A Discourse Analysis of *Say* versus *Tell* in Spoken American English*

Yong-Yae Park

This study explores semantic and functional differences between *say* and *tell* in authentic spoken discourse using contextual analysis (Celce-Murcia, 1980). Through frequency distribution and qualitative analysis, the paper demonstrates that the difference between the two is not only in the syntactic constraint as most reference grammars emphasize, but also in their semantic and distributional distinctions, which is closely related to the interactive functions of each verb and the speaker’s communicative goals. For ESL/EFL teaching, this paper will suggest that grammar be presented at the discourse level, illustrating how the use of a form contributes to the speaker’s communicative goals.

1. Introduction

This paper concerns two quotative verbs, *say* and *tell*, in spontaneous spoken discourse. It will focus on their semantic and functional differences with reference to reported speech in authentic discourse, by employing contextual analysis as the methodology (Celce-Murcia 1980, 1990).

The two quotative verbs, *say* and *tell*, have been presented predominantly in the context of reported speech by most ESL reference grammars (e.g., Alexander 1988, Maclin 1981, Murphy 1989). Yet, similar to other characteristics of reported speech such as back-shifting of tense; changes of pronouns; and deixis, including adverbials of time and place; what is available is not much more than sets of mechanical rules. Additionally, it is noted that the verb *tell* is syntactically constrained in that it must be

---

*I am grateful to Marianne Celce-Murcia for her insightful comments. I would also like to thank anonymous reviewers for their careful reading and suggestions. I am thankful to Susan Strauss for her helpful comments and editing. However, I am responsible for any errors.*
followed by an NP indicating who the person is talking to, as suggested in Alexander (1988: 286-287):

"... tell must be followed by a personal indirect object (i.e., tell somebody). Say can be followed by an optional to+the person who is addressed."

"If we need to mention the listener, tell+indirect object is generally preferable to say+to someone."

While Alexander's account seems to capture the difference between these two verbs at the surface level, thus serving as a practical guideline for ESL learners at the beginning stage, it does not provide an adequate explanation of the semantic or the pragmatic functions of the two verbs. In addition, it does not consider authentic language contexts, which would be crucial to the understanding of the actual use of these verbs in discourse. In this study, I will attempt to explore and identify the differences between the two quotative verbs, say and tell, through a frequency analysis to demonstrate their syntactico-semantic differences; and a qualitative analysis to examine their actual uses in both conversational discourse and elicited narrative discourse. The research questions I will pursue are: (1) How are the two verbs used in authentic spoken discourse?; (2) What are the semantic differences between say and tell?; (3) How do the semantic differences contribute to explaining the pragmatic differences between the two forms in discourse?

2. Literature Review

2.1. On Say versus Tell

When we look for the semantic differences between any two forms used in similar ways in a particular language, one method of examining such differences is to analyze the syntactic distributions of the target forms, and thereby deduce the meaning(s) of each. In this regard, the study by Dirven,  

1 Perhaps it is enough to mainly introduce syntactic characteristics of the two verbs to learners at the beginning level. However, for a better understanding of the use of these verbs in different contexts, learners at the later stage should be aware of the semantic and pragmatic differences of them.
Goossens, Putseys, and Vorlat (1982), which is based on a syntactico-
semantic analysis of four linguistic action verbs including *say* and *tell*,
deserves our attention.

Using Fillmore's (1977) framework for Scenes-and-Frames Semantics,
Dirven et al. performed an extensive analysis of *speak*, *talk*, *say*, and *tell* as
linguistic choices or frames for the scene of linguistic action. In so doing,
they attempt to determine how each verb perspectivizes the scene of [that]
linguistic action (p.2). According to Dirven et al., the scenes are related to
our conceptual world and the frames [are related] to the linguistic means of
evoking the concepts (p.1). To establish a comparison of each verb in
relation to the scene of the linguistic action under investigation, they
provide components and a schema for the scene, employing the categories
of Jacobson's (1960) communication model: an addressor or speaker, a
message, an addressee or hearer, a channel, a code, and the linguistic action
itself. Each communicative component, then, is connected to its syntactic
function in a sentence such as a sender to a subject or message to an
object in a sentence. By examining the syntactic characteristics through a
syntactico-semantic analysis, Dirven et al. provide the basic meaning for
each verb and explain the differences between them. In terms of the
meanings of *say* and *tell*, they suggest the following: The verb *say* has a
basic meaning of "express or utter something by means of language (p. 169)," whereas the verb *tell* has the basic meaning of "inform" or narrate.
Their findings, based on this syntactico-semantic approach, contribute to a
better understanding of the differences between the four verbs and also
shed significant light on the differences between *say* and *tell*.

In terms of the difference between *say* to and *tell*, unlike the descriptive
distinction mentioned earlier which indicates that *tell* must be followed by a
personal indirect object, while *say* can be followed by an optional *to* + the
person addressed, Dirven et al. present the following examples which make
the differences between the two verbs more salient (p. 141):

(1) a. He *said* to them that traffic was rather heavy.
(1) b. He *told* them that traffic was rather heavy.

(2) a. He *said* to them that he had been offered the job.
(2) b. He *told* them that he had been offered the job.
In sentences (1) and (2), *say* and *tell* seem to be in free variation. However, when additional context is provided as in (3) and (4), *say* seems to be preferred in (3), whereas *tell* is preferred in (4):

(3) After they had been watching the intersection for several minutes he **said to** them that traffic was rather heavy.

(4) As they did not know it yet, he **told** them that he had been offered the job.

Dirven et al. explain the reason why *tell* is preferred over *say* in 4) by indicating that “the basic function of *tell* is to impart new information to the addressee (p.141).” By contrast, *say*, as in 3), may also be used “to express in words what both speaker and addressee already know (p.141).”

