted as denoting the act of speaking NON-COMMUNICATIVELY in the presence of someone. This is impossible with the real communication predicates:

(15) George { howled something at me, but he wasn't saying anything to me. / *said something { to / at } me, but he didn't tell me anything. }

Most importantly, a negation of one manner-of-speaking predicate does NOT narrow down the set of possible truth claims made by the subject of the predicate. If (16) is true,

(16) John didn't { whisper / scream / howl } that Bill was an undercover agent.

we are not entitled to infer that there is a certain type of truth claim which John did not make. He may have *suggested*, or *claimed*, or *revealed* to us, or *admitted* that Bill was an undercover agent. All we know is that the truth claim was not made by means of a specific type of sound.

To the extent that a reported *mumbling*, *howling*, or *murmuring* event is interpreted communicatively, the communicative interpretation is always contextually-based. A sentence like *George mumbled that he was lonely and miserable* is perfectly compatible with a scenario where George was sitting by himself, mumbling in the dark. This is not the case with real communication predicates. To the extent that we want to interpret the sentence above as a report of a communicative act, we have to add elements to the context which go beyond the obligatory components of the meaning of *mumble*: an interlocutor; an intention on George's part to make the interlocutor believe that he was indeed lonely and miserable; and so on. The notion of truth claim is not semantically entailed by manner-of-speaking predicates, but can be pragmatically implied in certain usages of the predicates.

There are, however, at least four manner-of-speaking predicates — *yell*, *scream*, *whine*, and *whisper* — whose meanings seem to approximate that of real communication predicates. This seems to result from the fact that they are pragmatically associated with the communication of certain expressive meanings — shouting and yelling are associated with the expression of anger, whisper is associated with secretive communication, and whine is associated with complaining. This is why the following sentences are pragmatically wierd:

(17) ?? John { yelled that he was perfectly calm. / whined that he was the happiest guy in the world. / whispered that there was no need to keep it a secret. }

In line with the above semantic characterization, manner-of-speaking predicates disallow *that*-deletion:

(18) John { mumbled / murmured / howled / shrieked } { that Bill was an undercover agent. / *Bill was an undercover agent. }

For some speakers, however, *that*-deletion is acceptable with those predicates which approximate the meanings of real communication predicates (although these permissive speakers still prefer the version with the complementizer):

(19) John { screamed / yelled / whined / whispered } { that Bill was an undercover agent. / Bill was an undercover agent. }

3.2. *Predicates of instrument of communication*

Cable, *e-mail*, *fax*, *netmail*, *radio*, *relay*, *satellite*, *signal*, *telecast*, *telegraph*, *telex*, *wire*.

Predicates of instrument-of-communication denote the usage of different technical means for sending messages. Just like manner-of-speaking predicates, they can be used to refer to acts of communication by means of technical instruments, but they do NOT necessarily denote the communication of a truth claim. Here are a few examples where the message sent by the subject does not amount to a claim that some proposition is true:

(20) Mary { cabled Bill a few questions. / e-mailed John an empty message. / faxed George a picture of her old school. }

Just like manner-of-speaking predicates, however, predicates of instrument of communication can be pragmatically interpreted as reporting acts of making a truth claim — this time by means of technical instruments:

(21) Heather { cabled / e-mailed } Sara that the party would be tonight.

Like manner-of-speaking predicates, then, predicates of instrument of communication do not entail the notion of truth claim, but may be pragmatically associated with it.

Do predicates of instrument-of-communication allow *that*-deletion? Speakers' intuitions vary. Restrictive speakers judge the following sentences totally ungrammatical (the example is based on Levin [1993: 207]):

(22) *Heather { cabled / e-mailed } Sara the party would be tonight.

The more permissive speakers feel that the above sentences are "almost good," or "quite acceptable." However, they prefer the same sentences WITH the complementizer. In a sense, then, the difference between the permissive and restrictive judgments is a matter of degree, not a binary distinction. This is the same pattern we have seen when we looked at manner-of-speaking predicates.

3.3. *Emotive predicates*

Be amazed, be surprised, be annoyed, be amused, be appalled, be confused, be delighted, be disappointed, be dismayed, be embarrassed, be flabbergasted, be horrified, be pleased, be proud, be puzzled, be relieved, be shocked, be sorry, be worried.

Emotive predicates denote different emotional attitudes to different types of stimuli. They cannot be used to report eventualities of the truth claim type. Moreover, the stimuli do not necessarily have to be states-of-affairs in the world. They may be physical or abstract entities, or their properties:

(23) a. The children were amused by the clown.
b. George was appalled by John's behavior at the party.
c. Mary was delighted by the beauty of the mountain.

When an emotive predicate appears with a *that*-clause complement, however, it DOES carry an implication of the truth claim type — this time as a preparatory condition for the emotional attitude: if John was amazed that Bill gave up on the job, then John must have believed that he did it. If Mary was embarrassed that George behaved like a fool, then she must have believed that he did. In this sense, emotive predicates are similar to manner-of-speaking predicates — with an extra twist: they do not necessarily entail the notion of truth claim, but when they are used to report an emotional attitude towards a state-of-affairs, they imply that their subjects PRESUPPOSE the truth claim.

In line with this semantic characterization, most speakers find *that*-deletion totally unacceptable under most emotive predicates:

(24) George was { appalled / flabbergasted / annoyed } { that Bill behaved like a clown in that party. / *Bill behaved like a clown in that party. }

Very few speakers find these sentences marginally acceptable, but the testify that they definitely prefer the versions with the complementizer, which are perfectly natural.

There are, however, two small subsets of emotive predicates which, according to most speakers, allow *that*-deletion, or at least allow it in more freely (although in most cases, speakers still prefer the *that*-clause to the bare clause). I believe this is so for a good reason. The meanings of these predicates, as they are conventionally used, approximate those of predicates of belief, predicates of acquiring knowledge, or predicates of acknowledgment.

The first subset includes *be amazed* and *be surprised*, and possible some others. Consider the following examples:

(25) George was { amazed / surprised } that John behaved like a fool at that party.

If George was *surprised*, or *amazed* that John behaved like a fool at that party, we may infer that John's behavior somehow CONTRADICTED George's prior assumptions about John, his behavior, or his behavior at parties. However, if George was *appalled*, or *annoyed* that John behaved like a fool at that party, we cannot make the above inference. Consider the following exchanges:

(26) Do you think John will behave like a fool at the party?
  a. I'll be surprised if he does.
  b. I'll be appalled if he does.

In (26a), *be surprised* is used as a predicate of belief. (26a) constitutes a proper answer to the question: it means that the speaker has a strong reason to believe that John will not behave like a fool at the party. (26b), on the other hand, does not constitute an informative answer to the question: it does not imply that the speaker has any beliefs about the possibility of John's behaving like a fool. Unlike *be appalled*, predicates like *be surprised* and *be amazed* encode both the change-of-state of belief or knowledge of its subject, and the subject's own emotional attitude reaction to this change-of-state.

Consequently, for most speakers, *be surprised* and *be amazed* allow *that*-deletion:

(27) George was { amazed / surprised } John behaved like a fool at that party.

The second subset of predicates includes *be sorry*, *be proud*, and possibly some others. These predicates are conventionally used as predicates of acknowledgment:

(28) a. I'm sorry that you had to come here on such short notice.
  b. I'm proud that you managed to pull it off.

