A Comparative Analysis of Discourse Markers in English Conversational Registers[1]

Naoko Taguchi
Northern Arizona University

Using corpora of spoken American English conversations, the present study examines the use of discourse markers in different spoken registers. Three conversational corpora were selected for analysis: 12 family conversations, 11 professor-student conversations during office hours, and 10 server-customer conversations. Twelve discourse markers were identified based on previous literature, and their occurrences in context were analyzed using the Monoconc concordancing program. Quantitative and qualitative analyses show that there are considerable differences in the frequency distributions of discourse markers. These distribution patterns are interpreted in light of the functions of each discourse marker interacting with the typical characteristics of different conversational registers.

The routine use of discourse markers in conversation is no doubt a major characteristic of spontaneous and interactional spoken language. Discourse markers have attracted wide research interest and have been analyzed from a variety of perspectives (see, e.g., Blakemore, 1987; Brinton, 1990; Erman, 1987; Fraser, 1993; Jucker & Ziv, 1998; Östman, 1981; Schiffrin, 1987; Schourup, 1985; Stenstrom, 1998).[2] Some studies have analyzed a range of discourse markers, drawing general conclusions about the role of discourse markers as a class, while others have concentrated on the identification of individual markers and their core functions in discourse units. Both lines of research generally agree that discourse markers are inserts that are largely independent of the propositional content of an utterance and fulfill pragmatic functions in promoting interactiveness and coherence.

Despite the ample research findings on the types and functions of discourse markers, questions remain as to whether discourse markers exhibit similar distribution patterns across conversational contexts or registers. Recent developments in corpus linguistics have made it feasible to analyze large-scale authentic data (see, e.g., Biber, Conrad, & Reppen, 1998), and such analyses have been extended to conceptualize characteristics of spoken English (Leech, 2000). However, contrastive analysis across spoken registers is still underrepresented due to the limited availability of representative corpora. The present study intends to fill this gap by examining the use of a number of discourse markers across three conversational registers: conversations among family members, professor-student office hour conversations, and service encounters. Because specific characteristics of each situation (e.g., interlocutor relationships, purpose of communication) can greatly influence the patterns of spoken language, a comparison of language use over

Issues in Applied Linguistics
ISSN 1050-4273
© 2002, Regents of the University of California
Vol. 13 No. 1, 41-69

these different registers could expand our knowledge of how discourse markers vary in different conversational contexts. Therefore, this study aims to determine if there are differences in the use of discourse markers according to register and, if so, what situational characteristics may contribute to such differences.

LITERATURE REVIEW

Numerous studies have analyzed discourse markers in naturalistic conversations, examining their general linguistic and functional properties. An overview of previous research is provided below. The purpose of the literature review is to establish a collective definition of discourse markers in order to rationalize the selection of target discourse markers analyzed in the present study.

Linguistic Properties of Discourse Markers

Early work by Östman (1982) presents four basic linguistic features of discourse markers. Discourse markers are (a) short, (b) prosodically subordinate to another word, (c) independent from the content of the sentence, and (d) syntactically separate from the sentence. Hölker (1991) also states that discourse markers do not affect the truth conditions of an utterance, nor do they add anything to the propositional content of an utterance. Similarly, Fraser (1990, 1993, 1999) stresses that discourse markers do not affect the grammaticality of a sentence, nor do they create propositional meaning.

Brinton (1996) provides more detailed descriptions of the syntactic, phonological, lexical, and semantic characteristics of discourse markers. According to Brinton, the syntactic features of discourse markers include the following: (a) They are restricted to utterance-initial position, (b) they occur outside the syntactic structure or are loosely attached to it, and (c) they are optional. The phonological and lexical features include the following: (a) They are short and phonologically reduced, (b) they form a separate tone group, and (c) they are marginal forms and difficult to place within a traditional word class. Finally, semantically, discourse markers have little or no propositional meaning.

However, Brinton's (1996) formulation of discourse markers could be challenged. Although in Brinton's operationalization the position of discourse markers is restricted to utterance-initial position, several analyses show that discourse markers appear in various positions in utterances (Biber, Johansson, Leech, Conrad, & Finegan, 1999; Helt, 1997; Jucker & Smith, 1998; Schourup, 2001). Schourup (2001), for instance, argues that the occurrence of the marker *well* differs from that of discourse connectives (e.g., *moreover, after all*) because *well* can appear before any word in a sentence (e.g., *I was the most, well, experienced adventurer in the group*) and can highlight a single word or phrase, while discourse connectives do not have such distributional flexibility.

Similarly, Brinton's (1996) claim that discourse markers are optional may underestimate the importance of functional properties of discourse markers. It is

true that discourse markers do not affect the grammaticality of a sentence and could be removed from the sentence without changing its propositional meaning. However, pragmatically, discourse markers are often necessary devices in conveying speakers' attitudes and emotions. For instance, Bolinger (1989) provides examples of multiple occurrences of the marker *well* when commenting on someone's misbehavior (e.g., *Well, well, well!*), or of *well* used as a prompt to elicit another response (e.g., *Well?*). These examples clearly show that *well* fulfills the indispensable function of conveying attitudes in certain discourse contexts. Thus, a detailed analysis of the functional properties of discourse markers in context is necessary in order to determine the contextual features that relate to particular discourse markers.

Functional Properties of Discourse Markers

Researchers generally agree that discourse markers fulfill pragmatic functions: They promote cohesion between utterances and affect the degree of involvement among speakers. Discourse markers are also believed to signal the relation of an utterance to its immediate context (Redeker, 1990; Schourup, 1985). These pragmatic functions were highlighted by Schiffrin (1987), who emphasized the transitional function of discourse markers. Schiffrin defines discourse markers as "sequentially dependent elements which bracket units of talk" (p. 31). Based on naturally occurring interview talk, she analyzed the distribution of 11 discourse markers over five linguistic classes: lexicalized clauses (*y'know, I mean*), particles (*oh, well*), conjunctions (*and, but*), time deictics (*now, then*), and complements (*so, because*). She concluded that discourse markers provide coordinates to the discourse contexts in which participants produce and interpret meaning. The contextual coordinates integrate different components of talk and contribute to discourse coherence. For example, the marker *y'know* functions to invite a hearer to attend to specific information, while the marker *now* indicates a speaker's orientation toward an upcoming subtopic and signals the shift in the flow of discourse.

Hölker (1991) also claims that discourse markers are related to the speech situation, and have an expressive rather than a referential function. Fraser (1993) explicitly distinguishes between content meaning and pragmatic meaning. Content meaning is the literal meaning, while pragmatic meaning is the underlying intentions or attitudes that the speaker tries to convey through the literal message. Discourse markers belong to pragmatic meaning: They do not represent the propositional content of the sentence, but signal the speaker's belief towards, or evaluation of, the message. Stenstrom (1994, 1998) claims that discourse markers create discourse boundaries throughout spoken interaction.

Previous research has also examined how discourse markers influence a hearer's comprehension of a message. Discourse markers have been discussed in Relevance Theory (Blakemore, 1992; Sperber & Wilson, 1986; Wilson & Sperber, 1993), which emphasizes that understanding an utterance is a process of seeking relevance in what the speaker said. Efficient communication is achieved when a

great amount of information is conveyed with the least processing effort by the hearer. Discourse markers are claimed to help reduce the processing load: They are used to indicate the relevance of one discourse segment to another, signalling how the utterance should be interpreted and what should be expected in the upcoming discourse (Anderson, 1998; Blakemore, 1987; Blass, 1990; Jucker, 1993; Stenstrom, 1998).[3] For example, *like* informs the hearer that some kind of example or approximation follows (Anderson). Similarly, Jucker (p. 440) shows that *well* is a "signpost" that tells the hearer that the upcoming utterance is not optimally coherent with respect to the previous one. It indicates a shift in the conversational context, such as embarking on a new topic or mitigating some sort of confrontation. Similarly, Lenk (1998) demonstrates how the markers *anyway, however, still, incidentally, actually*, and *what else* signal the relevance of the utterance within the conversational context.

Summary of Discourse Marker Properties

To sum up, the literature claims that discourse markers are separate from the propositional content and syntactic structure of an utterance. They do not encode lexical meaning, but convey the pragmatic meaning of an utterance. Discourse markers function as interactive devices for speakers and hearers to help develop continuity and coherence in communication. They allow speakers to highlight important elements in a conversation, and convey their emotive and attitudinal stance toward the message. They also help hearers to follow speakers' trains of thought and assist in the interpretation of the utterance.