While Dirven et al. seem to provide a comprehensive explanation of the semantic characteristics of the two verbs based on their syntactic distribution, their study still has some limitations: First, their use of examples remains at the sentence level. In other words, it lacks an explanation of the use and pragmatic functions of each verb in various contexts such as direct reported speech at the discourse level. Second, their analysis is based on sentences taken from the present-day English theater corpus. One serious drawback of this type of data is that they are not truly authentic, in the sense that they are aesthetic written constructions and not spontaneous spoken language. It will be the goal of this study to attempt to supplement Dirven et al.’s insightful research on the two quotatives *say* and *tell* by examining authentic spontaneous spoken data at the discourse level, using the same semantic distinctions that they propose.

The next section will examine the most recent studies on direct and indirect reported speech, since *say* and *tell* occur primarily in these contexts.

### 2.2. On Direct and Indirect Reported Speech

**2.2.1. General Studies**

The different views on reported speech seem to follow a continuum. At one extreme end is the view that direct speech is “giving the exact words that someone utters or has uttered in speech or in writing (Quirk et al. 1985:1021).” Coulmas (1986) also views direct reported speech as a verbatim rendition of what was said. These would be the most traditional...
and typical views on reported speech, and it is these views that provide a firm basis for the drills and practices in most ESL/EFL textbooks.

At the other end of the continuum is the viewpoint that reported speech is “constructed dialogue,” as suggested by Tannen (1989). Tannen indicates that the distinction between direct and indirect speech is not really as clearcut as it may seem in actual discourse. According to Tannen, even seemingly indirect discourse can be, in a way, “a representation of the speaker’s actual words (p.99),” and what might appear to be ‘direct’ quotation is really “constructed dialogue,” that is, primarily the creation of the speaker rather than the party quoted (p.99). “She asserts that taking information uttered by someone in a given situation and repeating it in another situation is an active conversational move that fundamentally transforms the nature of the utterance (p.105).” Tannen then presents examples of reported speech taken from casual conversations between family members and friends to show that it is not in fact reported speech but constructed dialogue, and demonstrates how this constructed dialogue is used to create listener involvement.

Views suggested by Wierzbicka (1974) and Clark and Gerrig (1990) would fall in between these two extreme ends. It seems that Wierzbicka’s view is closer to verbatim reproduction, while Clark and Gerrig’s view gives the speaker more power to adjust what was originally said. Wierzbicka calls attention to the theatrical, playful and imaginary features of reported speech. According to her, “the person who reports another’s words by quoting them, temporarily assumes the role of that other person, ‘plays his part,’ that is to say, imagines himself as the other person and for a moment behaves in accordance with this counter-factual assumption (p.272).” In terms of the distinction between direct/indirect speech, Wierzbicka explains, “in direct quotation one assumes the role of the original speaker, i.e. one imagines oneself as that original speaker; in indirect speech one undertakes to state the content of the speech as though one were prepared to assert it oneself, that is to say, one imagines that one wants to assert (ask, etc.) here, now, to the present addressee, whatever the first speaker asserted (asked, etc.) when he spoke to his addressee (p.284-285).” This view is adopted and developed later by Li (1986), which will be examined in the next section.

Clark and Gerrig consider quotations as demonstrations among the three devices of communication or language use, i.e., demonstrations, descriptions, and indications. According to Clark and Gerrig, quotations share the characteristics of demonstrations in that they are non-serious actions and
selective depiction, as opposed to serious actions, which are really or actually occurring. Since a quotation cannot possibly be what is actually or literally occurring, it is a transformation of serious actions. In the same vein, Clark and Gerrig consider quotations as selective depictions in that speakers can depict the words, the linguistic acts, or the meaning of delivery of the original speaker, though only selectively.

Having presented this continuum of current views on reported speech, it seems that one cannot completely exclude the feature of verbatim reproduction of reported speech, even if one were to strictly adhere to Tannen's view of "constructed dialogue." Instead, I would rather adopt the term "reconstructed dialogue" to capture the essence of direct reported speech, and henceforth use this term exclusively to refer to direct reported speech. This will become particularly salient after the examination of authentic narrative data.

2.2.2. Functional Approaches

In this section, I examine recent functional approaches on reported speech such as those by Li (1986) and Yule (1992). Supporting Wierzbicka's (1974) theatrical feature of reported speech, Li (1986) presents the communicative situations for both direct and indirect speech as follows (p. 38, modified):

Figure 1.

a) in direct speech: i) reporter-speaker identifies reported speaker
   ii) reporter-speaker acts as the reported speaker, providing the reported speakers form, content, and non-verbal messages.

b) in indirect speech: i) reporter-speaker identifies reported speaker
   ii) reporter-speaker provides his/her own form and non-verbal messages, while taking the reported speakers content.

According to Li, in indirect speech, the reporter-speaker can convey his/her own voice and comments through the form and non-verbal messages. In other words, indirect speech is a transformed form of the message with the reporter-speaker's comment or attitude about the content of the reported speech. We could infer from this, then, that each quotative used for indirect speech also carries the function of expressing the reporter-speaker's attitude. In this sense, the reporter-speaker can choose a quotative as a part of the form of reported speech for his/her communicative goal, while
still conveying the same original content. Li also reports that direct speech occurs more frequently in narrative and that it expresses more of the reporter-speaker’s involvement in the events s/he is reporting, following Chafe (1982).

Other studies by Yule and Mathis (1992; also Yule, to appear; Yule et al., 1992) deserve our attention in that they examine the phenomenon using natural, spontaneous conversation data at the discourse level. Yule and Mathis (1992) introduce the concept of staging and constructed dialogue, explaining how reported speech functions to establish the speaker’s topic in narratives. They also show how the quotatives, say, tell, and be:like are used to contribute to that function in discourse. They observe a pattern of reported speech, involving reported events, reported speech, and constructed dialogue as in the following extract (pp. 207-208):

Figure 2.