The speaker in (28) does not simply express his or her emotional attitude towards the state-of-affairs described by the embedded clause. The speaker actually declares his or her knowledge of this state-of-affairs, and lets the hearer understand that this state-of-affairs, and its emotional implications, are part of their common knowledge. For most speakers, *be sorry* and *be proud* allow *that*-deletion:

(29) a. I'm sorry you had to come here on such short notice.
b. I'm proud you managed to pull it off.

4. Type III predicates

4.1. *World-to-word predicates*

Provide, *arrange*, *scheme*, *plot*, *legislate*, *contrive*, *decide*.

All these predicates can take *that*-clauses (the examples in this subsection are from Bolinger [1972: 46–47]):

(30) a. They provided that a special agency would look after it.
b. We schemed (plotted) that they would be betrayed by one of their own men.
c. Congress has legislated that all cars must have smog devices.
d. I'll arrange (contrive) that they all get the same orders.

The distinction between these predicates and the ones we have looked at so far is very clear. The subjects in (30) are NOT reported to have believed or claimed that the propositions denoted by the embedded clauses are true. On the contrary: they are reported to have acted with the goal of MAKING them true. To the extent that their acts were successful, they caused the world to change in such a way that it fitted the proposition — the exact opposite direction of fit from that of communication, belief, and knowledge predicates. World-to-word predicates do not entail a truth claim, and cannot be used as such.

World-to-word predicates very clearly disallow *that*-deletion:

(31) a. *They provided a special agency would look after it.
b. *We schemed (plotted) they would be betrayed by one of their own men.
c. *Congress has legislated all cars must have smog devices.
d. *I'll arrange (contrive) they all get the same orders.

Bolinger brings another interesting example, which demonstrates the robustness of the syntax–semantics correlation: the predicate *decide* has

two meanings. First, it may be used as a propositional attitude predicate of the truth claim type, and mean something similar to *realize* or *understand*. When used with this meaning, with the word-to-world direction of fit, *decide* allows *that*-deletion:

(32) Why didn't you take the job? – I decided I was too old.

However, *decide* can also be used as a world-to-word predicate when the deciding agent has the power or authority to change reality simply by making the decision. When used with this meaning, the predicate does not allow *that*-deletion:

(33) Yesterday the Supreme Court decided { that wiretaps must stop. / *wiretaps must stop. }

4.2. *Predicates of physical manipulation of text*

Write down, *read aloud*, *record*, *publish*, *translate*, *jot down*.

Predicates of physical manipulation of text denote different physical manipulations of meaningful written expressions. *Write down* denotes the recording in writing of some meaningful text; *read aloud* denotes the oral presentation of some meaningful written text; *translate* denotes the transfer of some meaningful text from one language to another. In all these cases, the epistemic attitude of the subject — the person doing the writing down, the reading aloud, the translating — towards the truth or falsity of the propositional content of the text is completely irrelevant. The subject does not take responsibility, so to speak, for the claim made in the text. In this sense, these predicates are all very different from *write* and *read*, whose meanings do include the component of truth claim: if John wrote that George was engaged in some illegal activity, then he, John, can be held accountable for the claim that George was breaking the law. We know that one of the following should be true: Either the claim made in the written statement reflected John's belief that George was breaking the law, or John simply lied. Writing is just like *telling* and *claiming* — it is only done by other means. But if all we know is that the statement was recorded in writing by John, we cannot hold him accountable for its content. The same distinction can be made with respect to *read* and *read aloud*. Reading is just like *hearing*, or *figuring out*, or *realizing*: it is in its essence an activity of acquiring information — which is done by means of deciphering written signs. This is why it makes sense to warn naive readers "not to believe what they read in the papers." *Reading aloud*, on the other hand, is nothing more than the transmission

of the text from its written form to spoken words. If John read aloud (from the paper) that George was engaged in some illegal activity, he does not commit himself in any way to the claim that George was breaking the law.

The distributional facts correlate with the above semantic distinction: *Read* and *write* allow *that*-deletion. Predicates of physical manipulation of texts do not (all examples are from Bolinger [1972: 45]):

(34) John wrote { that it was raining where he was. / it was raining where he was. }

(35) He wrote down { that he was beginning to feel the symptoms. / *he was beginning to feel the symptoms. }

(36) He read { that it was the President who had made the move. / it was the President who had made the move. }

(37) He read aloud { that it was the President who had made the move. / *it was the President who had made the move. }

(38) Did you record { that he had been rewarded? / *he had been rewarded? }

(39) He jotted down { that the shipment was complete. / *the shipment was complete. }

(40) Look, you've translated { that Dumas didn't pay the debt. / *Dumas didn't pay the debt. }

4.3. *Predicates of issuing a guarantee*

Guarantee, warrant.

These predicates have double meanings. Under the first reading, their meaning is rather similar to that of *swear* or *assure* — they are truth claim predicates, which imply that their subject expresses a very strong commitment to the truth of the embedded proposition. When used with this meaning, *guarantee* and *warrant* allow *that*-deletion (examples from Bolinger [1972: 47]):

(41) a. I'll warrant he's the greatest fighter that ever lived. (He's the greatest fighter that ever lived, I'll warrant.)
b. I guarantee it's true. (It's true, I guarantee.)

In (42), however, the predicates are used with a very different meaning:

(42) a. Our company warrants that its product will outlast any other on the market.
b. Before I buy it they're going to have to guarantee that the motor will last for 150,000 miles.

In (42a), the company does NOT in any way claim that it's true that its product will outlast any other on the market. The car dealers in (42b) are NOT expected to promise or claim that the motor will last for 150,000 miles. The company and the car dealers in the above examples are not committed to the truth of the relevant propositions. The issue at hand is not whether the propositions turn out to be true in the real world, but what type of financial compensation the company and the car dealers offer to the customer in the possible scenario where the propositions turn out to be false.

The distinction between these two interpretations is easy to state in terms of possible worlds. The predicates in (41) involve the claim that the actual world is included in the set of possible worlds where the proposition expressed by the *that*-complement is true. The predicates in (42), however, do not involve any such claim: the real world is not claimed to belong to the set of possible worlds where the product outlasts all the others and the motor lasts for 150,000 miles. To be sure, the company and the car dealers are making a truth claim about the actual world, but the claim they are making is a very different one: they claim that the actual world belongs to the set of possible worlds where, if the product does not outlast all the others, or if the motor does not last for 150,000 miles, the customer is financially compensated. So, if a company guarantees that P, it actually promises that M, where M is the proposition "if not P, then Q."

In their second interpretation, then, *guarantee* and *warrant* do not encode a truth claim with respect to the embedded proposition. Consequently, they disallow *that*-deletion:

(43) a. *Our company warrants its product will outlast any other on the market.
b. *Before I buy it they're going to have to guarantee the motor will last for 150,000 miles.

4.4. *Predicates of occurrence*

Happen, *occur*, *come about*, *develop*.

Consider the following contrast (based on Bolinger's examples [1972: 19]):

(44) a. It occurred that the captain had lost his orders.
b. *It occurred the captain had lost his orders.
c. It occurred to John the captain had lost his orders.