Despite general agreement on the functions and linguistic properties of discourse markers, there is great disagreement as to which items should be considered as discourse markers. There is no generally accepted list of discourse markers in English. Table 1 displays a list of items that have been identified as discourse markers in the literature.

PURPOSE OF THE STUDY

Although previous studies have provided a reasonably well-formulated analysis of the linguistic and functional properties of discourse markers, little research has compared their use in different conversational situations. Erman (2001), Jucker and Smith (1998), and Fuller (2003) are among the few studies that directly address the relationship between patterns of discourse markers and contextual characteristics. Using a corpus of British English conversations, Erman (2001) compared the functions of *you know* in teenage and adult talk. The study documented that *you know* in teenage talk serves a comprehension-securing function, the speaker making sure that the listener understands the specific references made (e.g., *You know that orthopedic doctor?*). However, in adult conversations *you know* occurs more as a text monitoring device, introducing a change of topic or allowing the speaker to stall for time when engaged in a self-repair, such as a word search. As

Table 1: Items Identified as Discourse Markers

Study	Items Identified as Discourse Markers
Östman (1982)	*well, like, kind of, sort of, you know, I mean, oh, now, but, and, uh*
Schourup (1985)	*well, like, kind of, you know, I mean, oh, now, ah, mind you, uh*
Schiffrin (1987)	*well, you know, I mean, and, but, or, so, because, now then, oh*
Fraser (1990)	*well, you see, now, but, so, ah, all right, anyway, OK, or, then*
Redeker (1990)	*well, you know, I mean, oh, but, so, ah, all right, OK, because, mind you*
Stenstrom (1994, 1998)	*well, you know, you see, I mean, oh, now, all right, anyway, yeah, OK, like, really*
Biber et al. (1999)	*well, right, now, I mean, you know, you see, look*

Erman (2001) concludes, the use of *you know* could be influenced by the type of discourse and subject matter being discussed rather than by the relationship between the speakers. Her conclusion implies the need for subsequent research that examines the occurrences of discourse markers over different discourse contexts.

Jucker and Smith (1998) compared tape-recorded conversations between pairs of strangers and pairs of friends. Two types of discourse markers were identified: reception markers, which signal a reaction to a message (e.g., *yeah. OK*), and presentation makers, which modify the information presented by the speaker (e.g., *you know, I mean*). The results showed that reception markers were more common between strangers than friends, reflecting differences in the nature of the interaction. Jucker and Smith's analysis is exclusively based on the dichotomized classification of reception and presentation markers. However, individual discourse markers may not necessarily fall within these two discrete categories; some discourse markers may serve both functions. Thus, additional research in this area could prove valuable in order to determine empirically whether discourse markers serve reception or presentation functions or both across different conversational registers.

The presentation and reception markers introduced in Jucker and Smith (1998) were further examined by Fuller (2003) in two speech contexts: semi-formal interviews and casual conversations. The study revealed that contextual characteristics such as the roles of speakers and the relationship between the interlocutors could

shape the distribution patterns of certain discourse markers. For instance, the markers *oh* and *well* occurred more frequently in casual conversations. Although *oh* was initially categorized as a presentation marker by Jucker and Smith, Fuller found that it also functioned as a reception marker during interviews when the interview participants were in the role of listening and responding. The marker *you know*, on the other hand, appeared more often in the interview data because it functioned to frame information and to enhance common ground between the interlocutors.

Since discourse markers are purported to assist the speaker and hearer in managing the flow of communication, situational characteristics such as interlocutor relationships or the purpose of communication could greatly affect their use. Thus, in order to develop better insights into the nature of discourse markers, there is a need to characterize their patterns across various spoken registers, expanding on the previous work cited above.

Register is defined as "a language variety viewed with respect to its context of use" (Biber & Finegan, 1994, p. 4). Ferguson (1994, p. 20) states that register is the language used in a communicative situation that recurs regularly in a society (in terms of participants, setting, and functions) and which tends to develop identifying characteristics over time. Thus, the core of register analysis is the exploration of the link between linguistic features and situational characteristics. Discourse markers, a class of discourse devices, are subject to such investigation because they may serve as pervasive indicators of register differences. Thus, the purpose of the present research is to investigate whether register variation exists in the use of discourse markers. The study aims to quantify the distributions of discourse markers across registers and to examine how their occurrences vary according to situational variables.

Among many potentially useful situations for analysis of discourse markers, this study focuses on three conversational registers: family conversations, professor-student office hour transactions, and service encounter conversations.[4] Biber (1994, p. 40) proposes a set of situational parameters that characterize registers, such as interlocutor relationship, setting, and communicative purposes, which are useful in characterizing these three registers. Family conversations usually take place in private settings, while service encounter conversations take place in public. The extent of shared knowledge and emotional involvement tends to be greater in family conversations than in service encounters. Family discourse usually involves expressions of personal feelings and attitudes, while service encounter conversations are designed for the transfer of information or goods. Professor-student office hour interactions, on the other hand, seem to stand between family conversations and service encounters because they take place in semi-public settings, but, similar to conversations among intimates, the interaction may often reflect a great degree of personal involvement between speakers. Given these situational differences, these three registers are a useful database for exploring the link between the relative distributions of discourse markers and contextual characteristics.

The present study has two purposes: to quantify discourse markers across registers and to explore the link between their distribution patterns and situational variables. The study is guided by the following two research questions:

(1) Do the frequency and use of discourse markers differ among the three conversational registers: family conversations, professor-student office hour transactions, and service encounter interaction?

(2) If so, to what extent can the differences be attributed to the situational characteristics of individual registers?

METHOD

Selection of Discourse Markers for Analysis

Operational Definitions of Discourse Markers

It is beyond the scope of this paper to analyze all the discourse markers identified in previous studies. In order to rationalize a selection of discourse markers for analysis, a principled set of criteria was established based on previous literature. Because there is no unified definition of discourse markers, the items that are considered as qualifying as discourse markers are numerous. Thus, the purpose of establishing the working definition is to specify the subgroup of discourse markers to be analyzed in the current study.

According to previous research, linguistic properties of discourse markers include the following: (a) they are inserts (i.e., single words, verbal formulae) which can be prosodically subordinated to another word and are syntactically independent from an utterance, and (b) they convey little lexical meaning. Functional properties include the following: (a) they convey pragmatic/expressive meaning and indicate the speaker's attitudes, and (b) they are interactional devices which contribute to the evolving progress of discourse continuity and discourse coherence.

Selection of Target Discourse Markers

In order to limit the number of target discourse markers, these operational definitions were applied to each of the discourse markers identified in the eight previous studies (see Table 1). Based on their linguistic properties, linking adverbial and coordinating conjunctions (i.e., *like, because, then, and, so, anyway, but*) were excluded from the analysis because they can be part of syntactic structures and signal connections or transitions between elements in a text (Biber et al., 1999). The marker *mind you* was also excluded because it does not occur frequently in American English. In addition, interjections such as *ah*, which indicate purely emotive moves (e.g., hesitation, surprise), were excluded from the analysis. As a result, 12 discourse markers were left for the analysis: *well, kind of (sort of), you know, I mean, OK, now, oh, right, look, you see, yeah,* and *really.*

Table 2: Corpora for this Study

Corpus Description	Number of texts	Number of words
Family conversations	12	54,694
Professor-student office hour transactions	11	50,412
Service encounter conversations	10	56,478

Data Collection

Three corpora of transcribed naturalistic conversations of American English were selected for the current study, representing each register: family, professor-student office hour transactions, and service encounter interaction (see Table 2).[5]

Family conversations were taken from the *Longman Spoken and Written English Corpus*.[6] Professor-student and service encounter interaction came from the *TOEFL 2000 Spoken and Written Academic Language Corpus*.[7]

Data Analysis

Monoconc software was used to sort data in this study. Monoconc is a publicly available concordancing program that allows the user to search for specific words in a corpus and provides lists of the occurrences of the words in context, along with frequency information. Each lexical item was entered as a search word in the corpora in order to generate a list of its occurrences in context. Then, each list was checked manually by the researcher in order to exclude the items which did not function as discourse markers.

The criteria for these judgements were based on the operational definitions of discourse markers established above. Although these definitions helped to identify the group of target discourse markers, more detailed criteria were needed for individual markers. The lexical items identified as discourse markers often occur in different linguistic functions, and some lexical items that function as discourse markers also have clear grammatical uses. Because these linguistic and grammatical functions are relatively easy to identify, this study established a set of linguistic functions that disqualified an occurrence of a particular item as a discourse marker. That is, when the items appeared in the linguistic contexts summarized in Table 3, they were not counted as discourse markers. The established criteria are mainly focused on grammatical functions, but for some markers, fixed expressions (see the case of *oh* in Table 3 for instance) were also excluded from the count. Although it is true that the expressions such as *oh my god* or *oh no* also serve pragmatic functions, indicating speaker attitudes toward a previous proposition, this study limited its analysis to discourse markers as single entities, when they are stand-alone units and do not combine with other inserts or interjections.