<table>
<thead>
<tr>
<th>Reported events</th>
<th>Reported speech</th>
<th>Constructed dialogue</th>
</tr>
</thead>
<tbody>
<tr>
<td>And Allan told him</td>
<td><strong>and he said</strong> Don fell out laughing</td>
<td><strong>and he’s like</strong>: oh really?</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>And Allan said him</td>
<td><strong>and he said</strong> Don was just falling out</td>
<td></td>
</tr>
</tbody>
</table>

Here, Yule and Mathis point out that the reporting verb told is used to summarize the least important material and as such falls under the category of reported events; for the category of reported speech, the verb said, in the past tense only, is used to attribute descriptions of salient events to the reported speaker. For Yule and Mathis, the most theatrical of the three constructions, be:like, in the present tense, is used to introduce constructed dialogue and to create an impression of a two-party conversation. However, while Yule and Mathis maintain that say only occurs under the category of reported speech, it will be shown later, that it may also be used for constructed dialogue.

Later, Yule (to appear), pointing out the lack of attention given to the functions of reported speech in previous research, suggests that such
changes in form, (i.e., from *told* to *said* to *be:like*) are aimed at making the indirect speech report more "distant from the speaking event being reported (p.2)." This distance provides indirect speech with a function like "a narrative account of an event" or a report, and renders it distinct from direct speech which provides the dramatic presentation of the event (cf. Wierzbicka's (1974) theatrical function). Yule further explains that when the indirect speech is simplified into a summary of a speaking event, it creates more distance between the speaking event itself and the report, yielding greater control to the reporter for the interpretation of the event. This was also supported by Li (1986), as evidenced earlier in Fig.1. This distancing effect, then results from the reporter-speaker providing the form and non-verbal report, after taking the reported speaker's content. The following illustrates the functional distinction between three types of reporting formats in English proposed by Yule (to appear):

**Figure 3.**

<table>
<thead>
<tr>
<th>Summarized Report</th>
<th>Indirect Speech</th>
<th>Direct Speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lara called me on the phone about the project.</td>
<td>She asked if they could finish that project the following week.</td>
<td>&quot;Oh, Dan, can we please finish this project next week.&quot;</td>
</tr>
</tbody>
</table>

Yule observes that the three reporting formats listed here (Summarized Report, Indirect Speech, Direct Speech) constitute the type of sequence, often found in spoken reports beginning with a summary, and continuing to indirect speech and then moving into direct speech. This type of sequential development will be observed in section 4.2.

With reference to the studies on direct and indirect speech examined above, we can relate the basic meanings of *say* and *tell* suggested by Dirven et al. (1982) to the use of each verb in a pattern of reported speech in discourse observed by Yule and Mathis (1992). That is, when the speaker uses direct speech, the speaker would use *be:like* or *say* to provide the reported-speaker or the original speaker's form, content, and non-verbal messages (Wierzbicka, 1974; Li, 1986). This means that the speaker would use direct reported speech and employ quotatives to make the least transformation of what was said. When we consider the quotative *say*, the meaning of utter or express seems to be appropriate for this purpose. In
other words, for the speaker to make the least transformation of what was said, s/he would employ the verb say with its basic meaning of utter. In contrast, tell is used in reported events to summarize the least important material (Yule and Mathias, 1992). From its basic meaning of inform, we can infer that tell is used when the speaker needs more transformation on what was said. This also conforms to Li’s (1986) explanation on indirect speech (Fig.1) or to Yule’s (to appear) explanation that indirect speech creates more distance between the event and the report, giving more control to the reporter for the interpretation of the event.

As demonstrated thus far, it is clear that speakers employ and manipulate different linguistic vehicles to best serve their communicative goals. Therefore, in studying the use of competing forms, we should consider the various contexts in which these forms occur to see precisely how each form contributes to the speaker’s particular communicative goal.

3. Data and Methodology

3.1. Data

The data for this study consist of the following:
1. Personal Conversation data: The data were originally collected and transcribed by graduate students for conversation analysis courses at a university in the U.S. The database includes naturally occurring telephone and face-to-face conversations between close friends or family members in relatively casual situations. The data have many story-telling sequences or narratives inserted in the conversation, and all are either two-party or multi-party conversations. The speakers in each data set are all native speakers of American English.
2. Elicited Narrative data: These data were taken from 1) Terkel’s book, Working, in which people from different professions talk about their jobs in their personal narratives; 2) transcripts of personal narratives: These narrative data were collected through interviews.

Table 1 gives the breakdown of the total number of words for each type of data in the corpus:
<Table 1> Corpus Size

<table>
<thead>
<tr>
<th>Data</th>
<th>Total Number of Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>(I) Conversation</td>
<td>26,665</td>
</tr>
<tr>
<td>(II) Elicited Narrative</td>
<td>61,324</td>
</tr>
<tr>
<td>Total Words</td>
<td>87,989</td>
</tr>
</tbody>
</table>

3.2. Methodology

This study was conducted within the framework of Contextual Analysis (Celce-Murcia (1980, 1990)). Contextual Analysis is a type of discourse analysis which examines a linguistic form (or two ostensibly similar forms) using an appropriate corpus or other means in order to determine where, why, and how frequently that form occurs in discourse, be it written or spoken, planned or unplanned, spontaneous or elicited. A major advantage of Contextual Analysis is that it avoids the researcher's reliance on his/her intuition, which is "inevitably influenced by the researcher's theoretical orientation (Coates 1983: 3 quoting Labov, 1971)." It is, therefore, extremely important to examine real language data to connect the rather abstract concepts of grammatical items with the contexts in which they are realized and to show the actual use or function of such linguistic forms in everyday language.