At first sight, the facts in (44) seem rather mysterious. A closer look, however, reveals that they constitute one of the most remarkable pieces of evidence for my theory. (44a) refers directly and objectively to an event in the real world, namely, the event in which the captain had lost his orders. Although the sentence implies that the event actually happened, it does not imply that a cognitive agent has made a truth claim about the proposition describing this event. In other words, (44a) does not base the fact of the described event on a truth claim made by an agent. (44c), on the other hand, reports an event in which John realized that the proposition denoted by the embedded sentence was true. (44c) reports an event of the truth claim type, whereas (44a) does not. Consequently, the complementizer can be deleted in (44c), but not in (44b).

Additional examples (based, again, on Bolinger) reveal the same pattern. Predicates which refer directly to eventualities in the world — and not to the claim, made by a cognitive agent, that the propositions describing these eventualities are true — do not allow *that*-deletion:

(45) It { came about / happened / developed } { that the captain had lost his orders. / *the captain had lost his orders. }

4.5. *Complex expressions*

Consider the following contrasts (from Bolinger 1972: 20):

(46) a. We got the message they were coming.
b. *I lost the message they were coming.

(47) a. They came to the conclusion they had to act.
b. *They stressed the conclusion they had to act.

(48) a. Where did you get the idea I didn't want to?
b. *He was referring to the idea I wouldn't help him.

The first sentences of the above pairs imply a truth claim, but the second sentences do not: (46a) is equivalent to a sentence like *we heard they were coming*, but (46b) reports an event in which the subject lost a physical object (a piece of paper, a fax, a message on an answering machine) which recorded the message that they were coming. (47a) is equivalent to a sentence like *they decided they had to act*, but (47b) does not imply that the

agents who stressed the conclusion were necessarily the ones who arrived at it. The subjects of the sentence are not necessarily responsible for the truth claim made in the conclusion. (48a) is equivalent to a sentence like *what made you think I didn't want to?*, but in (48b), the subject is referring to an idea which was not necessarily his own. Consequently, the first sentences allow *that*-deletion, but the second ones do not.

In summary, we have looked at a wide array of sentence-embedding predicates and expressions, and have examined both their meanings and their *that*-deletion behavior. We discovered a clear pattern of correlation: virtually all truth claim predicates, which semantically entail that a cognitive agent (most often their subject) has made a claim regarding the truth of the embedded proposition, freely allow *that*-deletion. Non-truth claim predicates — which do not entail that such a claim has been made — do not allow *that*-deletion. Predicates which can be pragmatically interpreted as implying a truth claim, although their semantics does not necessarily entail it, have a more variable behavioral pattern: Some speakers judge *that*-deletion under these predicates totally unacceptable, and some speakers allow *that*-deletion under subclasses of these predicates, which seem to conventionally approximate the meanings of real truth claim predicates. Moreover, when *that*-deletion is possible, it is always optional. I take this to be a very significant fact: I have not been able to find a single predicate which can taken the bare clause but rejects the *that*-clause.

5. Corpus results

The correlations presented in the above sections were strongly supported by the results of a large-scale search of the British National Corpus. The distributional patterns of 32 verbs were checked. A thousand random sentences per each verb were downloaded, and the appearances of *that*-clause and bare clause complements were counted. As is often the case with corpus-based analyses, some of the verbs yielded zero results. That was the case with three verbs of manipulation of text and three verbs of instruments of communication. These verbs were not found with sentential complements of either form. The remaining 26 verbs, however, yielded results which correlated with those based on the speakers' judgments. They are presented in Table 1.

The most interesting results have to do with the ambiguous verbs, *guarantee*, *warrant*, and *decide*, which, as we have seen before, can be used as either TYPE I or TYPE III verbs. Here the results were strikingly clear, as can be seen in the lower part of Table 1. The TYPE I usages of these

Table 1. *The distribution of* that-*clauses and bare clauses in sets of a 1000 token sentences per each verb, randomly downloaded from the National British Corpus*

Type	Verb	*That*-clauses	Bare clauses	TYPE	Verb	*That*-clauses	Bare clauses
TYPE I	think	126	347	TYPE II	shout	8	2
	suggest	161	80		scream	5	0
	affirm	58	2		mumble	2	0
	hypothesize	35	0		be sorry	4	22
	tell	106	86		be surprised	21	4
	argue	512	40		be shocked	19	0
	occur to x	9	2		be relieved	29	8
	realize	364	172		be amazed	9	1
	find out	147	84	TYPE III	happen	11	0
	guarantee[1]	27	13		guarantee[2]	37	0
	warrant[1]	4	9		warrant[2]	9	0
	decide[1]	77	38		decide[2]	18	0
TYPE II	whisper	16	0		provide	17	0

verbs appear with the bare clauses in a considerable number of cases, whereas the TYPE III usages only appear with *that*-clauses.[9]

6. Explaining the correlation

Capturing the correlation between the lexical semantics of the embedding predicates and their *that*-deletion behavioral patterns constitutes the first step towards an explanation of the phenomenon of *that*-deletion. The explanation, however, cannot be based exclusively on the semantics of the predicates. To understand the correlation, we have to understand the role played by the meanings of the sentential complements — the bare clause and the *that*-clause. A full explanation should ideally explicate the terms of compatibility between the lexical semantics of the different predicates and the meanings of the complements they are capable of embedding. Translating our empirical findings into the language of compatibility, a full explanation of the facts should capture the following set of generalizations:

(i) The meanings of Type I predicates are fully compatible with the meanings of both types of complements, the bare clause and the *that*-clause;
(ii) The meanings of Type III predicates are fully compatible with the meaning of the *that*-clause, but are incompatible with the meaning of the bare clause;

(iii) The meanings of Type II predicates are fully compatible with the meaning of the *that*-clause; for some speakers, the meanings of some Type II predicates are partially, or almost fully compatible with the meaning of the bare clause — where the level of compatibility seems to be determined on pragmatic grounds;
(iv) There are NO attested predicates whose meanings are compatible with the meaning of the bare clause, but not with the meaning of the *that*-clause.

What, then, is the meaning difference between *that*-clauses and bare clauses? To answer this question, let us turn our attention to a very fundamental syntactic fact: matrix clauses are ALWAYS bare clauses, never *that*-clauses:[10]

(49) a. John tried to steal money from the bank.
b. *That John tried to steal money from the bank.

What does this fact mean? Searle and Vanderveken (1985) use it to make an important distinction between two components of speech acts — the illocutionary act and the propositional act:

> In the performance of an illocutionary act the speaker performs the subsidiary act of expressing the propositional content and this act we will call the propositional act. A propositional act is an abstraction from the total illocutionary act in the sense that the speaker cannot simply express a proposition and do nothing more. The performance of the propositional act always occurs as part of the performance of the total illocutionary act. Syntactically this fact is reflected in natural languages by the fact that 'that' clauses, the characteristic form of isolating the propositional content, cannot stand alone; they cannot make complete sentences. One can say 'I promise that I will leave the room,' but one cannot say simply 'That I will leave the room.' (Searle and Vanderveken 1985: 9).