Table 3: Linguistic Properties that Disqualify a Token as a Discourse Marker

I mean *you know* *you see (see)*	Subject and predicate in a sentence (e.g., *You know* what?; *You see* my point? *I mean* it.)
look	Phrasal or prepositional verbs (e.g., *Look up* the vocabulary.)
well	Adverb (e.g., You sang really *well*.) Fixed expressions (e.g., Is it summer session as *well*?)
now	Adverb with temporal meaning (e.g., I'm doing my homework *now*.)
kind of/sort of	Partitive of quality (e.g., It is a new *kind of* soup.) Fixed expressions (e.g., It's a good sign poster for that *sort of* stuff.)
yeah/oh yeah *OK* *right/all right*	Response to a question, request, or order (e.g., A: you have a pen? B: *Yeah*.) Adverb of exactness (e.g., You'll initial *right* here.) Confirmation check when *right* or *OK* are transcribed with rising intonation (e.g., You just gave me $10, *right?*)
really/oh really	Adverb (e.g., I *really* like your sweater.) Confirmation check when *really* is transcribed with rising intonation (e.g., A: She missed class today. B: *Really?*)
oh	Fixed expressions (e.g., *Oh* my god.; *Oh* no.)

In order to confirm the accuracy of data sorting, the regrounding technique (Seliger & Shohamy, 1989) was used, and the original data were sorted twice. The frequency of each discourse marker was compared between the first and second sortings, showing a 98.5% agreement rate. In addition, when functional analysis was required, 20% of the data from each list was randomly selected and independently coded by a second rater who had experience in discourse analysis. The agreement rate between the two raters was 94.5%.

Finally, because the three corpora had different lengths, normalization was applied to make the frequency counts comparable with each other. The raw frequency count of each discourse marker was divided by the total number of words in the corpora and multiplied by 50,000. Thus, frequencies are reported as the occurrence of each marker per 50,000 words, since all corpora are approximately this long. Normed counts, rather than raw counts, are reported in this study because normed counts convert the number of occurrences of a particular discourse marker to a standard scale, informing us how often individual discourse markers are found in a fixed amount of text (See Biber et al., 1998, p. 263-264 for an explanation of normalization).

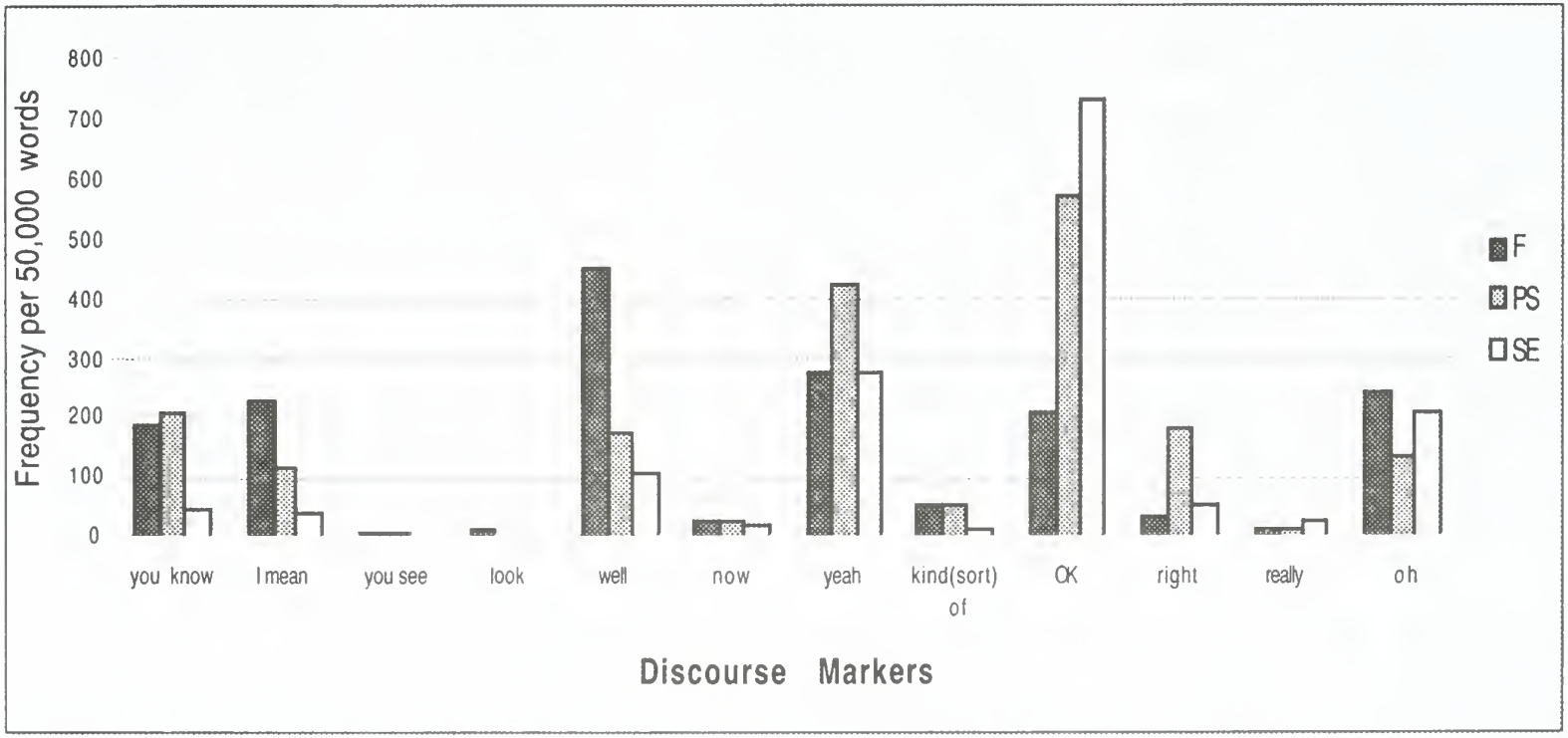
RESULTS AND DISCUSSION

Overall Distribution Patterns of Discourse Markers

Figure 1 reports the frequency distributions of the twelve discourse markers across discourse contexts. As seen in Figure 1, the markers *you know* and *I mean* are much more common in family and office hour conversations compared with service encounter situations. *Well* is twice as frequent in family interaction, compared with office hour and service encounter interactions. The markers *you see, look*, *now*, and *really* are not common in the corpora.[8] The marker *OK* is much more frequently used in office hour and service encounter interactions than in family conversations. *Yeah* and *oh* show similar distribution patterns across registers. *Right* appears most frequently in the office hour corpus.

The following section provides more detailed discussions of a selection of six discourse markers *I mean, you know, OK, right, yeah,* and *oh*, focusing on how register characteristics may interact with their distribution patterns. These six discourse markers were selected in part because they were relatively prevalent in the data. Originally the marker *well* was also considered for the analysis; however, due to the extensive and complex functional characteristics of *well* claimed in the previous literature, it was decided that this marker requires individual, separate analysis, possibly in a future study.[9]

Figure 1: Distribution of Individual Discourse Markers



Note: F= Family conversations; PS = professor-student office hour transactions; SE = service encounter conversations.

Analyses of Individual Discourse Markers

The Discourse Marker I mean

The marker *I mean* has the major function of signalling repair and consequently often prefaces an expansion or clarification of the speaker's prior utterance (Schiffrin, 1987, p. 295). Therefore, *I mean* contributes to conversational continuity by helping to orient the hearer to the upcoming message.

Excerpt (1) below, from the family corpus, illustrates the use of *I mean* to signal a repair.

(1) Family conversation (123701.txt)[10]

1		Mother:	That room, that we had had our church in has returned to what was originally its purpose which was to be an all purpose parish hall.
2		Son:	Yeah.
3		Mother:	It would be all right.
4		Son:	Yeah.
5	→	Mother:	We would still need a chapel of some sort. The idea is okay, **I mean**, we could do a lot of with space.
6	→	Son:	Oh, yeah, it's a good idea, it's a good idea. **I mean** it would bring, it would probably make us more visible.
7		Mother:	As well as increasing our facilities, the only thing is twenty thousand dollars is just barely giving the architect enough. Well, they won't kick in until there's a matching ten thousand. So if they have twenty thousand dollars that still isn't going to build the church. Do you think we should do it?
8		Son:	I don't know . . .