For this study, all tokens of *say* and *tell* were isolated, counted, and classified in terms of various syntactic categories. A frequency analysis was performed to examine the syntactic distributions of the two quotatives, adopting Dirven et al.'s categories, and a qualitative analysis was performed to examine their functions in context.

4. Analysis

4.1. Frequency Analysis: Syntactico-semantic Analysis

4.1.1. Examples of each coding categories

In order to perform the frequency analysis, I replicated the framework used in Dirven et al.'s study. Since Dirven et al. examined four linguistic action verbs, *speak, talk, say*, and *tell*, and since each verb was analyzed by a different person, the exact categories for each verb were different from each other. For the sake of consistency, I adopted the categories
established by Goossens (1982) for the verb *say*, and then applied them to both *say* and *tell*. These categories can be divided into three types: direct speech, indirect speech (embedded clauses), and miscellaneous. This last category includes a variety of syntactic forms which follow both verbs. Examples (5) through (20) are actual excerpts from the data illustrating how both *say* and *tell* were categorized and analyzed:

**A. Direct speech**

As evident by the title, this category is for utterances which constitute a direct (or quasi-direct) enunciation of the original speech:

(5) (Gee) [conversation]
It's suddenly they *said* well, we want cha to work– days now full: time b'cuz we really need you.

(6) (Working) [Elicited Narrative]
I *told* him, "Who the hell are you, Hitler? What is this Yes, sir bullshit? I came here to work, I didn't come here to crawl. There's a fuckin' difference."

As in Goossens (1982), I also included nominal phrase objects as reported speech in this category, e.g., "say thank you" and "say hello," and categorized them as exceptional cases. However, I did not include expressions like *say things right/tell me a story/tell me a joke*, but instead placed them in the category of nouns/nominal phrases, since they are more related to the gist of what was said or told rather than to the words that were actually uttered.

**B. Indirect speech**

This category includes the occurrences of indirect speech either with complementizer *that* as in example (7), or without a complementizer as in (8).

(7) (TG) [conversation]
yihknow, the guard jus' *told* me that this wz the building.

(8) (Working) [Elicited Narrative]
The little girl, after she did it, she *said* she was sorry.
C. to-infinitive

There is only one occurrence of *say* in this category (example (9)), which functions in a similar way as *tell* in example (10), and is thereby considered an exception:

(9) (ME) [conversation]
But we had some leftover. [En he said to freeze it.

(10) (Working) [Elicited Narrative]
The agency said, "She's too fat, tell her to lose weight."

D. Wh-clause

As will be shown in Tables 2 and 3, the frequency of *say* in this category is very low. Utterances in this category are all comprised of the verb *say* or *tell*, followed by some type of wh-headed clause.

(11) (Working) [interview]
They wouldn't let you *say* how you'd like your hair cut, they wouldn't let you have your own personality, your makeup, your clothes.

(12) (Working) [Elicited Narrative]
For forty-five minutes they *tell* you what they want. They explain and explain and you sort of tune out and do the same thing.

E. Pronominal items

This category basically includes instances if *say* or *tell* followed by some type of pronoun as in:

(13) (Working) [Elicited Narrative]
I said, "Holiday Inn." I *said* it because we're not Holiday Inn, I was just fooling around.

(14) (After the Movie) [conversation]
W: I loved the Alamo. The Alamo doesn't have a basement.
H: =M-hm-hm-hm.
(1.0)
W: hh They won't *tell* you that in school hh[hh
C: [Buenos dias.
F. Nouns/ Nominal phrase

This category contains the use of _say_ and _tell_ as they are used with the same syntactic category, i.e., nouns or nominal phrases, in different ways: With _say_, the noun(s) or noun phrases which follow are quantifying pronouns such as _something, anything_, or noun phrases with the words _things_ or _words_. This type of noun or noun phrase is different from the nouns which were included in the category of direct speech in that they focus more on content rather than on form. That is, they provide a summary or an account of the characteristics of what was said rather than what was actually or literally said. On the other hand, the meaning of _tell_ in this category is clearly "narrate" (Putseys, 1982), and as can be seen in example (16), the verb is used with nouns such as "stories" or "jokes":

(15) (TG) [conversation]
I finally said something right.

(16) (AD) [conversation]
Mike: Hey I started t' tell you a joke
Ryan: Hey tell yer jo:ke,

G. Prepositional phrases

The fact that there is no occurrence of _say_ in this category shows a clear difference in both the syntactic constraints as well as the semantics of the two verbs. The meaning of _tell_, being to inform, as proposed by Dirven et al. would explain its use with prepositional phrases (See example (17)).

(17) (TG) [conversation]
Bee: hhhhh So she tol' me of a place on Madison Avenue 'n Seveny Ninth Street.=

---

2 This category includes the use of prepositional phrases such as "tell someone about something," "tell someone of something," "tell on someone," and "tell against someone" where they introduce the topic of speech act (Putseys 1982).
H. Miscellaneous

This category deals with exceptional cases for each verb, such as say to, functioning to deliver what was said without an actual transformation to direct speech:

(18) (Working) [Elicited Narrative]
After a week later, I said to her, “You ng lady, I was gonna Quit, but I wouldn’t do it for the likes of you.”

Instances of tell in this category include its idiomatic use, meaning to distinguish, to understand:

(19) (AD) [conversation]
Mike: You know why a Polak paints iz garbage cans red’n yellow,
Gary: So ‘e c’n tell it from iz flowerpots?

(20) (MD) [conversation]
D: It’s hard to tell.