Searle and Vanderveken (1985: 9), then, characterize the *that*-clause as the object of the propositional act, the "characteristic form of isolating the propositional content". By implication, we may characterize the bare clause as the object of the assertive speech act — the propositional object which is expressed (as the object of the propositional act) and claimed to be true (as the object of the illocutionary act). This explains the fact that the bare clause, as opposed to the *that*-clause, can "stand alone" in matrix position. Let me, then, suggest the following distinction: *that*-clauses denote propositions (or, maybe, more accurately, propositional contents), and bare clauses denote propositions qua objects of assertive acts. Let us call them "asserted propositions."[11]

Let us turn back to embedded contexts. Consider a regular speech act predicate, like *tell* or *claim*:

(50) John { told us / claimed } { that Bill tried to steal the money from the bank. / Bill tried to steal the money from the bank. }

The act of *telling* or *claiming* consists of exactly the two components of speech acts we have just looked at. (49) entails that John expressed the proposition that Bill tried to steal the money from the bank (the propositional act), and that he claimed that it was true (this is what we called the truth claim). Thus, in line with the distinction between propositions and asserted propositions, we may conclude that the meanings of predicates like *tell* and *claim* are semantically compatible with the meanings of both types of complements: the object of *telling* or *claiming* is necessarily a proposition, and this proposition is necessarily an asserted one. Consequently, *tell* and *claim* can take both types of complements.

The same line of argumentation holds for belief and knowledge predicates. These predicates do not denote speech acts, but they do denote eventualities which consist of propositional acts (in the mental domain), and truth claims (again, in the mental domain). Consider the examples in (51):

(51) John { believed / realized } { that Bill tried to steal the money from the bank. / Bill tried to steal the money from the bank. }

(51) entails that John had (or came to have) a mental representation of the proposition that Bill tried to steal the money from the bank, and that he made the mental claim that this proposition was true. Consequently, predicates like *believe* and *realize* are semantically compatible with the meanings of the two types of complements, and are capable of embedding both.

All Type I predicates share the two meaning components — the propositional act and the assertive act — with *tell*, *claim*, *believe*, and *realize*. This is why they are compatible with both types of complements. Note that this is no coincidence: all predicates which denote truth claims necessarily denote the subsidiary propositional act. One CANNOT *think*, *understand*, *guess*, *insist*, *reveal*, or *divulge* that P if one does not represent P, either mentally or publicly. In other words, every predicate which is compatible with the meaning of the bare clause is NECESSARILY compatible with the meaning of the *that*-clause. Thus, every predicate which is capable of embedding bare clauses must necessarily be able to embed *that*-clauses. This is why *that*-deletion is always optional, and there are no predicates which obligatorily delete the complementizer.

The opposite relation, however, does not hold: Not every predicate which is compatible with the meaning of the *that*-clause is compatible with the meaning of the bare clause. This is so for an obvious reason: propositions — descriptions of states-of-affairs which may or may not be true — may participate in many types of eventualities which do not involve truth claims. Consider, for example, Bolinger's examples with *warrant* and *guarantee*:

(52) a. Our company warrants that its product will outlast any other on the market.
b. Before I buy it they're going to have to guarantee that the motor will last for 150,000 miles.

In these examples, the warranted and guaranteed propositions are definitely the objects of a propositional act. The company, or the car dealer, describe a state-of-affairs where the product will outlast any other on the market, or the motor will last for 150,000 miles. As we have already seen, they do not make the claim that these propositions are true. Thus, the two predicates are semantically compatible with the meaning of the *that*-clause, but not with the meaning of the bare clause. Consequently, they can only embed *that*-clauses, but not bare clauses.

The different Type III predicates we have looked at denote different types of eventualities, in which propositions play different roles. With world-to-word predicates, for example, propositions describe states-of-affairs which result from the activities denoted by the predicates. With predicates of physical manipulation of text, propositions describe states-of-affairs represented by the physical products of writing down, reading aloud, and so on. With predicates of occurrence, propositions describe factual states-of-affairs in the world. As we have seen, none of these cases involve truth claims: the different predicates do not denote eventualities in which a proposition is asserted by a cognitive agent to be true. Consequently, Type III predicates are semantically compatible with *that*-clauses, but not with bare clauses, and can only embed the first type of complement, but not the second.

The more complex behavioral patterns of Type II predicates follow naturally from the explanatory framework suggested above. As we have seen, Type II predicates do not entail truth claims, but can sometimes be pragmatically interpreted as truth claim predicates. Moreover, some Type II predicates are conventionally associated with truth claim usages. This semantic–pragmatic pattern suggests the following predictions: (i) speakers' judgment should vary to a higher degree than with Type I and Type III predicates; (ii) individual speakers should find it harder to make grammaticality judgments concerning *that*-deletion with Type II

predicates than with Type I and Type III predicates; and (iii) speakers should find *that*-deletion more acceptable with those Type II predicates which are more conventionally associated with truth claim usages. As we have already seen, these predications are borne out by the facts.

How can speakers' variation be explained in terms of compatibility? I suggest that the difference between permissive and restrictive speakers should be formulated as follows. Restrictive speakers employ a strict notion of SEMANTIC COMPATIBILITY: Type II predicates do not entail truth claims, and are thus incompatible with asserted propositions. Permissive speakers employ a looser notion of SEMANTIC–PRAGMATIC COMPATIBILITY: Type II predicates are sometimes pragmatically interpreted as truth claim predicates, and when they are, they are compatible with asserted propositions. Thus, for restrictive speakers, the complementizer can only be deleted when the predicate semantically entails a truth claim. For permissive speakers, on the other hand, the complementizer can be deleted when the predicate can be interpreted, semantically or pragmatically, as a truth claim predicate. Obviously, more speakers will find *that*-deletion possible with a predicate which is more closely associated with truth claim usages, and in this sense, Type II predicates form a continuum of acceptability gradations, all the way between Type I and Type III predicates. The fact that clear, nonvariable judgments fall in the domain of semantics (Type I and Type III predicates), whereas the continuum of unclear, variable judgments is found in the domain of pragmatics (Type II predicates) seems to me to be a very welcome result of the analysis.

6. Additional explanatory factors

6.1. *The role of register*

As we have seen in the beginning of Section 2, there are a few speech act predicates which do not naturally allow *that*-deletion, although their semantics seems to license it. The predicates are *hypothesize*, *assert*, *affirm*, and probably some others. These predicates have at least one property in common: they are very formal, technical predicates, and they are rarely used in informal conversation. This fact rings a familiar bell. Bolinger suggests that frequency of usage plays a role in determining the distribution of *that*-deletion: the more frequently a predicate is used in regular conversation, the more likely it is to allow *that*-deletion. As a sweeping generalization, I do not see that Bolinger's hypothesis captures the facts. As we have seen, for example, Type III predicates, some of which are used very frequently and regularly in informal conversation, nevertheless

disallow *that*-deletion all across the board. However, as a recent study (Biber 1999) convincingly shows, *that*-deletion as a general phenomenon is dramatically associated with informal register. According to Biber's corpus-based analysis, the complementizer is deleted in about 85% of all the occurrences of complement clauses in conversation register; in about 65% of all occurrences in fiction register; in about 30% in news register; and in about 5% (!) of all occurrences of complement clauses in academic prose register (cf. Biber 1999: 145). We are thus entitled to reformulate Bolinger's original suggestion in following manner: *that*-deletion, as a grammatical phenomenon, is register-bound. It is associated with informal conversation. Consequently, predicates allow *that*-deletion (i) only if they are (or can be pragmatically interpreted as) truth claim predicates, AND, (ii) to the extent that they are associated with conversation register. Thus, register contributes an additional constraint on the distribution of *that*-deletion, beyond the constitutive level of semantic determination which lies at the foundation of the phenomenon: informal Type III predicates do not allow *that*-deletion, but some Type I predicates, which are very clearly associated with higher, formal register, disallow *that*-deletion in spite of their semantics.

