Running head: DISCOURSE MARKERS ACROSS SPEAKERS AND SETTINGS

Discourse Markers across Speakers and Settings

Jean E. Fox Tree

University of California Santa Cruz

Abstract

Viewing discourse markers as one of a class of signals that communicators use to manage conversation can help make sense of the varied and contradictory functions ascribed to them. Issues for further testing include demonstrating that speakers and listeners ascribe particular functions to particular markers, tracking the development of these functions as children age, tracking how these functions are adopted by second language learners, and documenting how discourse marker use changes across technological settings.

Discourse Markers across Speakers and Settings

Spontaneous communication differs dramatically from prepared. With carefully prepared talk or writing, there are no unplanned stops, no off-the-cuff re-framings, and no rambling verbiage. Discourse markers take center stage in spontaneous communication, with a much more limited role in prepared communication. In this review, I will cover (1) the history of discourse markers, and the thorny issue of defining what they are, (2) a theoretical approach to discourse markers that views them as one of a class of signals that communicators use to manage conversation, including a review of evidence supporting these signaling properties, (3) a review of what is known about the acquisition of discourse markers by children and by second language learners, and (4) a discussion of how discourse markers might change across different communicative media. By viewing discourse markers as a conventionalized system, hypotheses can be made about how markers might vary across settings, including formal versus informal, public versus private, and face-to-face versus computer-mediated.

Historical Overview

A fundamental characteristic of discourse markers is that they operate beyond the propositional content of the communication. But there is widespread disagreement about what constitutes a discourse marker and what discourse markers do.

In the research literature, discussion of discourse markers can be found under a variety of names, including *pragmatic expressions* (Erman, 1987), *pragmatic markers* (Andersen, 1998; Redeker, 1990), *pragmatic devices* (Stubbe & Holmes, 1995), *pragmatic particles* (Holmes, 1990), *discourse particles* (Schourup, 1985; Aijmer, 1988), *discourse connectives* (Schiffrin, 1987), *phatic connectives* (Bazzanella, 1990), *ritualized*

speech (Lalljee & Cook, 1975), and *interjections* (James, 1972; Wilkins, 1992), although researchers have centered on the term *discourse markers* in recent years (Jucker & Ziv, 1998).

Although the study of related phenomena such as silent pauses and *um*s and *uh*s can be traced back to at least the 1950's (Goldman-Eisler, 1968; Maclay & Osgood, 1959; Mahl, 1956), it was in the 1980's that a proliferation of work on discourse markers hit the stage, including Schourup (1985), Schiffrin (1987), and Erman (1987). The researchers varied widely in which expressions, exactly, qualified as discourse markers. Schourup (1985) focused on *like*, *well*, and *you know*, but also discussed *oh*, *now*, *I mean*, *mind you*, *and everything*, *sort of*, and *kind of*, among others. Schiffrin (1987) focused on *oh*, *well*, *and*, *but*, *or*, *so*, *because*, *now*, *then*, *you know*, and *I mean*. Erman (1987) highlighted *you know*, *you see*, and *I mean*. The discussion of what counts as a discourse marker is still not resolved. *Well* for example is an obvious discourse marker to some (Fox Tree, 1999; Hellerman & Vergun, 2007) but not included for others (Wilkins, 1992).

Part of the confusion undoubtedly rests in the fact that discourse markers seemingly do so many things. The function of discourse markers can be conceptualized as a long list of varying roles, including contributing to local coherence of adjacent phrases, assisting in turn-taking or repair, or contributing to social solidarity (for review see Fox Tree & Schrock, 2002 or Jucker & Ziv, 1998). This approach can lead to functions that are widely disparate and even contradictory. For example, discourse markers can been seen as turn-initiators (Fung & Carter, 2007) or turn-relinquishers (Duncan, 1972). They can be used more among more familiar (Jucker & Smith, 1998;

Redeker, 1990) or less familiar (Östman, 1981) people. They can be used in anxious

(Lalljee & Cook, 1975) and casual (Broen & Siegal, 1972) environments. They can be

seen as grammatically superfluous (Redeker, 1991) or playing some grammatical role

(Schiffrin, 1987). They are sometimes equated with hedges or fillers (Aijmer, 1984; Fung

& Carter, 2007), and at other times kept distinct from hedges and fillers (Fox Tree, 1999).

They are sometimes seen as idiosyncratic products of a particular group of speakers

(Lalljee & Cook, 1975), and at other times as quite universal (Hellerman & Vergun,

2007).

       Observations like these led researchers to argue that discourse markers are

interchangeable or meaningless, with reference to either markers as a whole (O'Donnell

& Todd, 1991; Russel, Perkins, Grinnell, 2008) or to a subset of markers (for example,

Redeker, 1991, discusses similarities between *oh* and *well* and Stenström, 1990b,

discusses similarities between *well* and *um*). At the same time, other researchers have

argued that the same discourse marker can be used differently by different people; for

example, that men use *you know* to express uncertainty but that women use it to express

confidence (Holmes, 1990).