4.1.2. Discussion

The results of the frequency analysis are shown in Tables 2 and 3, with Table 2 providing the distribution of say and tell in conversation, and Table 3, their distribution in interviews:

<Table 2> Distribution of Say and Tell in Conversation (percentage (Number of tokens))

<table>
<thead>
<tr>
<th></th>
<th>SAY (N=119)</th>
<th>TELL (N=56)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. direct speech</td>
<td>0.58 (69)</td>
<td>0.02 (1)</td>
</tr>
<tr>
<td>II. indirect speech</td>
<td>0.30 (36)</td>
<td>0.35 (20)</td>
</tr>
<tr>
<td>III. Other syntactic Patterns</td>
<td></td>
<td></td>
</tr>
<tr>
<td>to-infinitive</td>
<td>0.01 (1)</td>
<td>0.09 (5)</td>
</tr>
<tr>
<td>wh cl.</td>
<td>0 (0)</td>
<td>0.07 (4)</td>
</tr>
<tr>
<td>pronominal items</td>
<td>0.07 (9)</td>
<td>0.14 (8)</td>
</tr>
<tr>
<td>nouns/nom. phrase</td>
<td>0.02 (2)</td>
<td>0.13 (7)</td>
</tr>
<tr>
<td>prep. phrase</td>
<td>0 (0)</td>
<td>0.13 (7)</td>
</tr>
<tr>
<td>misc.</td>
<td>0.02 (2)</td>
<td>0.07 (4)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>
As expected, there is a high frequency of use of *say* in direct speech in both data types. This reflects the nature of spontaneous spoken data in which speakers tend to use more reconstructed dialogue to make their stories more vivid and effective as Chafe (1982) and Li (1986) observe. This can also be related to the findings that more direct reported speech is used in narratives, especially in the climax of the narratives to provide more vividness (Li 1986).

In both tables, it is also notable that the use of *say* is skewed toward direct speech, occurring 58% in the conversational data and 67% in the elicited narrative data. By contrast, the use of *tell* seems to be skewed more toward indirect speech, showing a frequency of 35% in the conversation data and 30% in the elicited narrative data, yet the overall distribution is more varied. Compared with the results of the Dirven et al.'s study, the frequency of use of the verb *say* in direct speech in these data is strikingly high (58% in the conversation data and 67% in the interview data), while in the Dirven et al.'s study, the frequency was only 29.85%. This difference could be accounted for by the different type of data used: Since the Dirven et al.'s data, taken from a theater corpus, are not truly authentic spoken discourse, it could be presumed that this corpus is planned as literary discourse and not spontaneous, and hence would have fewer and less vivid story telling narratives than more authentic conversational data. This will be examined further in the next section.

It is striking that the use of *say* is skewed toward the first and the
second categories, i.e., direct and indirect speech (88% in both data types), whereas the use of tell is skewed toward the third category (63% in the conversational data and 65% in the elicited narrative data). The third category is established for the use of say or tell with various syntactic patterns. What the speaker can convey by using these syntactic patterns such as to-infinitive or prepositional phrases is the gist or the summary of what was said. The fact that say is used much less in the third category demonstrates that speakers would prefer not to use say for this purpose because of its basic meaning "utter." It can be seen that in using these syntactic patterns, the basic meaning of say "utter" would be weakened. Rather, a marginal meaning such as "state" can be proposed to explain the occasional use of say to convey the content of what was said, as Dirven et al. indicate. In contrast, tell is used more frequently with these categories, to which the underlying meaning of tell conforms well. That is, the meaning of "inform," or "impart new information," becomes clear when the speaker uses tell with syntactic patterns such as prepositional phrases in that the speaker performs more of a transformation in reformulating what was said. Thus, the different types of skewing of say and tell in their syntactic distributions actually reflect the semantic distinctions between the two verbs, confirming Dirven et al.'s study.

Considering the use of say with various syntactic patterns in the third category in additional detail, note that there is no use of prepositions with say either in the conversation and elicited narrative data or in Goossens' data. As was indicated (Goossens 1982), this tendency is related to the underlying meaning of say. That is, say with the meaning of "to utter or express" seldom co-occurs with prepositional phrases such as "*say about." It is also observed that in each data set, there are only a few uses of say to + indirect object, as a counterpart of tell, as in "I said to her that I would go." Other categories such as to-infinitive and wh-clauses also occur more frequently with the verb tell than with the verb say, as in "He told me to come back early." In fact, of the 119 total tokens for say in the conversational data, only one token occurred with a to-infinitive and Ø with a wh-clause; and of the 312 tokens of say in the elicited narrative data, Ø occurred with a to-infinitive and only two with wh-clause.

Next, Table 4 illustrates how frequently different quotatives are used in direct reported speech, which seems to be one of the major contexts where say is used:
In fact, say is used with an extremely high frequency in the conversation data (79%) and even more frequently in the elicited narrative data (87%). This shows that for direct speech, say is the most preferred form. This highly salient use of say in direct speech can be related to the underlying meaning of say, in the sense of “to express or utter something by means of language” (Goossens 1982). Note also the strikingly low frequency of use of the verb tell in direct speech. This is related to the function of the verb tell, which is to summarize or paraphrase, derived from its basic meaning of “to inform” (Putseys 1982). In this sense, it is difficult to imagine that tell, with its function of summarizing or paraphrasing, would fit into direct speech as (re)constructed dialogue. In accordance with the findings of the previous literature (e.g., Yule et al., 1992; Yule, to appear), quotatives such as go and be:like are used only in reconstructed dialogue. In the data for the present study, these two quotatives were used far more frequently in the conversation data than in the interview data, reflecting the interactional nature of conversation, which allows for more vivid and informal use of quotatives. Another reason for this skewing would be the age variables of the speakers: In the conversation data, most speakers were in their twenties, whereas in the interview data, the speakers were much older. In fact, Yule and Mathis (1992) report that the main quotatives for constructed dialogue are go and be:like. In my data, however, say is the most preferred form for quotatives in direct reported speech.

The next section will examine how the two quotatives are used to serve the speakers communicative goals in actual contexts.