6.2. *Complementhood, adjacency, and complexity*

As indicated in the introduction to this article, the phenomenon of *that*-deletion seems to be severely restricted on the structural level: *that*-deletion is impossible in subject-position, under derived nominals, and in topicalized clauses. These facts have inspired a series of ECP-related analyses (Stowell 1981; Kayne 1981; and Snyder and Rothstein 1992), and at least two optimality-based analyses (Pesetsky 1998; Grimshaw 1997). The general spirit of these analyses is most clearly captured by Pesetsky (1998: 357): "The head ... of a CP may be deleted only if that CP is a complement."

How should this descriptive principle be theoretically interpreted in relation to the semantic findings discussed in the present article? Two possibilities come to mind. First, it may be argued that the entire set of *that*-deletion facts, including those which are lexically-bound, are syntactically determined — *that*-deletion is only possible in complement position — and that the semantic factors described in this article actually help determine complementhood. In other words, the sentential objects of such predicates as *think* and *claim* are genuine complements, whereas the sentential objects of such predicates as *murmur*, *write down*, and *legislate* are

actually adjuncts. (This seems to be Stowell's position with respect to manner-of-speaking predicates, for example.) This view is problematic, for at least three complementary reasons. First, I do not see any good reason to assume that the sentential objects of Type III predicates are adjuncts, rather than complements. Second, as we have seen, Type II predicates (including manner-of-speaking predicates) seem to be contextually sensitive: when pragmatically used as communication predicates, they tend to allow *that*-deletion — at least for some speakers. If *that*-deletion is semantically-determined, the pragmatic influence seems natural; if, however, *that*-deletion is determined by complementhood, we are forced to assume that complementhood is somehow pragmatically sensitive. That this should be the case is at least mildly surprising. Third, the structural explanation strips the distinction between bare clauses and *that*-clauses from its semantic significance: the compatibility-type contention that asserted propositions are only compatible with truth claim predicates has an explanatory value which its structural counterpart seems to lose. I do not see any good reason to assume, a priori, that asserted propositions should be structurally restricted to complement position. The second approach might be that of a theoretical division of labor: (i) *that*-deletion is only possible in complement position; and (ii) in complement position, the distribution of *that*-deletion is determined by semantic compatibility. There is, however, a certain level of redundancy in this approach. After all, both sides of the equation — the syntactic restriction and the semantic considerations — have to do with the relations (or lack of relations) between predicates and their (syntactic or semantic) complements. It seems reasonable to look for a theoretical way to reduce this apparent level of redundancy — and capture the structural restriction without the reduplication of the semantics in the syntactic explanation.

A possible solution may be based on the fact that syntactic complementhood is actually a combination of two separate properties. On the one hand, it is a hierarchical relation between a predicate and an object. On the other hand, it is a fact about syntactic positioning — essentially a fact about an adjacency relation between the predicate and the object. As a hierarchical relation, however, syntactic complementhood is quite obviously intimately related to semantic complementhood. It may prove useful, then, to replace the syntactic complementhood restriction with an adjacency constraint, and let the semantic theory developed in this article do the rest of the work — along the following lines: *that*-deletion is possible only when (i) the bare clause is (semantically or pragmatically) compatible with the embedding predicate; and (ii) the bare clause is positioned to the right of, and adjacent to the embedding predicate. Formulated in this way, the structural constraint in (ii) seems to fall within the domain of

pronunciation principles, in the general spirit of Pesetsky (1998): it is a constraint on the pronunciation of the bare clause.

The above constraint immediately rules out *that*-deletion on the left side of the predicate, that is, in subject position and in topicalized clauses. Crucially, however, it allows for a certain amount of variability with respect to clause shifting on the right side of the predicate — because adjacency is a relative notion. As is well known, intuitions concerning *that*-deletion in shifted clauses are much more flexible and variable. It seems, for instance, that the following examples are unacceptable for some speakers, but quite grammatical for others:

(53) a. The president said in his press conference yesterday he was not planning to run for re-election.
b. Mary swore on a stack of bibles she would not eat the cookies.

Note that this variability in judgments cannot be explained by the syntactic complementhood restriction: it does not seem reasonable that speakers vary with respect to the complementhood status of the sentential objects in the above examples. Some kind of adjacency constraint, so it seems, is thus necessary even within the syntactic approach. As it seems to me, the adjacency constraint, coupled with the semantic analysis developed in this article, renders the complementhood analysis redundant.

Having isolated the adjacency constraint as the relevant causal parameter, we may attempt to account for it in deeper, explanatory terms, and here the COMPLEXITY framework of Rohdenburg (1996) and Hawkins (2001) seems especially promising. Dealing explicitly with *that*-deletion, Rohdenburg (1996: 161) shows that "any elements capable of delaying the processing of the object clause and thus of the overall sentence structure favor the use of an explicit signal of subordination." Moreover, "the incidence of *that* generally correlates with the overall complexity of both the superordinate and subordinate clauses." In more general terms, Hawkins (2001: 2) maintains that "the preferences for adjacency appear to reflect the number of syntactic and semantic relations that hold between two categories, and the speed and efficiency with which each such relation can be processed in an adjacent order versus a non-adjacent one." Assuming that the appearance of a bare clause in subject position would unnecessarily delay (and complicate) its processing as a complement, we may indeed expect to only find *that*-clauses in subject positions. Assuming that clause shifting may delay and complicate the processing of the complement to a certain degree, we may indeed expect to find more *that*-clauses in circumstances where more material intervenes between the verb and its complement.

The interesting question, however, is whether we interpret the complexity claim as a claim about grammaticalization, or as a claim about online processing in real time. Hawkins (2001: 2) seems to think in terms of grammaticalization: "grammars have conventionalized the preferences of performance." In this sense, then, Hawkins' framework actually complements Pesetsky's principle, rather than rendering it redundant. Rohdenburg, on the other hand, seems to think in terms of online processing, and here it seems to me that the robust fact that bare clauses never appear in subject position is not accounted for. We may imagine situations in which the semantic status of the clause in subject position is quite clearly indicated by the context (by previous utterances, for example), but this does not seem to make a difference in terms of the overall prohibition of bare clauses in subject position. The same problem arises with respect to the suggestion that structural constraints of this type may be related to such pragmatic factors as focus and topic. My understanding of this suggestion is that, at least as a first approximation, those contexts where *that*-deletion is totally impossible are usually associated with topichood (subjects and topicalized clauses), whereas those contexts where it is possible are prototypically associated with focus. I am generally sympathetic to this suggestion, as I am to many similar attempts to develop an explanatory theory of grammar on a pragmatic basis (e.g. Erteschik-Shir 1998). The recurrent problem with this approach, however, is that the explanatory tools are not as robust as the empirical facts in need of explanation. Thus, for example, although subjects are prototypically associated with topics, they are not always topics — sometimes they are focused. Crucially, however, *that*-deletion in focused sentential subjects is not easier than it is in topic-subjects.