Using *I mean* at line 5, the mother clarifies and expands her prior utterance, explaining why "the idea [using the space as an all-purpose parish hall] is okay." In the subsequent turn, at line 6, the son also uses *I mean* as he clarifies why "it's a good idea" to use the parish hall. Excerpt (2) also illustrates the repair/clarification function of *I mean*.

(2) Family conversation (123701.txt)

1		Son:	Did he have Mexican hair or did he have American hair?
2	→	Mother:	What do you mean? He had black, black hair. **I mean**, he had black man's hair.
3	→	Son:	**I mean**, did he have, if it was different, some black people have Mexican hair, some black people have nappy hair. Nappy hair is like the nappy hair is like the real short curly stuff, you know, like Mondero has.
4		Mother:	Oh.
5		Son:	Like, and Mexican hair is like straighter. It's straight. It's like Jeff Jones.

At line 2, using *I mean*, the mother clarifies what she means by "black hair." The son begins his turn with *I mean* at line 3. The marker *I mean* here is also used to clarify meaning: The son uses it to clarify his initial question (line 1) after the mother's question in line 2 reveals that that was unclear or misunderstood.

Register characteristics can help us understand why the discourse marker *I mean* is more frequent in the family and professor-student corpora than in service encounters. Conversational exchanges in service encounter interaction tend to be short and to exhibit less personal involvement due to their business-oriented purpose involving the exchange of goods, money, and information. Conversely, family and professor-student office hour interaction often consist of more extended pieces of discourse, and participants tend to be engaged in conversation at a more personal level. Thus, it is possible that these two registers would require more instances of clarification or discourse repair.

The Discourse Marker you know

The major function of *you know* is to establish a mutual base of knowledge between the interlocutors and to promote cooperative interaction. *You know* asserts that there exists a shared orientation between the speaker and hearer (Schourup, 1985, p. 103). By using *you know*, the speaker is imparting information to the hearer, and at the same time appealing to the hearer to cooperate and to accept the information as shared knowledge (Biber et al., 1999, p. 1077; Erman, 1987, p. 169; Östman, 1981, p. 17; Schiffrin, 1987, p. 268). With *you know*, the speaker stresses the listener's role in conversation, drawing his or her attention to a piece of information.

The interpersonal functions of *you know* are also highlighted in Holmes's (1986, p. 7) two broad categories of *you know*: *you know* expressing speaker certainty and *you know* expressing speaker uncertainty. The former *you know* is used to reassure the listener about the validity of the information presented. In the latter function, the speaker expresses his or her uncertainty about the validity of the message or the listener's understanding of the message and thus seeks reassurance or agreement from the listener. Holmes distinguishes these two functions according to a variety of factors, such as intonation patterns, the speaker's relationship to the addressee, and the degree of shared knowledge.

Similarly, Schourup (1985, p. 136) states that *you know* in the middle of the turn signals that the speaker is actively engaged in conversation and is concerned with whether the speaker and hearer are on the same track. In extended discourse, *you know* helps sustain the listener's attention to the ongoing interaction. It assists the speaker in checking whether the hearer is still tuned in to what is being said.

In Excerpt (3) below, a professor is imparting his knowledge about Chinese history and is presenting it as knowledge shared by the student by using *you know*.

(3) Professor-student conversation (humhioh 068.txt)

1	→	Professor:	**You know**, the Chinese never went around trying to convert people to Confucianism. This, you know....
2		Student:	Yeah.
3	→	Professor:	**You know**, they were, again, inherently conservative.They just thought **you know** this is great, got everything set up, everybody else is screwed, **you know**. Leave them to their own devices. **You know** the Chinese are busy building a big wall, **you know**, to keep people out. Not trying to take other people over. And, oh **you know**, the Mongols would, **you know**, invade China.

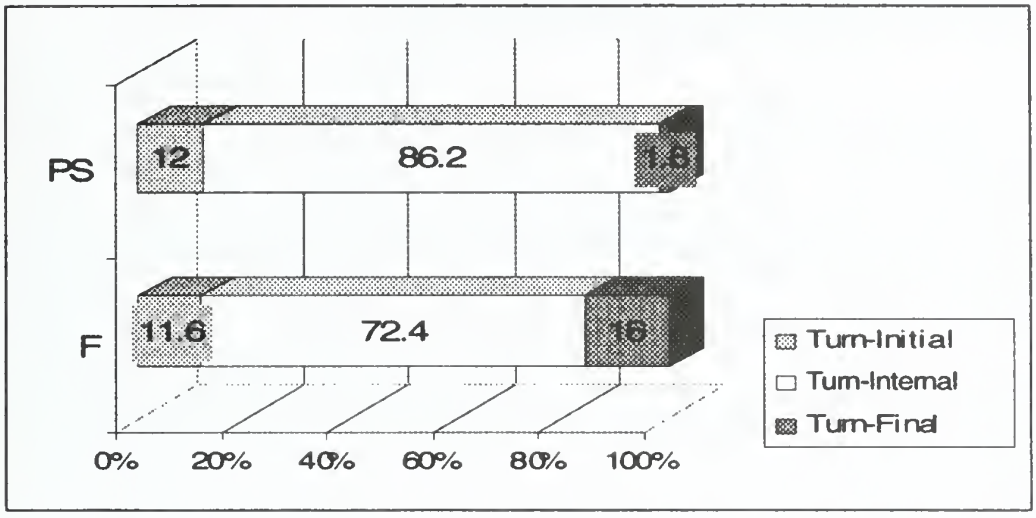
The frequency of *you know* in the conversation reflects the professor's tacit effort to get the student to accept the presented information and to integrate it as mutual knowledge. This effort is acknowledged by the student, as seen by the student's response, *yeah,* at line 2 following the professor's first utterance. With *yeah*, the student indicates that the information provided by the professor is being attended to and understood (see also the discussion of the discourse marker *yeah* below).

To sum up, *you know* is an indicator of striving to attain rapport and mutual understanding between the interlocutors (Östman, 1981). This function of *you know* becomes important in this analysis because the family and professor-student corpora have more incidences of *you know* than the service encounter corpus. Conversations among family members and between professors and students are usually topic-oriented and tend to involve longer stretches of discourse, often in the form of a narrative, with a greater degree of interlocutor involvement in the interaction. Such conversational features are quite different from those of service encounter situations, in which conversational exchanges tend to be short, business-transactional, and highly structured and routinalized. Due to these differences, *you know* is more common in the family and professor-student interaction, possibly because speakers use it to sustain the interest of listeners, to draw them into the topic of conversation, and to check their understanding.

Although the frequency of *you know* is similar in family and professor-student conversations, there seems to be some difference in terms of where in an utterance *you know* occurs. Figure 2 shows the percentages of *you know* identified either at the beginning, middle, or end of a speaker turn in the family and professor-student corpora. The percentages in the two registers are considered comparable because the raw counts of *you know* are very similar (frequencies of 190 and 210 in the family and professor-student corpus, respectively). The figure shows that *you know* occurs in turn final position nine times more frequently in family interactions than in professor-student conversations.[11]

According to Holmes (1986), turn-final *you know* indicates the speaker's confidence that the hearer knows the thing being referred to. Erman (2001) calls this function of *you know* a "comprehension-securing function"; the speaker marks

Figure 2: Distributions across turn positions: *you know*



Note: F= family conversations; PS = professor-student office hour transactions; SE = service encounter conversations.

termination of a turn with *you know*, and the hearer responds by showing acknowledgement. Turn-final *you know* attempts to prompt a response from listeners, appealing to their (presumed) shared knowledge, and urging them to accept information as given or known. Turn-final *you know* could thus be interpreted literally as "as you know," indicating more presumed certainty in the information provided (Östman, 1981, p. 21).

An example of turn-final *you know* in family conversation is given in Excerpt (4) below.

(4) Family conversation (122301.txt)

1	→	Wife:	I just, I thought he'd do a really bad job, but I think he was great. I was I was expecting him to do a worse, **you know**.
2		Husband:	Yeah, I guess.

After providing her personal opinion on someone's performance, the wife finishes her turn with the marker *you know*, which is followed by the husband's hedged *yeah* in line 2. Even though the husband does not seem to totally agree with his wife's statement, the wife's *you know* prompts an acknowledgement from him. Thus, the marker *you know* here urges the listener to accept the information as given or known (whether or not it actually is).