4.2. Qualitative Analysis: Functional approach

First, with the examples from the data along with the categories in the frequency analysis, I will try to recapitulate in further detail the differences
between *say* and *tell*:

The examples for direct and indirect reported speech, the first two categories used in the frequency analysis, do not seem to reveal the differences between *say* and *tell* at the sentence level. The examination of these forms occurring in segments at the discourse level will be presented later.

Example (21) is the example which belongs to the category of direct speech, showing that *say* is used with the actual utterance or the pronouncing of an actual word. It shows that what H meant by "say adobe" was the actual pronunciation of the word, confirming the meaning of *say* as "utter" or "express." As was also indicated in Dirven et al., the difference between *say* and *tell* can be found in "Say your name, please." versus "Tell me your name."

(21) (AM) [conversation]

W: ’n they all just kept repeating everything. Tortillas. Hhhhh ’ts like an ESL classes. Hhh

(H)

H: Can you *say* ado:be:?

W: =h[h

C: [Eh heh heh heh hh hh

H: A- DO:BE::=

Next, let us examine the following examples which belong to the third category in the frequency analysis, that is, the use of *say* and *tell* with various syntactic patterns:

As shown earlier, there is only one use of *say* with to-infinitive as in example (22), and this usage basically constitutes an exception in that *say* in this example is used in the same manner as the verb *tell* in (23). In contrast, the use of *tell* with to-infinitive, functioning as a directive, is more frequent (9% in the conversation data and 10% in elicited narrative data).

(22) (ME) [conversation]

Marsha:-> But we had some leftover. [En he *said* to freeze it.

(23) (Working) [elicited narrative]

The agency said, She’s too fat, *tell* her to lose weight.
As was mentioned, with complement categories such as to-infinitive, wh-clause, and prepositions, the frequency of say is very low. This low frequency in syntactic distribution also seems to stem from the meaning of say. That is, with these categories, the role of quotatives as condensations of message or delivery of contents rather than actual form is more prominent. The quotative say in this use can be paraphrased as “express” or “state” not as “utter” (Goossens 1982). Yet, the wh-clause with say in example (24) seems to undergo less transformation of actual speech, compared with the wh-clause with tell in example (25), where the wh-clause seems to be the summary of what was said.

(24) (Working) [elicited narrative]
They wouldn’t let you say how you’d like your hair cut, they wouldn’t let you have your own personality, your makeup, your clothes.

(25) (Working) [elicited narrative]
For forty-five minutes they tell you what they want. They explain and explain and you sort of tune out and do the same thing.

If we examine the use of the quotatives with prepositions, no occurrence of say with prepositions is found. By contrast, tell in this use seems to fit well with its underlying meaning of inform as in example (26):

(26) (TG) [conversation]
Bee: hhhhh So she tol’ me of a place on Madison Avenue ’n Seventy Ninth Street.=

Lastly, it is interesting to see that the quotative verb tell can go through a semantic change taking on the meaning of recognition or cognition as in examples (27) and (28), similar to the semantic change of the verb see, which shifted from a meaning of actual perception to cognition as in I see what you mean.

(27)
Mike: You know why a Polak paints iz garbage cans red’n yellow,
Gary: So ’e c’n tell it from iz flowerpots?

3 Marianne Celce-Murcia, personal communication
No such semantic transformation occurs with say.

Thus far, I have tried to examine the syntactico-semantic distinctions between say and tell in detail with actual examples from the data. In the following section, I will attempt to demonstrate the use of the two quotatives in various discourse contexts. One of the frequent and notable contexts involves story-telling, or narrative sequences. In many narratives, speakers frequently employ direct reported speech to establish what Tannen (1989) refers to as “constructed dialogue,” and it is the quotative say that is used predominantly in this format. Let us examine the following segments taken from the elicited narrative data:

(29) (Working)
It was one o’clock in the morning. A phone call came in. I worked the night shift. And I said, “Holiday Inn.” I said it because we are not Holiday Inn, I was just fooling around. The little girl I worked with turned me in. So the boss called me in. She said, “Why did you do it?” I said, “Just for a lark. It was quiet, nothing to do.” She said, “Fran, you are a good operator and we all love you, but I don’t know why you did it.” I said, “I wanted to have a little fun.”

The little girl, after she did it, she said she was sorry. After a week later, I said to her, “Young lady, I was gonna quit, but I wouldn’t do it for the likes of you.” And she says, “Fran, I admire you because you didn’t say anything to me in front of the boss.” And I said, “Well you’re pretty low because you have done things I would never have told on.”

In segment (29), the speaker is a telephone operator and is talking about an episode at her workplace. Notice that the whole episode consists of reported speech, most of which is direct speech. This use of direct speech makes the story more vivid to demonstrate the speakers involvement in the story, as noted by Tannen (1989) and Chafe (1982). This also supports Tannen’s notion of “constructed dialogue.” That is, the story is (re) constructed as a frame of dialogue either between the speaker (an operator) and the girl or between the speaker and her boss. Here, the speaker employs the quotative say with its basic meaning of “utter” to make the least transformation of what was said. The use of say in the second
instance of reported speech "I said it because..." also indicates that the "utterance" meaning of say is at work here. That is, the pronoun it is used to replace the words "Holiday Inn," thus engenders a minimum transformation.

The same tendency toward the use of say in reconstructed dialogue is found in story-telling in conversational data as in the following segment. In segment (30), the speaker is talking about how she got to her classroom on the first day of class:

(30) (TG)

B: e - en:-d, there wz a ha - et - there wz a hole in the wall in the back a' the building en, eh - there wz another girl walking around

-> she says

hhh uh - dihyou know

where the science building is

-> 'n I said - well I said-

=>>> yihknow, the guard jus' told me

that this wz the building.