6.3. *A note on the optional status of* that-*deletion*

The optional status of *that*-deletion raises a very interesting performance question: provided that a certain predicate allows *that*-deletion, under which conditions will the complementizer actually be deleted? When will speakers choose to say, for example, *John knew that Bill planned the whole thing*, and when will they choose to say *John knew Bill planned the whole thing*? Thompson and Mulac (1991) deal directly with this question, and they conclude that certain grammatical parameters — first and second person subjects, the predicates *think* and *guess* as main predicates, the existence of auxiliaries and indirect objects in the main clause — are, probabilistically speaking, significant predictors of *that*-deletion. They explain this probabilistic pattern by claiming that some combinations of

main clause subjects and predicates in English are being reanalyzed as "unitary epistemic phrases." Thompson and Mulac claim that in expressions like *I think P*, "the distinction between 'main' and 'complement' clauses is being eroded" (Thompson and Mulac 1991: 237). Thus, in sentences of this type, where the subject of the main clause refers back to the speaker, and sometimes to the listener, the "bleached-out main predicates and their subjects behave very much like single epistemic morphemes in other languages" (Thompson and Mulac 1991: 239). This, in turn, explains the fact, discussed at length by Hooper (1975), that these "bleached-out" combinations can be moved to positions where they do not seem to function as complement embedders, as in (54b):

(54) a. I believe it's going to rain.
b. It's going to rain, I believe.

Underhill (1988), quoted by Thompson and Mulac, makes a similar claim with respect to *that*-deletion in journalistic English: probabilistically speaking, *that* is deleted when the subject of the lower sentence is the topic of the utterance, and when the writer makes or endorses the assertion of the lower sentence.

Two observations need to be made with respect to these discourse-based analyses. First, the identification of the predicates which allow *that*-deletion with the predicates which allow complement preposing is misleading. Thus, for example, manner-of-speaking predicates seem to allow complement preposing, but do not easily allow *that*-deletion:

(55) a. *I shrieked we were getting a new chairman.
b. We were getting a new chairman, I shrieked.

Second, and more importantly, even to the extent that speakers most frequently take advantage of the option of *that*-deletion when they use the predicates *guess* and *think*, when they talk about their own beliefs or thoughts, or when they endorse the embedded proposition, this should not obfuscate the fact that speakers accept *that*-deletion as perfectly grammatical with a much larger set of predicates, with second and third person subjects, and with embedded propositions which they take to be completely false. Moreover, as we have seen, speakers judge *that*-deletion as perfectly ungrammatical with a large set of other predicates, even when these predicates are used with a first person subject, and when they take the embedded proposition to be completely true. Thus, the relevant probabilistic tendencies cannot be taken to constitute an explanation of the facts of *that*-deletion. Rather than constitute the explanation, these tendencies should be explained within a more general theory of the

phenomenon: we should find a way to explain both the fact that speakers make certain judgments about their grammatical options, and the fact that they tend to use some of their grammatical options more frequently than others. Let me try to sketch such an explanatory framework, based on the semantic–pragmatic theory presented above.

First, we have already seen that the option of *that*-deletion — properly thought of as part of speakers' competence — is determined as a function of the compatibility between the lexical semantics of the embedding predicates and the meanings of the two complement types, the bare clause and the *that*-clause. Thus, the performance question of the real-time choice between the two complements is only relevant for the subset of predicates which allows *that*-deletion in the first place.

Now, to the extent that a certain predicate does allow *that*-deletion, the real-time choice between the two complement types does NOT make a difference in terms of truth conditions. Thus, the sentences *John knew that Bill planned the whole thing* and *John knew Bill planned the whole thing* have the same semantic interpretation. This is so, because the object of predicates like *know* is necessarily a proposition, and this proposition is necessarily an asserted one. This specification is part of the lexical semantics of the predicate itself.

There is, however, a non-truth-conditional difference between the options. In the case of the bare clause, the proposition is structurally marked as an asserted proposition. When the embedded object is the *that*-clause, on the other hand, the proposition is not structurally marked as an asserted proposition. It is interpreted as such in virtue of the fact that it is embedded under a truth claim predicate. This structural difference may, in principle, be used by speakers to make more fine-grained pragmatic distinctions of different types. Here are two possible examples:

(i) The structural distinction should allow speakers to distinguish between two types of contexts: (a) contexts in which they focus on the proposition as the object of the propositional act, that is, the idea that some proposition was acquired, or is being stored, by a cognitive agent, and (b) contexts in which they focus on the proposition as the object of the assertive act, that is, the idea that some proposition is taken by the cognitive agent to be true. In the first case, speakers may use the *that*-clause, as opposed to the bare clause, to mark their choice of focus. In the second case, speakers may use the bare clause. This pattern of choice seems to underlie Bolinger's observations regarding the usage of the predicate *learn*. As we have already seen, *learn* allows *that*-deletion. Notice, however, that there are contexts which favor the deletion, whereas in other contexts the complementizer seems almost obligatory (examples adapted from Bolinger [1972: 44]):

(56) What's new? – I just learned { that Mr. Smith got nominated as the new C.E.O. / Mr. Smith got nominated as the new C.E.O. }

(57) What did you learn in school today? –
I learned { that Johnson was president 63′–68. / *Johnson was president 63′–68. }

In (56), the speaker very clearly focuses on the epistemic status of the new piece of information, that is, the fact that a certain proposition — *Mr. Smith got nominated as the new C.E.O.* — turned out to be true. In this context, the bare clause is perfectly natural. In (57), on the other hand, the speaker focuses on the very acquisition of the new piece of information, that is, the fact that a certain proposition — *Johnson was president 1963–68* — was the object of teaching in school. In this context, the complementizer seems almost obligatory. Note, moreover, that although (57) carries the pragmatic implicature that the proposition *Johnson was president 1963–68* is true — this implicature is very easily canceled:

(58) I learned that Johnson was president 1963–68, but later on it turned out this wasn't true.

In (56), on the other hand, canceling the implicature is much more difficult:

(59) ??I learned Mr. Smith got nominated as the new C.E.O., but later on it turned out this wasn't true.

(ii) The structural distinction may also be used by speakers to mark the proposition as either new information, that is, a proposition which has not yet been "asserted" in the discourse, or as part of the common ground, that is, a proposition which has already been "asserted" in the discourse. To focus on the status of the proposition as a part of the common ground, speakers may choose to use the bare clause. To focus on the status of the proposition as new information, speakers may choose to use the bare clause. Consider the following scenario (partially based on examples from Bolinger [1972: 60]): A police officer has just been told that a robbery has been committed at the local gas station. When she gets there, she sees broken glass all over the place. The attendant is ostensibly shaken. The fact that the place has just been robbed is by now very obviously part of the common ground. The police officer is likely to say something like:

(60) They told us you've been robbed.

Imagine, by contrast, a scenario where the police officer arrives at the gas station, and finds no signs of robbery. The attendant seems relaxed and happy. The police officer approaches the attendant and tells him about the robbery report. In this context, the police officer is more likely to say something like:

(61) They told us that you've been robbed.