It is possible that because family members share more knowledge and experiences, speakers can more easily appeal to hearers' shared knowledge with *you know,* thus accounting for the higher frequency of this discourse marker in the family corpus. Speakers might end their turns with *you know* more often in family discourse because they are more confident that hearers, who are more likely to

share the speakers' current state of knowledge or feelings, will accept the information and understand the speakers' feelings or ideas.

The Discourse Marker OK

The major function of *OK* is signalling the reception of a message. According to Biber et al. (1999), *OK* usually serves as a compliant response to directives, suggestions, offers, and permission-giving. For example, a common speech act in professor-student interaction is professors giving advice to students about the classes they should enroll in or about the projects they are working on. Using *OK*, the student can signal to the professor that his or her advice is being attended to and accepted. This function of *OK* is illustrated in Excerpt (5), a student-professor conversation during a regular academic advising session. In each turn the student uses *OK* in order to acknowledge the professor's suggestions about choices of history classes:

(5) Professor-student conversation (humhioh 068.txt)

1		Professor:	History 47 is offered.
2	→	Student:	**OK.**
3		Professor:	But history 489 is offered.
4	→	Student:	**OK.**
5		Professor:	So, if you're interested in taking something with me.
6	→	Student:	**OK.**
7		Professor:	that would be something you could do.
8	→	Student:	**OK.**

Similar to professor-student conversations, in service encounters, *OK* is also used as a response form signalling that the message has been accepted and understood. This feedback signal is frequently observed as a response to a request for service or information, which is a common speech act in this register. For instance, in Excerpt (6), the server's *OK* at lines 3, 5, and 7 functions as a response to the customer's requests or instructions regarding the type and number of photocopies to be made. The customer also uses *OK* at lines 6 and 10, showing that the server's information is understood and accepted.

(6) Service encounter conversation (en130)

1		Server:	Hi.
2		Customer:	Hi. Um...I need four copies of each of these.
3	→	Server:	**OK.**
4		Customer:	Double sided.
5	→	Server:	Double sided. **OK**. This one . . I might have a hard time double siding because, um, unless you cut the fringes off, otherwise the, machine'll eat it.
6	→	Customer:	Oh, **OK**. Well I can cut [3 syllables]. It doesn't have to be double sided.
7	→	Server:	Yeah. **OK**. And this-just four copies?

8		Customer:	Yeah.
9		Server:	I'm not going to be able to do these double sided cos the machine is, [2 syllables] paper, [4 syllables].
10	→	Customer:	**OK**.
11		Server:	Is that OK?
12		Customer:	Yep.

This example illustrates that *OK* assists in the interactive achievement of discourse purposes, in this case, the completion of copy-making service. Frequent use of *OK* contributes to the efficient and prompt completion of business and information exchange in service encounters.

Another major function of *OK* discussed in the previous literature is *OK* as a transition marker (Beach, 1993; Condon, 2001; Merritt, 1984). In this function, *OK* signals the reception of a message, but at the same time it indicates a transition in the discourse and marks a shift in the direction of conversation. Merritt observed this dual function of *OK* in service encounter interaction, emphasizing that service encounters are routinized conversational acts in which the transition to each successive stage is highly expected. According to Merritt, service encounters typically consist of five phases: access, selection, decision, exchange, and closure. *OK* characterizes the routine nature of service encounters and functions as a bridge or linking device between phases of an encounter. Condon elaborates further on this use of *OK* and considers *OK* as a "default verification device" which signals a transition in discourse. The default verification device becomes most salient when linkings of utterances appear in routines (Condon, p. 497). Condon found in her analysis that *OK* frequently initiates routines of decision making and is followed by fixed expressions such as suggestions or requests. This default use of *OK* could be more common in service encounter discourse because these interactions usually consist of a series of structured and routinalized exchanges.

In the present analyses of the service encounter corpus, *OK* is also common during closing routines in the interaction. It often co-occurs with *thank you* and functions as a transition marker when the service transaction is nearly completed and the customer is ready to leave, as shown in the use of *OK* at line 4 of Excerpt (7) below:

(7) Service encounter conversation (en210)

1		Server:	. . . would you put your student I. D. on there?
2		Customer:	It's already on there.
3		Server:	**OK**, great thanks. [types on keyboard] [printer sounds]
4	→	Customer:	**OK**, thank you.
5		Server:	You have a good day.
6		Customer:	You too.

This transition function of *OK* may be more common in service encounter discourse because service encounters consist of a series of expected, structured,

and routinized exchanges. Service encounter interactions are often goal or task-oriented and aimed at the prompt completion of particular actions, namely exchanges of goods, information, and money. The marker *OK* seems to facilitate the completion of these tasks by marking transitions between the stages involved in service encounter interaction, such as opening, bidding, transaction of goods, and closing. As Beach (1993) claimed, *OK* displays an understanding of the prior utterance, but at the same time it indicates readiness for a transition in the discourse. *OK* serves to set up a new turn and topic, and thus facilitates the current speaker's actions. This function of *OK* thus seems important in service encounter interaction that includes numerous discourse transitions and exchanges.

In the present analysis, *OK* appears least frequently in the family corpus. This could be because family members have less need for explicitly signalling a reaction to a message because of their shared background and knowledge. As shown in Jucker and Smith (1998), the more distant the interlocutor relationship is (e.g., strangers), the greater the need for the use of backchannel signals such as *OK* in order to assure that the communication is progressing well and that the information has been accepted and integrated into the listener's current state of knowledge.

Another possible reason for the different distribution of *OK* is the different purposes of interaction across these three registers. As noted above, *OK* often signals the reception of a message and completes a conversational sequence. While family interaction is less likely to be focused on exchanging new information, service encounter and professor-student conversations include numerous short information exchanges and frequent turn exchanges, consequently providing greater opportunities for *OK* to occur in the reception function.

The Discourse Marker right

The marker *right* resembles the use of *OK* in that it indicates understanding, compliance, and agreement with the previous remark (Biber et al., 1999; Jucker & Smith, 1998). This function of *right* can be seen in the following professor-student conversation:

(8) Professor-student conversation (busatoh088.txt)

1		Student:	well what I kind a did is jump down to here thinking Ok I know the, the uh, the figure under row zero under my slide variable X seven
2	→	Professor:	**Right.**
3		Student:	is gonna be that number so I just started
4		Professor:	[many unclear syllables] right.
5		Student:	doing the simple
6	→	Professor:	**Right right**, so
7		Student:	Instead of realizing that if everything else is zero it cuts them out, a lot of your work [few unclear syllables].

8	Professor:	Yeah, that's, and that's what they're after on this. It's that you're,recognize that if if exponents five are the basic variables then you can figure out what the values are.
9 →	Student:	**Right.**
10	Professor:	Without going through anything else.
11	Student:	OK.
12	Professor:	Because in effect everything else can be dropped out of the problem.
13 →	Student:	**Right.**

In Excerpt (8), the professor uses *right* at lines 2 and 6, indicating that the student's solution to a mathematical problem is accurate and adequate. The student also uses *right* at lines 9 and 13, indicating that the professor's explanation is adequate and thus understood.

Right is approximately six times more frequent in the professor-student corpus than in family and service encounter interaction. These results may be caused by the fact that professors and students use *right* not only as a confirmation marker when information is attended to and accepted, but also as an evaluative response when some answers or solutions are presented regarding the problem under discussion. Thus, in the professor-student register, the marker *right* seems to have a literal use and may be interpreted as "You're right," confirming the degree of accuracy or truth value of the information presented. This function of *right* could be more pervasive in professor-student interaction because conversations in this register are often oriented toward the exchanges of academic information or facts gleaned from textbooks or class lectures. These exchanges might invite the use of more evaluative confirmation markers, such as *right*, as well as the simple reception markers (e.g., *OK, yeah*).

The Discourse Marker yeah

Jucker and Smith (1998) suggest that the major function of *yeah* is a backchannel cue or reception marker, marking a response to a statement. Similarly Schegloff (1981) notes that vocalizations such as *uh huh* or *yeah* are signals of attention, interest, or understanding on the listener's part and assist in the smooth continuation of a conversation. Schegloff considers *yeah* as an indication of passing on the opportunity to take a turn, which therefore encourages the speaker to continue. *Yeah* is thus a signal that a conversation is accomplished collaboratively between participants. *Yeah* as a reception marker is a useful device to show that communication is in progress, and new information is being integrated into the ongoing interaction.