-> So she seh -

hh are you goin here fer 'n hh fer in'dian class by any chance

-> 'n I said yes. hh

-> So I said c'mon

we'll fi:nd it together

en we craw::led through the hole in the wa:ll, tch!

Here, speaker B is (re)constructing the episode by the use of dialogue with the almost exclusive use of direct speech. Notice that the dialogue consists of a scheme with contrasting tokens of "she said" and "I said" utterances. What is striking is that there is only one instance of tell. This usage of tell clearly illustrates the difference in focus between say and tell, because the two verbs could basically be interchangeable in this utterance. B is imparting new information to a girl she had met. The information provided by the guard is used to assure the girl that that building is the right building. The important thing is not what the guard "uttered" or actually "said" but the information itself. This line is also embedded in a scheme of "I said" and "she said." Furthermore, this instance of tell is the
only occurrence of indirect speech embedded in this scheme. This instance of *tell* clearly illustrates how the speaker can choose either *say* or *tell*, depending on his/her communicative goals. In other words, B chooses *tell* as a way of providing a more reliable source of information.

Another instance in which *say* and *tell* contrast interestingly in discourse can be seen in segment (31), which takes place at the dinner table, EM is talking about her friends' break-up:

(31) (Halloween Dinner)
Liz and E.M. are in their teens and close friends to each other. Nancy is Liz's mother; and Betsy is E.M.'s mother.

1-> EM: Uh ((ingressive)) O~h did you **tell** your Mom?
2 Liz: what
3 EM: That Jessie and Mellissa broke up,
4 Liz: Oh Mom, Jessie and Mellissa broke up, cause Jessie was beating up
5 Mellissa
6 Betsy: [WHAT!]
7 Nancy: [Jessie [and Mellissa,
9 EM: [You ( )
10 Liz: I like a little action in there,
11-> Betsy: Who **said** that?
12-> EM: That's what Gabby **said** t'm n then she goe's Ingo: (I know) what he di'd, he just threw her against a wa'll n pulled her hair n like .nhh then she was like .nh then um he goes she like was crying like a 15 half cry like
16 .hhh I lo:ve you I lo:ve you I try:

Here, when EM first proffers the topic by asking Liz to talk about the actual story, she uses *tell* in the way that Yule (to appear) suggests for summarized reports, i.e., by summarizing the whole event regarding the break-up. The next instances of *say*, in "who said that?" and "That is what Gabby said to me" draw our attention in that *tell* is also a possible form. Let us examine the utterances more closely. In line 10, Liz is about to talk more about a story she launched at lines 4 and 5. Then, in line 11, her mom Betsy interrupts, asking her who the source of information was. Here, what concerns Besty seems to be not only the source of the information
itself but also who is responsible for the actual wording or the details of the gossip. In line 12, using the same verb *say* in a pseudo-cleft construction, EM emphasizes that what was said was from Gabby; or that Gabby is the person who actually uttered all this, implying that she is just delivering the story and not responsible for it. In contrast with “The guard just told me this was the building” in segment (30), where the speaker uses *tell* to make the provider of the information as a more reliable source, here, EM uses *say* to detach herself from any responsibility for the information. It is evident that the speaker chooses a particular verb to serve his/her particular communicative goal. In this segment, the speakers also use various quotatives such as *go* and *be*:like, which Yule (to appear) reports as typical quotatives for constructed dialogue.

Segment (32), illustrating a rich usage of various quotatives, particularly *say* and *go*, is an excerpt of a narrative in a telephone conversation, where John, the speaker, is telling the listener a story about how he and his roommate’s friend, Donna, made fun of his roommate.

(32) (PJ)

1 John: 'n so finally he **said**, “You talk to Donna, I’m goin to the bathroom.” hhh=so he disappeared to the bathroom, an I **said**, well let’s (.3) see how far this can go. I **said**, “Donna, you wanna step outside for a few minutes?” hhh she **goes**, that’s great. So we ran outside, n we hid (.7) hh **telling** our other roommate to **tell** him that w-we were outside (.4) in the front talking, hh so then he came out of the bathroom n her went out front to find us an hh he couldn’t find us hh so we went back in n we sat n waited for him to come back an he came back, an he’s rully pissed off. he **said**, she **said** something that was really simplistic. I forget what 12 it was. hhh He **goes**, “Oooo:, you are so smart, what were you, a cheer leader in a high school? hhh an she **goes** I can’t believe you **said** that and so I, so I can’t believe he **said** that either.

hh Then, I **said**, well, I’m leaving now= if you wanna go home, Donna, 17 I’ll take you
As is in the previous segments, the speaker uses direct speech almost exclusively, and the quotative verb is mainly say and occasionally go. This alternation also shows what Yule and Mathis (1992) suggested as a format for constructed dialogue may not always be the case. Unlike Yule and Mathis's (1992) observation about the use of be:like for “constructed dialogue,” speakers frequently use the quotatives say or go for the same purpose. This shows that the quotatives used for “constructed dialogue” may not always be the same and can be changed, depending on speakers or contexts. Line 17, “So I, so I can’t believe he said that either,” employs the “zero quotative” (Yule et al. 1992). That is, speakers sometimes delete quotatives in reported speech.

The verb say is also used in discourse to compete for or to claim authority over some type of information. This usage is clearly shown in example (33), where two speakers are planning a surprise party. Lines 1 and 2 show that two interlocutors have different information on whether an invited person, Evie can come or not, which engenders the following sequence of competing for the accuracy and claim of authenticity for the information:

(33) (Kamunsky)

1  K:  Evie can’t come=
2  A: =Well she’ll come la:te.
3
4  K:  We’ll,
5-> A:  Why.=She said not et a:ll?
6
7  K:  Well she’s goin’ out with er folks afterwards hh
8
9  A:  Oh yea:h: We:re,
10
11-> K:  I don’ know, but she says ez long ez her folks er
12  coming to the game
13  they probably wanna go ou:t.=
14-> A:  =’hnh oh because I j- I talked to er Sunday en she
15  said she (c’)/(d’) come after.