Note, again, that the choice between the two options does not make a difference in terms of truth conditions. Moreover, the pragmatically-based choice is a matter of tendency, not of obligatory selection: Speakers may use *that*-clauses to refer to common-ground propositions, or bare clauses for new information. Probabilistically speaking, however, we may hypothesize that speakers will tend to delete the complementizer more often when focusing on the propositional object as an asserted proposition. The prototypical cases where this is the case are exactly the ones discussed by Thompson and Mulac: in sentences with a first person subject, for example, the speaker reports his or her own belief or claim. In this case, where the cognitive agent licensing the proposition is the speaker, and the deletion of the complementizer focuses on the proposition as a truth-bearing unit, the utterance of a sentence like 'I think P' virtually equals an utterance of 'P' by the same speaker, amended by the restriction that the speaker's epistemic commitment to P is slightly weaker than absolute knowledge. In such a case, we are indeed witnessing an effect of semantic bleaching. This effect, rather than explain the phenomenon, actually results from the interaction between the semantics of the embedding predicates, the semantics of the embedded complements, and the pragmatics of their usage.

7. Conclusion

The first and primary goal of this article was strictly empirical: to establish the hypothesis that the distribution of *that*-deletion across sentence-embedding predicates (in complement position) is non-arbitrarily determined by the lexical semantics of the embedding predicates. We have looked at a large number of predicates, and have seen that the distribution is indeed correlated with the semantic notion of truth claim: virtually all Type I predicates, which entail that their subject has made an epistemic claim concerning the truth of proposition denoted by the embedded clause, allow *that*-deletion. (Some Type II predicates, which are associated with formal register, do not allow *that*-deletion.) Type III predicates, which do not entail truth claim and cannot be pragmatically used as truth

claim predicates, do not allow *that*-deletion. Type II predicates, which do not entail the notion of truth claim but can sometimes be pragmatically used to imply a truth claim, display a more complex behavioral pattern: some speakers judge *that*-deletion under these predicates to be totally ungrammatical. For other speakers, *that*-deletion is quite acceptable under a subset of Type II predicates — which seem to be frequently used with extended meanings, approximating those of Type I predicates.

As a first step towards the explanation of this correlation, I suggested that a distinction should be made between the meanings of the bare clause and the *that*-clause in English: whereas *that*-clauses denote regular propositions, bare clauses denote asserted propositions — propositions which have a putative truth-value, assigned to them by a cognitive agent. Following Searle and Vanderveken (1985), I suggested that this semantic distinction explains the fact that bare clauses, but not *that*-clauses, can stand alone in matrix position. The semantic distinction allows us to explain the empirical results described in this article in terms of compatibility of meaning. To achieve this theoretical goal, I suggested that speakers use two different notions of compatibility. Some speakers employ a strict notion of semantic compatibility: According to these speakers, a predicate will be able to embed a certain type of complement only if the meaning of the predicate is semantically compatible with the meaning of the complement. These restrictive speakers allow *that*-deletion under Type I predicates, and disallow it under Type II and Type III predicates. Other speakers employ a more permissive notion of compatibility, according to which a predicate can embed a complement if the meaning of the predicate is semantically or pragmatically compatible with the meaning of the complement. Just like restrictive speakers, permissive speakers allow *that*-deletion under Type I predicates, and disallow it under Type III predicates. They differ from restrictive speakers, however, in their acceptance of *that*-deletion under some Type II predicates, which are pragmatically compatible with the meaning of the bare clause.

Beyond explaining the distributional facts, this analysis makes two important predictions, which are borne out by the facts. First, it predicts that *that*-deletion is always optional, that is, that there is no English predicate which takes the bare clause but rejects the *that*-clause. This is so because every predicate whose meaning is compatible with a propositional object of the asserted type must, by definition, be compatible with a propositional object as such. Remember that the opposite relation does not hold: as we have seen, there are many predicates whose meanings are compatible with a propositional object, but not with an asserted proposition. Second, the analysis predicts that speakers will display a clear, non-variable pattern of syntactic judgments in the domain of semantics, and a

continuum of unclear, variable judgments in the domain where pragmatic compatibility is applicable. This prediction, which is indeed borne out by my findings, strengthens the hypothesis that the syntactic patterns of *that*-deletion are meaning-based, exactly because it correlates degrees of syntactic vagueness with degrees of semantic vagueness: The claim that a certain syntactic domain is determined by meaning does not imply that speakers will never make quasi-arbitrary syntactic decisions in that domain. What it does imply is that cases where speakers make such syntactic decisions will be those where the semantic parameters do not allow for a clear judgment in terms of semantic compatibility.

To account for the well-known fact that *that*-deletion is severely restricted in noncomplement position, I suggested that the above semantic analysis should be supplemented with an adjacency-type constraint on the pronunciation of the bare clause. This constraint, in turn, may have emerged on the basis of considerations of processing complexity.

One final note of a more general nature: It is sometimes assumed (cf. Farkas 1990: 72), that universal and stable syntactic phenomena are more likely to be meaning-based than typologically rare and unstable ones. According to Farkas, for example, cases of arbitrariness do not falsify the hypothesis of the non-autonomy of syntactic structures, as long as "the cases which are claimed to be arbitrary do indeed bear the marks of arbitrariness, namely instability of various types: historical, cross-linguistic, dialectal, idiolectal" (Farkas 1990: 72). Although I agree with Farkas' general argument, I believe that the results presented in this article show that cross-linguistic and dialectal instability, and interpersonal variability, do not necessarily point at syntactic arbitrariness. Typologically speaking, the phenomenon of *that*-deletion is very rare indeed, but, as we have seen, at least in the English case, it is very clearly meaning-based.

Received 7 November 2001
Revised version received
5 June 2003

Tel Aviv University

Notes

* Correspondence address: Dept. of Communications, Tel Aviv University, Ramat Aviv 69978, Israel. E-mail: danield@post.tau.ac.il.

1. As *that*-deletion is also licensed by adjectival predicates (e.g. *I consider it obvious he will win*), not just by verbal ones, the notion of predicate, rather than the notion of verb, will be used throughout the article.

2. *That*-deletion phenomena are sometimes related in the literature to other sets of facts, such as those collectively known as the "*that*-trace effect." Moreover, there seem to be parallels between the syntactic environments which allow *that*-deletion and those where raising is possible with infinitival clause complements. I will not discuss these issues in this article.
3. In the case of *know*, of course, the speaker also presupposes that the proposition is true. In other words, *know*, as opposed to *think*, is also a FACTIVE verb (Kiparsky and Kiparsky 1979). The notion of factivity will not be put to use in the present analysis.
4. Note that the notion of TRUTH CLAIM is a property of the verb, not of the entire proposition. Thus, a sentence such as *John never told me Bill was going to do this* is perfectly fine, although it indicates that John did not in actual fact make the claim that Bill was going to do this. This sensitivity to the verb itself is one of the major insights emerging out of the literature on lexical semantics and its syntactic manifestations.