       Nonetheless, there is some evidence that people discriminate among discourse

makers. A content analysis of folk definitions of the meanings of three frequently

conflated inserts, *um* and *uh*, *like*, and *you know*, indicated that people did hold intuitive

understandings of the meanings of these words that differentiated them from each other

(Fox Tree, 2007). Kindergarten-aged children from different language communities who

spoke several different languages used discourse markers differently depending on

whether they were pretending to be a low-status or high-status character (E. Andersen, 2000).

While there is widespread disagreement about what constitutes a discourse marker and what discourse markers do (Jucker, 1993; Fuller, 2003; Wilkins, 1992), researchers do agree that discourse markers focus on the way communication is negotiated rather than on its content. This negotiation process goes by many names including *metalanguage* (Maschler, 1994), *procedural meaning* (Blakemore, 2002), *negotiating strategies* (Jucker & Smith, 1998), and *collateral signals* (Clark, 1996). These functional approaches to spontaneous communication take as a starting point the concept that people communicate on multiple levels. In a two-level form, one level conveys particular words and a second level conveys how those words should be interpreted. For example, speakers can say *oh* to indicate to addressees that the immediately following words are not meant to be connected to the immediately preceding words, which would be the usual state of affairs in the absence of an *oh* (Fox Tree & Schrock, 1999).

Conversation Management

Discourse markers are only one way of commenting on a primary message. They can be conceptualized as part of the category of *inserts* (Biber, Johansson, Leech, Conrad, & Finegan, 1999; Clark, 2004). Inserts include *um*s and *uh*s and discourse markers (words like *well*, *oh*, *you know*, and *I mean*). Communicators can also comment on primary messages via *juxtapositions* (providing information by the way talk is presented; for example, the speaker indicates a change by abutting "he said that" with "he asked if"), *modifications* (changes in the production of speech such as prolonging syllables), and *concomitants* (other information conveyed at the same time as speech,

such as facial expressions and manual gestures; Clark, 2004). Collectively, these are called *collateral signals*.

Collateral signals of different types can convey similar things. For example, the insert *um* can forewarn upcoming delay in speech (Clark & Fox Tree, 2002; Fox Tree, 2001) and the modification of prolonging syllables can also forewarn an upcoming delay (Fox Tree & Clark, 1997). The concomitant of gaze can be used in turn-taking (Clarke & Argyle, 1982; Duncan, 1974) as can a variety of inserts such as *you know* (Erman 2001; Schegloff, 1987) and *well* (Stenström, 1990b). The juxtaposition of a restart can convey that information in the discourse record should be altered (Fox Tree, 1995) as can the use of the insert *I mean* (Schiffrin, 1987). The modification of a prolonged high terminal rise can direct attention forward when listening to talk (Tomlinson & Fox Tree, 2009); forward focus has also been proposed for the insert *now* (Aijmer, 1988).

To date, the bulk of existing research on discourse markers has focused on production, generally adopting the method of analyzing speech corpora for regularities from which functional hypotheses can be drawn. This approach lends itself to the proposing of a proliferation of meanings for any particular discourse marker, to the overlapping of meanings, and indeed to the meaninglessness of meanings – with so many meanings, how is it possible to know which applies in a particular situation? One solution is to treat discourse markers as similar to other words, with a conventionalized component that shifts in manifestation depending on context. This conventionalized component has been described as a *core* meaning (Jucker, 1993), *underlying* meaning (Fox Tree & Schrock, 1999), *abstract* meaning (Östman, 1995), *generic* meaning (Heritage, 1984, 1998), or *basic* meaning (Clark & Fox Tree, 2002; Fox Tree & Schrock,

2002).

One example of how this solution has been used is with respect to the inserts *um* and *uh*. These fillers have been studied in much greater depth than any other inserts, including all discourse markers. Like discourse markers, research on *um*s and *uh*s is rooted in production, and the proposed meanings of *um* and *uh* have varied widely (for overview, see Fox Tree, 2000). But the medusa of meanings can be tamed with a single conventionalized use that shifts in apparent meanings depending on the context. The myriad uses of *um* can be distilled to the basic meaning of indicating upcoming delay (Clark & Fox Tree, 2002). All other uses can be seen as interpretations built on this base. For example, *um*s and *uh*s, in conjunction with pauses, made defendants in mock trials appear more guilty (Hosman & Wright, 1987). But this does not mean that *um* means that a speaker is lying. Instead, *um* suggests an upcoming delay, and in the context of answering a question from a prosecutor, the delay may be hypothesized to be because the speaker is lying. As another example, hearing *um* in picture descriptions primed listeners to expect difficult-to-name objects, but not when the speaker was thought to have trouble naming objects (Arnold, Hudson Kam, & Tannenhaus, 2007). For the non-object-agnosic, the delay forewarned by *um* was interpreted to be caused by naming difficulty for particular, difficult-to-name objects. For the object-agnosic, the delay forewarned by *um* was interpreted to be caused by generalized naming difficulty, and therefore not useful to a listener for determining whether the object being described was new or old.