In Excerpt (9) from a service encounter interaction, by responding with *yeah* at lines 5 and 7, the customer is signalling to the server that the information is being understood and accepted.

(9) Service encounter conversation (n210)

1		Customer:	I just need this hold like released. I just paid for this but I went overthere and they said.
2		Server:	It didn't, it didn't release?
3		Customer:	Um I'm sure if that one's I think that one was done, but this one's not done yet and they sent me back over here because I just paid this parking ticket but it wasn't released yet I guess.
4		Server:	[types on keyboard] it should have been let's take a look here.
5	→	Customer:	**Yeah**.
6		Server:	Unfortunately I'm not the one to release it [unclear word]
7	→	Customer:	**Yeah**.
8		Server:	Oh no student line which is this one. OK. [unclear word] What I'm going to ask you do
9		Customer:	Uh huh
10		Server:	Is step down to Student Accounts on the left hand counter.
11		Customer:	OK.
12		Server:	And give them this paper and then they'll take the hold off.
13		Customer:	OK.

However, the present analyses reveal an additional function of *yeah*. The marker *yeah* also occurs as a discourse link or connective in which the speaker indicates acknowledgement of information with *yeah*, but also uses *yeah* as a take-off for further talk. The speaker does not simply use *yeah* to acknowledge what the other has said, but, following *yeah*, continues with an elaboration of the ongoing topic. This function of *yeah* is different from the reception function of *yeah* noted by Jucker and Smith (1998). In this function, *yeah* seems to correspond with Jucker and Smith's definition of *presentation markers*, namely markers used to introduce new information. *Yeah* is used both as a confirmation signal, responding to the information presented, and at the same time it is followed by further expansion of conversation, introducing more information or comments into the ongoing discourse. Thus, *yeah* functions to present or introduce further discourse development.

In the following professor-student conversation, the professor uses *yeah* at line 5 in order to confirm the student's understanding of the naval conference while subsequently expanding upon the student's explanation of the role of the naval conference during World War II.

(10) Professor-student conversation (humhioh_n71.txt)

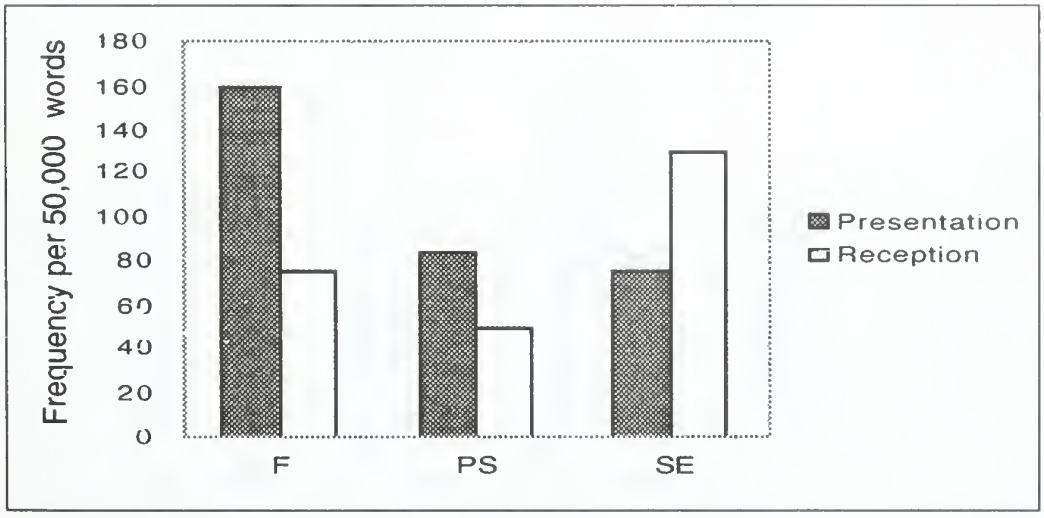
1	Professor:	You're confusing the League of Nations with the Washington Naval Conference.
2	Student:	OK, which one were we talking about
3	Professor:	Yesterday we were talking about the Washington Naval

4		Student:	Naval conference, that's the one where we were just getting rid of.
5	→	Professor:	**Yeah**, it was basically because during the World War II, there had been this great escalation
6	→	Student:	**Yeah**.
7		Professor:	Of, of arms of navy ships and all this stuff and so because we didn't join the League of Nations we were which Wilson's League of Nations was gonna be all encompassing and take care of all these things.
8		Student:	OK.
9		Professor:	And since we didn't join that then we had to go piece by piece and start thinking well what's important to us remember what I did I call a twenties foreign policy.

Thus, in the above excerpt *yeah* as a presentation marker in line 5 links the professor's utterance with what the student has said previously, in effect continuing and adding to the student's explanation. In contrast, the student's use of *yeah* at line 6 serves a reception marker function: It demonstrates the student's acceptance of the professor's information, but is not followed by additional talk by the student. In fact, the professor's talk is still in progress, as we see at line 7 where he continues the turn-in-progress explaining about the naval conference.

Figure 3 displays the frequencies of *yeah* serving the reception and presentation functions in the three registers. *Yeah* is coded as a simple reception marker when it appears as a single response or as an expression to show the reception of information alone (e.g., *Yeah, good*; *Yeah, I know*; *Yeah, thanks*), and passes on the opportunity to take a turn. When *yeah* does not comprise a turn in itself but is

Figure 3: Frequency of *yeah* functioning as a presentation and reception marker



Note: F= family conversations; PS = professor-student office hour transactions; SE = service encounter conversations.

followed by an elaboration of the ongoing topic or takeoff for further talk (i.e., *Yeah* followed by the further expansion of the discourse as in Excerpt [10], line 5) it is coded as a presentation function.[12]

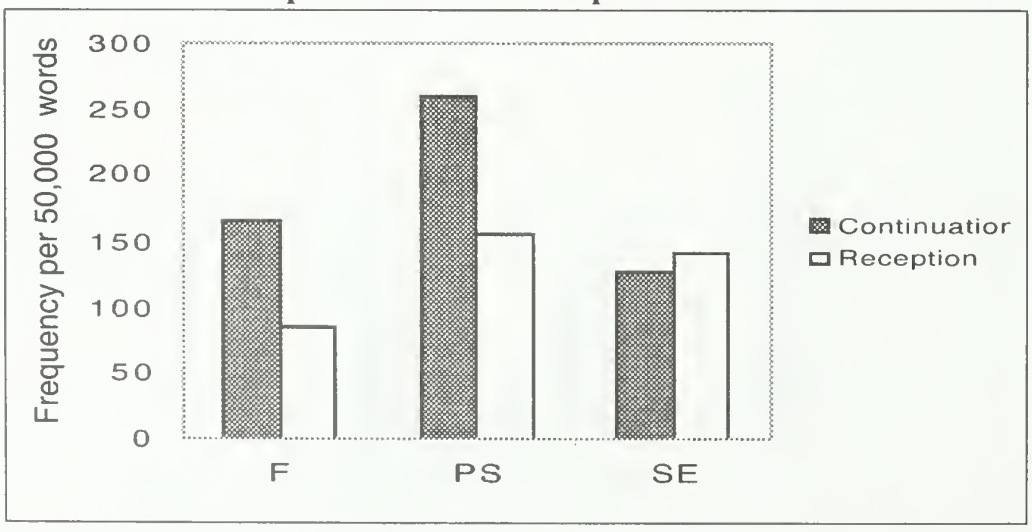
As shown in Figure 3, in the family and professor-student corpora, *yeah* occurs almost twice as often as a presentation marker, while in the service encounter conversations, the frequency is similar between the two functions. Because of the high personal involvement in the discussions and opinions exchanged in family and professor-student discourse, speakers seem to use *yeah* as a presentation signal. Speakers use *yeah* as a confirmation of what the other has said, but at the same time as a strategic device to take over the floor and develop communication further. This presentation use of *yeah* may be less common in service encounter situations, in which conversations are brief and are oriented toward specific goals.

The Discourse Marker oh

The reception use of *oh* is well noted in the previous literature (Biber et al., 1999; Heritage, 1984; Jucker & Smith, 1998; Schiffrin, 1987). According to Heritage, *oh* signals that the speaker's knowledge base is undergoing a state of change and is not continuative. Similarly, Schiffrin's analysis reveals that *oh* is a recognition display marker; it displays the speaker's reception of unanticipated or familiar information.

As with the marker *yeah*, in the present analysis *oh* is identified with a dual function: as a marker indicating the reception of information (reception marker) and as a marker shifting toward new information (presentation marker). Both functions of *oh* may encode the change of state claimed by Heritage (1984). The following example illustrates the reception function of *oh*.