   Note, moreover, that in the great majority of cases, the truth claim is about the real world, but this is not necessarily the case in all verbs. The verb *imagine*, for example, entails that its subject is committed to the truth of the embedded proposition in a possible world which the speaker decides is different from the real one.
5. The correlation between the lexical semantics of Type I predicates and their *that*-deletion patterns is not a perfect one: there seems to be a small set of speech act predicates whose semantics suggests that they should allow *that*-deletion, but they do not. I have not found, however, Type III predicates which freely and generally allow it despite of their semantics. Thus, although the semantic reality discussed in this article is the fundamental determinant of the distributional facts, there is probably an additional factor involved, which further constrains the domain of *that*-deletion. This factor may be REGISTER, and its contribution is discussed in Section 6.1.
6. I thank an anonymous referee for her/his discussion of this point.
7. Koenig and Davis (2001) discuss lexical semantic distinctions of this type, which they call "sublexical modalities." They hypothesize that semantically-based linking constraints are insensitive to sublexical modality, a claim which seems to be corroborated by the account presented in this article.
8. The predicate *forget* seems to provide an obvious counterexample to my claim, as it seems to imply that its subject is no longer committed to the truth of the proposition. This, however, is an inaccurate characterization of the meaning of *forget*. Obviously, the full characterization of the meanings of predicates needs to include the type of eventuality they denote. Thus, for example, PREDICATES OF BELIEF, such as *believe* and *think*, and PREDICATES OF RETAINING KNOWLEDGE, such as *know* and *be aware*, denote the state of maintaining the epistemic commitment. PREDICATES OF ACQUIRING KNOWLEDGE, such as *discover* and *find out*, denote the achievement of establishing such a commitment; they entail that their subjects made the commitment at some time t, the time of discovering, or of finding out. By the same token, *forget* is a CHANGE OF STATE PREDICATE, which denotes the change from the state of knowledge to the state of ignorance. Thus, it entails that its subject was in a state of epistemic commitment to the truth of the proposition before time t, where t is the time of forgetting. One cannot forget something one did not know in the first place.
9. In some cases, the sentences themselves betray the semantic distinctions which determine the distributional patterns. The following contrasts are especially revealing:

   (i) a. We would like to guarantee that our flights will never be delayed, and in fact the vast majority of flights do depart on time.
   b. If disaster strikes, such as a flood or a burst pipe, you can guarantee it will be at the most inconvenient time.

(ii) a. Literary, second-sighted, sick, she holds out a hand to him. He claps it, but decides he can't go on.
b. However, a recent case decided that this degree of formality is not essential.

10. Sometimes, *that*-clauses may appear on the surface, with the proper intonation, as zero-exclamatives, for example, *That I should live to see this!*. This is definitely not the same type of speech act. I thank an anonymous referee for making this note.
11. In an insightful semantic analysis of the distributional patterns of the sentence pronouns *it* and *so*, Cushing (1972) invokes a semantic distinction, which at first sight seems to be identical to mine, between propositions and asserted propositions. As far as I can tell, however, the two seemingly-identical distinctions capture two different semantic intuitions. For Cushing, asserted propositions are the objects of "acts of adopting a definite stance with respect to (their) truth or falsity" (Cushing 1972: 189), whereas propositions are the objects of "weaker," indefinite, passive states-of-mind. Thus, *theorize*, *expect*, *predict*, and *state* take asserted propositions, whereas *suppose*, *guess*, and *think* take propositions. Cushing's distinction, then, is a finer-grained explanatory notion than my own. For me, a proposition is asserted if a cognitive agent makes a truth claim with respect to it — REGARDLESS of the robustness of the truth claim. In other words, all the above predicates, the "stronger" and the "weaker" ones, take asserted propositions.

References

Biber, Douglas (1999). A register perspective on grammar and discourse: variability in the form and use of English complement clauses. *Discourse Studies* 1(2), 131–150.

Bolinger, Dwight (1972). *That's That*. The Hague: Mouton.

Cushing, Steven (1972). The semantics of sentence pronominalization. *Foundations of Language* 9, 186–208.

Erterschik-Shir, Nomi (1977). On the nature of island constraints. Unpublished doctoral dissertation, MIT.

—(1998). *The Dynamics of Focus Structure*. Cambridge: Cambridge University Press.

Farkas, Donka F. (1990). On the semantics of subjunctive complements. In *Romance Languages and Modern Linguistic Theory: Papers from the 20th Linguistic Symposium on Romance Languages*, Paul Hirschbueler and Konrad Koerner (eds.), 69–104. Amsterdam: John Benjamins.

Giorgi, Alessandra; and Pianesi, Fabio (2000). Sequence of tense phenomena in Italian: a morphosyntactic analysis. *Probus* 12, 1–32.

—; and Pianesi, Fabio (forthcoming). Imperfect dreams: the temporal dependencies of fictional predicates. To appear in *Probus*.

Grimshaw, Jane (1997). Projection, heads and optimality. *Linguistic Inquiry* 28(3), 373–422.

Gumperz, John; and Levinson, Steven (1996). *Rethinking Linguistic Relativity*. Cambridge: Cambridge University Press.

Hawkins, John (2001). Why are categories adjacent. *Journal of Linguistics* 37, 1–34.

Hooper, Joan. B. (1975). *On Assertive Predicates*. Syntax and semantics 4. New York: Academic Press.

Karttunen, Lauri (1978). Syntax and semantics of questions. *Linguistics and Philosophy* 1, 3–44.

Kayne, Richard (1981). ECP extensions. *Linguistic Inquiry* 12(1), 93–133.

Kiparsky, Paul; and Kiparsky, Carol (1979). Fact. In *Syntactic Argumentation*, Donna Jo Napoli and Emily Rando (eds.), 345–369. Washington DC: Georgetown University Press.
Koenig, Jean-Pierre; and Davis, Anthony R. (2001). Sublexical modality and the structure of lexical semantic representations. *Linguistics and Philosophy* 24, 71–124.
Levin, Beth (1993). *English Predicate Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.
Lucy, John (1992). *Language Diversity and Thought: A Reformulation of the Linguistic Relativity Hypothesis*. Cambridge: Cambridge University Press.
Pesetsky, David (1998). Some optimality principles of sentence pronunciation. In *Is the Best Good Enough*, Pilar Barbosa, Fanny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky (eds.), 337–383. Cambridge, MA: MIT Press.
Ramchand, Gillian Catriona (1998). Deconstructing the lexicon. In *The Projection of Arguments: Lexical and Compositional Factors*, Miriam Butt and Wihelm Geuder (eds.), 65–97. Stanford, CA: CSLI Publications.
Rohdenburg, Günter (1996). Cognitive complexity and increased grammatical explicitness in English. *Cognitive Linguistics* 7(2), 149–182.
Searle, John R.; and Vanderveken, Daniel (1985). *Foundations of Illocutionary Logic*. Cambridge: Cambridge University Press.
Snyder, William; and Rothstein, Susan (1992). A note on contraction, case and complementizers. *Linguistic Review* 9(3), 251–266.
Stowell, Tim (1981). Origins of phrase structure. Unpublished doctoral dissertation, MIT.
Thompson, Sandra A.; and Mulac, Anthony (1991). The discourse conditions for the use of the complementizer *that* in conversational English. *Journal of Pragmatics* 15, 237–251.
Underhill, Robert (1988). The discourse conditions for *that*-deletion. Manuscript, San Diego State University.
Zwicky, Arnold (1971). In a manner of speaking. *Linguistic Inquiry* 11(2), 223–232.