The age variable seems to be the strongest candidate for this variation.
What is interesting is that both speakers use the quotative *say* in their competition for the exact information: After they notice the information gap, A asks a question to find out what the potential guest, Evie, actually said in line 5. It turned out that both interlocutors did talk to her earlier. They, then, start to compare what they heard by saying “what she said” in lines 5, 11, and 14. However from line 14 on, it becomes an issue of the time when each one talked to her, i.e., whoever spoke to Evie more recently would presumably have the more accurate information. In this segment, again, what is important is what the original source, Evie, actually said. It is also noticeable that K’s use of the present tense in line 11 enables her to claim authority on the information, which is also contrasted with A’s use of the past tense in line 15; K’s use of the present tense underscores her claim of authenticity since the interaction seems to have occurred more recently that A’s.

So far, I have attempted to show the differences between the two quotatives, *say* and *tell*, in various contexts. The most representative use of *say* was found in narratives, where speakers reconstruct dialogue to make their stories more vivid. The use of *say* in direct speech in narratives can be attributed to this goal of making stories more vivid in that the verb *say* would effect the least amount of transformation on what was said. It was also observed that speakers use the quotatives *say* and *tell* in different ways in discourse contexts where they are theoretically interchangeable, depending on the speaker’s interactional moves. For example, the verb *tell* is used to make the information more reliable as in segment (30), while the verb *say* is used when the speaker wants to detach herself from any responsibility for the information as in segment (31). What is used when both speakers are competing to report the “accurate” information and thereby to claim authority is the verb *say* as in segment (33).
5. Conclusions

I have attempted to show the differences between the verbs *say* and *tell* with reference to reported speech in spontaneous spoken discourse within the framework of Contextual Analysis. For actual analysis, I performed a frequency analysis and a qualitative analysis: In the frequency analysis, the following points were made: The verb *say* is used almost exclusively to utter what was said either in direct speech or indirect speech. No use of the verb *say* was found before prepositional phrases, which serves more to report the gist or summary of what was said. By contrast, the verb *tell* seldom occurs in direct speech. The verb *say* is used mainly for the linguistic action of "uttering or expressing," while *tell* accomplishes other linguistic actions such as imparting new information, asking someone to do something, or summarizing information.

As I tried to show in the qualitative analysis, even in indirect speech, *say* and *tell* serve different functions: First of all, because of the nature of direct speech being "show as well as speech" i.e., its theatrical features, it is used almost exclusively in narratives or story telling. The primary verb used in these contexts is *say* with the meaning of "utter or express." This meaning of "utter or express" can match the function of direct speech very well, where the reporter speaker acts out the reported speaker for form, content, and even non-verbal messages. By contrast, in indirect speech, the reported speaker takes only content from the reported speaker and uses his own form and nonverbal messages. This establishes room for all of the changes in form such as pronouns, time and place adverbs, and tense-shifting according to the speaker perspective or deixis. The reporter speaker is allowed to adjust the words to serve his/her communicative goal. Various types of performative verbs can be used. Through these changes which serve the reporter speaker's goals, the indirect speech is transformed. In terms of the degree of transformation, the verb *say* undergoes the least change. It is related to the meaning, "utter or express" without adding other linguistic actions. When the verb *tell* is used, however, the reported speaker’s content undergoes more transformation; it is summarized and condensed by the reporter speaker. This seems to be more for indirect speech. Similarly, Quirk et al. (1985) observe that indirect speech often paraphrases or summarizes the original speech or writing. The reporter speaker can manipulate the use of *tell* in discourse for his/her own
A Discourse Analysis of *Say* versus *Tell* in Spoken American English

communicative purposes. For example, this use of *tell* is found in story telling, especially at the beginning of the story where the speaker launches the story by providing the summary or general picture of the event such as "Allan told him the story about Bobo."

It is certain that each of these verbs functions in discourse in distinctive ways. In contrast to the traditional descriptions appearing in most reference grammars, the difference is not only the syntactic constraint that the verb *tell* needs a direct personal indirect object, whereas *say*, in general, does not. It has also been shown in this study that the two verbs demonstrate certain semantic and distributional distinctions, which seem to be related to the interactive functions of each verb and the speaker's communicative goals.

This study has focused on conversation and interview discourse. However, it would be interesting to look at the different functions of reported speech and the use of these two verbs in different types of discourse, such as He (1993) has done using data from academic counseling sessions.

For ESL and EFL teaching, it should be considered that putting too much emphasis on mechanical changes at the sentence level might cause us to overlook other important points that also need to be taught (Celce-Murcia 1980). In addition to sentence-level facts, grammar should be introduced and presented at the discourse level, focusing on how the use of a form contributes to the speakers' communicative goals as demonstrated in the paper. The semantic and pragmatic differences suggested in this paper can benefit ESL/EFL teachers and high-intermediate/advanced learners in that the two verbs are used frequently not only in the context of reported speech but in situations as well. For example, teachers can help students differentiate them by using contexts such as meeting someone new. That is, this could constitute the kind of a context where the difference between the two expressions "Can you tell me your name?" and "Can you say your name for me?" becomes very clear: "Can you tell me your name?" would be used when the speaker does not know the hearer's name, while "Can you say your name for me?" would be used when the speaker actually knows the hearer's name but does not know how to pronounce it. This kind of awareness would certainly contribute to learners' communicative competence and would thus shed new light on ESL/EFL pedagogy.
References


Department of English Language and Literature
Seoul National University
San 56-1 Sillim-dong, Gwanak-gu
Seoul 151-745, Korea
E-mail : parky@snu.ac.kr