Figure 4: Frequency of *oh* functioning as a presentation and reception marker



Note: F= family conversations; PS = professor-student office hour transactions; SE = service encounter conversations.

(11) Service encounter conversation (en115)

1		Server:	. . . Forty eight cents.
2		Customer:	What is it?
3		Server:	Forty-eight cents.
4	→	Customer:	**Oh.** That's with this though too.
5	→	Server:	**Oh.**
6		Customer:	I need to buy that.
7		Server:	OK. . . . . A dollar twenty eight . . .seventy two cents.

At line 4 the customer uses *oh* to respond to the server's information about the price. Here *oh* expresses the change of state of knowledge in the customer regarding the price. In the same line, the customer informs the server that he would like to purchase an additional item, which is not included in the forty-eight cents. The customer's utterance is followed by the server's *oh*, which similarly indicates a change in the state of knowledge of the server.

In addition to the reception function, *oh* also occurs as a presentation marker, serving as a "take-off signal" and introducing comments or new information. In this sense, *oh* marks a transition of the discourse content. *Oh* as a presentation marker is observed at the point of topic transition, introducing new discourse content, as in Excerpt (12) below:

(12) Family conversation (123701.txt)

1		Mother:	Your cross-country shirt, this one.
2		Son:	Yeah.
3		Mother:	Here, I don't want her to lie on it.
4		Son:	You can throw it on my bed.
5	→	Mother:	Here's what I'll do. You can put it on whatever hangar you want, but I'm going to put it right here. Okay and then you can do . . . okay. **Oh**, yeah, this was lying outside on my way to the car, so I thought well if I do have the day off Monday
6		Son:	Uh huh.
7		Mother:	That's what I'm going to do.

In the middle of the turn in line 5 there is a transition of conversation topic. *Oh* indicates the mother's change of knowledge state because she has just remembered that something was lying outside. At the same time, *oh* marks a topic change from the cross-country shirt to the thing she found outside.

Figure 4 displays the frequency of the discourse marker *oh* identified as a reception marker (i.e., acknowledgement of information) or as a presentation marker (i.e., introduction of a comment or a new topic) in each register. *Oh* is coded as a reception marker when it appears as a single response form or as an expression showing the reception of information (e.g., *Oh good*; *Oh interesting*), while it is coded as a presentation marker when followed by new information or comments, as in Excerpt (12).[13]

Figure 4 shows that although the overall frequencies of *oh* are similar in family, professor-student, and service encounter discourse, *oh* is approximately twice as common in its function as a reception marker in the service encounter register. Conversely, in family and professor-student conversations, more occurrences of *oh* function as presentation markers. These results could reflect the differences in the purposes and subject matter being discussed in the three registers. Family and professor-student registers are similar because they include many separate extended pieces of discourse that may require multiple shifts in topic. For instance, common topics of family conversation include describing various incidents that occurred either inside or outside the home, providing opinions or beliefs on various matters, telling stories or jokes, or recalling past events. Thus, *oh* plays an important role as a presentation marker in marking the introduction of new topics. The presentation function of *oh* is less salient in the service encounter corpus perhaps because in this register conversations tend to consist of short exchanges that are oriented toward specific goals rather than extended narratives or detailed descriptions of events. A service encounter interaction is likely to be focused on only one topic, rather than having shifts in topic.

SUMMARY AND CONCLUSIONS

The present study has aimed to investigate register variation in the use of discourse markers. The study fulfills two purposes: to compare the frequencies of discourse markers among three conversational registers and to explore the link between certain functions of discourse markers and characteristics of these communicative situations in order to pursue an explanation of their use.

The findings clearly indicate a relationship between distribution patterns of discourse markers and situational characteristics of individual registers. Discourse markers such as *you know* and *I mean* are more prominent in the family and professor-student corpora, perhaps reflecting the length of discourse and high personal involvement in conversation topics. Some discourse markers (e.g., *OK*) are much more common in service encounter situations, reflecting the focus on information exchange.

Another important finding is that some discourse markers fulfill different functions even within individual registers, supporting the multi-functionality claimed in the literature. The present findings also show clearly that some discourse markers serve both as presentation and reception markers. *Oh*, for example, is used both to display recognition of information by the listener and to introduce new information by the speaker. Similarly, *yeah* signals the reception of a message, but also serves as a presentation marker, behaving as a continuation latch: Following *yeah*, speakers continue with further elaboration of the ongoing topic. This multi-functionality of individual markers interacts with register characteristics. *Yeah* as a presentation signal is more common in the professor-student corpus than in service encounters, where interactions are short, business-oriented, and

routinized in nature. Family members may use *oh* more as a presentation marker than as a reception marker because the nature of these interactions, often involving multiple topics, provides more occasions for signalling topic shifts.

Based on the limitations observed in this study, several implications for further research are suggested:

1. Due to the large amount of data involved, the current analysis excluded certain categories of discourse markers, such as interjections, linking adverbials, or coordinating conjunctions. Analysis of the marker *well* was omitted for the same reason. Future research should address these items in order to provide a more unified analysis of discourse markers. In addition, in order to confirm the relationship between discourse markers and particular discourse contexts suggested in this study, more detailed analyses of the functions of discourse markers over different conversational situations are needed.
2. The target registers for comparison should be expanded. It would be particularly interesting to analyze presentation markers in registers whose speech is predominantly unidirectional, such as in lectures. Such studies would provide insight into whether discourse markers are a distinctive characteristic of spontaneous interaction, that is, whether multiparty communication is a prerequisite for the occurrence of discourse markers.
3. Discourse markers are linguistically simple, but functionally complex. They have many different uses, and there is a range of contexts in which they can occur. Therefore, an inquiry into the first and second language acquisition of discourse markers could add to our understanding of how children and second language learners expand their functional repertoire of linguistic forms in the process of language development.
4. This analysis did not attempt to correlate intonation contours with the content and context of the conversation. How the intended function of the discourse maker changes with voice inflection might prove to be of further research interest.

In summary, an insight gleaned from the current quantitative and qualitative analyses is that discourse markers display different profiles according to register. Each discourse marker serves particular functions, and certain discourse markers are more salient in particular contexts than others, reflecting the characteristics of the specific conversational context, such as communicative goals, settings, and interlocutor relationships. These features interactively make up the register. That is, what determines the patterns and functions of discourse markers is not only the relationship between the participants, but also the nature or purpose of the discourse, all of which contribute to the characterization of certain registers and consequently affect the distribution patterns of discourse markers. Discourse markers are a type of linguistic device which could serve as significant indicators of register variation.

ACKNOWLEDGMENTS

I am very grateful to Professor Doug Biber at Northern Arizona University for his constructive feedback, guidance, and encouragement throughout this research. Thanks also go to three anonymous reviewers and the editors of *Issues in Applied Linguistics* for their helpful suggestions and comments on earlier drafts. I am solely responsible for the errors that may remain.

NOTES

[1] An earlier version of this paper was presented at the annual meeting of the American Association of Applied Linguistics in St. Louis in February, 2001.

[2] Previous studies have used a variety of terms to refer to the conversational devices that serve pragmatic functions (Jucker & Ziv, 1998): discourse marker (Brinton, 1990; Schiffrin, 1987; Stenstrom, 1998), pragmatic marker (Fraser, 1993), pragmatic particle (Östman,1981), discourse particle (Schourup, 1985), discourse connector (Blakemore, 1987), and pragmatic expression (Erman, 1987).

[3] Flowerdew and Tauroza (1995) demonstrated that a lecture with discourse markers was better comprehended than the same lecture without discourse markers. Fox Tree and Schrock (1999) also showed that the presence of discourse markers can reduce the hearer's processing effort and assist comprehension. The subjects in this study recognized the upcoming speech faster when they heard *oh* than when it was excised.

[4] Biber (1994, p. 51) notes that there is little consensus among researchers as to what counts as register, beyond the general association of register with situation variation. Following Biber (1994, p. 32), the present research uses *register* as a general cover term for all language varieties associated with different situations and purposes.

[5] The 12 texts of family discourse were from three families living in Ohio, California, and Oklahoma. All conversations took place among father, mother, and siblings in the morning when getting ready for the day. Professors in the 11 professor-student office hour texts were from four different disciplines: business, humanities, natural science, and social science. The 11 service encounter conversations were recorded in five different locations in a university: copy shop, library, bookstore, information technology service, and student business office.

[6] See Biber et al., 1999, for descriptions of the Longman corpus.

[7] See Biber, Reppen, Clark, and Walter (2001) and Biber, Conrad, Reppen, Byrd, and Helt (2002) for descriptions of the TOEFL 2000 Spoken and Written Academic English Corpus. The TOEFL 2000 Spoken and Written Academic English Corpus was sponsored by the Educational Testing Service.

[8] *Really* often occurs with other grammatical functions such as an intensifier (e.g., *It's really funny*).

[9] Some of the functional characteristics of *well* in the previous literature include marking insufficiency, mitigating face-threats, floor-holding, introducing new topics, introducing reported speech, building discourse coherence, prompting a response, and marking surprise (see, e.g., Bolinger, 1989; Fuller, 2003; Jucker, 1993; Schiffrin, 1987; Schourup, 2001).

[10] Interlocutors in each conversation are specified according to the speaker numbers in the transcriptions. Unintelligible sounds are transcribed either as [syllables] or [unclear word].

[11] Turn-initial position can be preceded by an interjection such as *ah, um*, or *uh*.

[12] The classification of the reception and presentation functions of *yeah* was confirmed by two raters for 20% of randomly extracted data, as explained in the methodology section. The agreement rate was 92%.

[13] The inter-rater agreement on the classification of reception and presentation functions of *oh* was 90% for 20% of randomly extracted data.

REFERENCES

Andersen, G. (1998). The pragmatic marker *like* from a relevance-theoretic perspective. In A. H. Jucker & Y. Ziv (Eds.), *Discourse markers: Descriptions and theory* (pp. 147-170). Amsterdam: John Benjamins.

Beach, W. (1993). Transitional regularities for 'casual' "okay" usages. *Journal of Pragmatics, 19*, 325-352.

Biber, D. (1994). An analytical framework for register studies. In D. Biber & E. Finegan (Eds.), *Sociolinguistic perspectives on register* (pp. 31-56). Oxford: Oxford University Press.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use.* Cambridge: Cambridge University Press.

Biber, D., Conrad, S., Reppen, R., Byrd, P., & Helt, M. (2001). Speaking and writing in the university: A multidimensional comparison. *TESOL Quarterly, 36*, 9-48.

Biber, D., & Finegan, E. (1994). Introduction. In D. Biber & E. Finegan (Eds.), *Sociolinguistic perspectives on register* (pp. 3-12). Oxford: Oxford University Press.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English.* Harlow, England: Pearson Education Limited.

Biber, D., Reppen, R., Clark, V., & Walter, J. (2001). Representing spoken language in university settings: The design and construction of the spoken component of the T2K-SWAL Corpus. In R. Simpson & J. Swales (Eds.), *Corpora in linguistics and language teaching* (pp. 48-57). Ann Arbor: University of Michigan Press.

Blakemore, D. (1987). *Semantic constraints on relevance.* Oxford: Blackwell.

Blakemore, D. (1992). *Understanding utterances.* Oxford: Blackwell.

Blass, B. (1990). *Relevance relations in discourse.* Cambridge: Cambridge University Press.

Bolinger, D. (1989). *Intonation and its uses: Melody and grammar in discourse.* Stanford, CA: Stanford University Press.

Brinton, L. J. (1990). The development of discourse markers in English. In J. Fisiak (Ed.), *Historical linguistics and philology* (pp. 45-71). Berlin: Mouton de Gruyter.

Brinton, L. J. (1996). *Pragmatic markers in English: Grammaticalization and discourse functions.* Berlin: Mouton de Gruyter.

Condon, S. L. (2001). Discourse *ok* revisited: Default organization in verbal interaction. *Journal of Pragmatics, 33*, 491-513.

Erman, B. (1987). *Pragmatic expressions in English: A study of* you know, you see, *and* I mean *in face-to-face conversation.* Stockholm: Almqvist and Wiksell.

Erman, B. (2001). Pragmatic markers revisited with a focus on *you know* in adult and adolescent talk. *Journal of Pragmatics, 33*, 1337-1359.

Ferguson, C. A. (1994). Dialect, register, and genre: Working assumptions about conventionalization. In D. Biber & E. Finegan (Eds.), *Sociolinguistic perspectives on register* (pp. 15-30). Oxford: Oxford University Press.

Flowerdew, J., & Tauroza, S. (1995). The effect of discourse markers on second language lecture comprehension. *Studies in Second Language Acquisition, 17*, 455-458.

Fox Tree, J. E., & Schrock, J. C. (1999). Discourse markers in spontaneous speech: Oh what a difference an *oh* makes. *Journal of Memory and Language, 40*, 280-295.

Fraser, B. (1990). An approach to discourse markers. *Journal of Pragmatics, 14*, 383-395.

Fraser, B. (1993). Discourse markers across language. *Pragmatics and Language Learning Monograph Series, 4*, 1-17.

Fraser, B. (1999). What are discourse markers? *Journal of Pragmatics, 31*, 931-952.
Fuller, J. M. (2003). The influence of speaker roles on discourse marker use. *Journal of Pragmatics, 35*, 24-45.
Helt, M. E. (1997). *Discourse marker and stance adverbial variation in spoken American English: A corpus-based analysis.* Unpublished doctoral dissertation, Northern Arizona University, Arizona.
Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 299-345). Cambridge: Cambridge University Press.
Hölker, K. (1991). Franzosisch: Partikelforschung. *Lexikon der Romanistischen Linguistik, 5*, 77-88.
Holmes, J. (1986). Functions of *you know* in women's and men's speech. *Language and Society, 15*, 1-22.
Jucker, A. H. (1993). The discourse marker *well*: A relevance-theoretical account. *Journal of Pragmatics, 19,* 435-452.
Jucker, A. H., & Ziv, Y. (1998). Discourse markers: Introduction. In A. H. Jucker & Y. Ziv (Eds.), *Discourse markers: Descriptions and theory* (pp. 1-12). Amsterdam: John Benjamins.
Jucker, A. H., & Smith, S. W. (1998). And people just you know like 'wow': Discourse markers as negotiating strategies. In A. H. Jucker & Y. Ziv (Eds.), *Discourse markers: Descriptions and theory* (pp. 171-201). Amsterdam: John Benjamins.
Leech, G. (2000). Grammars of spoken English: New outcomes of corpus-oriented research. *Language Learning, 50*, 675-724.
Lenk, U. (1998). *Making discourse coherence.* Tubingen: Gunter Narr Verlag Tubingen.
Merritt, M. (1984). On the use of 'okay' in service encounters. In J. Baugh & J. Sherzer (Eds.), *Language in use* (pp. 139-147). Englewood Cliffs, NJ: Prentice-Hall.
Östman, J-O. (1981). You know*: A discourse functional approach.* Amsterdam: John Benjamins.
Östman, J-O. (1982). The symbolic relationship between pragmatic particles and impromptu speech. In N. Enkvist (Ed.), *Impromptu speech: A symposium* (pp. 145-177). Åbo, Finland: Åbo Akademi.
Redeker, G. (1990). Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics, 14*, 367-381.
Schegloff, E. (1981). Discourse as an interactional achievement: Some uses of 'uh-huh' and other things that come between sentences. In D. Tannen (Ed.), *Analyzing discourse: Text and talk. Georgetown university round table on language and linguistics.* Washington DC: Georgetown University Press.
Schiffrin, D. (1987). *Discourse markers.* Cambridge: Cambridge University Press.
Schourup, L. (1985). *Common discourse particles in English conversation*: Like, well, y'know. New York: Garland.
Schourup, L. (2001). Rethinking *well. Journal of Pragmatics, 33*, 1025-1060.
Seliger, H. W., & Shohamy, E. (1989). *Second language research methods.* Oxford: Oxford University Press.
Sperber, D., & Wilson, D. (1986). *Relevance.* Cambridge: Cambridge University Press.
Stenstrom, A. B. (1994). *An introduction to spoken interaction.* London: Longman.
Stenstrom, A. B. (1998). From sentence to discourse: *Cos* (because) in teenage talk. In A. H. Jucker & Y. Ziv (eds.), *Discourse Markers: Descriptions and theory* (pp. 171-201). Amsterdam: John Benjamins.
Wilson, D., & Sperber, D. (1993). Linguistic form and relevance. *Lingua, 90*, 1-25.

Naoko Taguchi holds a Masters of Education Degree in Teaching English to Speakers of Other Languages from State University of New York at Buffalo. She is currently teaching at the ESL Department of Minnesota State University-Akita, Japan, and at the same time pursuing her Ph.D. in Applied Linguistics at Northern Arizona University. Her research interests include discourse analysis, interlanguage pragmatics, and learning strategies.