Comprehension studies of the basic meanings of discourse markers are sparse. Information is available for *oh*, *well*, *like*, *you know,* and *and*. The basic meaning of *oh* has been proposed to be indicating an upcoming change of state (Heritage, 1984), such as

when information that was forgotten is suddenly remembered and inserted into the ongoing dialogue (the change of state is from not knowing to knowing). In support of this proposal, hearing *oh* assisted people in making sense of what was heard (Fox Tree & Schrock, 1999). Listeners were faster at recognizing words in a speech stream after an *oh* than when the *oh* was edited out. This demonstrated that *oh* helped listeners integrate upcoming talk, either by informing listeners to expect upcoming change-of-state information or by informing listeners to create a mental block between what was said before and after the *oh*s (Fox Tree & Schrock, 1999).

Two proposals for the basic meaning of *well* are that it indicates that the seemingly most relevant interpretation is not quite right (Jucker, 1993) and that it indicates that a seemingly irrelevant interpretation is actually relevant (Blakemore, 2002). In support of these proposals, reading *well* changed the way people interpreted replies (Holtgraves, 2000). Listeners judged replies as face-threatening more quickly when they were prefaced by *well*. This demonstrated that *well* indicated to listeners that the turns were dispreferred (Holtgraves, 2000), which can be seen as another way of suggesting to the listener that a less-obvious interpretation of the upcoming information was warranted.

The basic meaning of *like* has been proposed to be that it indicates upcoming loose use of language (Andersen, 1998), a deliberate marker of vagueness (Jucker, Smith, & Lüdge, 2003). In support of this proposal, hearing *like* affected how people retold stories (Fox Tree, 2006). *Like* was recycled in a similar or exact same location about 20% of the time across two tellings of the same spontaneous story to different addressees. About 10% of the time, *like*s were recycled across speakers (listeners retelling the teller's story used the teller's *like*s in similar locations). This demonstrated that the seemingly

amorphous *like* played a nonrandom role in story telling (Fox Tree, 2006). At the same time, loose language is dispreferred in some circumstances where precision is valued, such as interviews; predictably, overuse of *like* reduced job applicant's chances of success (Russel, Perkins, & Grinnell, 2008). An alternative proposal for *like* is that it functions to focus attention on upcoming information (Underhill, 1988), but this proposal excludes manifestations of *like* that other researchers include (Fox Tree, 2006; Jucker, Smith, & Lüdge, 2003; Romaine & Lange, 1991) and thus may be considered too restrictive.

The basic meaning of *you know* has been proposed to be an invitation to the addressee to more fully specify the speaker's intentions (Jucker & Smith, 1998). In support of this proposal, a study of spontaneous definitions of *you know* found that 77% of respondents described *you know* as being used by speakers to ensure addressee comprehension (Fox Tree, 2007).

Finally, the non-conjunctive, discourse marker use of the word *and* has been proposed to link sections of talk (Schiffrin, 1987). In support of this proposal, *and* was enough to make a false start behave more like a *middle false start*, which reliably slows comprehension in both English and Dutch, than like a *beginning false start*, which does not slow comprehension in Dutch and tends to not slow comprehension in English when utterances prefaced by discourse markers are controlled (Fox Tree, 1995). An example of a beginning false start is "aren't you on- isn't the government paying for yours?" An example of a middle false start is "she asked if they- how long they lived." An example of a near-beginning false start introduced by *and* is "and uh their- the beginning- the first sentence is…." Marker-free false starts that occurred at the beginning of utterances had

less processing cost than false starts that occurred in the middle of utterances.

Acquisition

Like other words, the meaning of discourse markers has to be learned, both in first and second language acquisition. With respect to children, the overall pattern emerging is that children use a restricted range of discourse markers with a narrower range of functions. For example, two- to three-year-olds use discourse markers in creating local connections between utterances; they use *because*, *and*, *but*, and *well* to mark the answer to a question, to mark upcoming contradictions, and to connect turns (Sprott, 1992). As they age, children use the same markers plus *so* to accomplish more with their talk, including marking upcoming justifications and topic shifts (Sprott, 1992). This movement from the local to the general has also been observed with respect to the acquisition of *ah* in Spanish. The basic meaning of *ah* is "'Look at X' or 'I see X'" (Montes, 1999, p. 1301), and it is this meaning that the child studied used when about age 2. When about age 3, the child was able to use *ah* for other purposes, such as to refute what was said (Montes, 1999). As another example, three-year-old German children used *hm* and *na* in more restricted ways than adult German speakers (Meng & Schrabback, 1999). Similarly, by the age of six children used different discourse markers for high status and low status puppets, a distinction not made by most four year olds (E. Andersen, 2000). The picture of expanding range of use with maturity is echoed by frequency: adults used *look*, *listen*, *you know*, *I mean*, *well*, and *you see* four times as often as children did (Romero Trillo, 2002).

This progression of children's learning of markers has also been observed with the similar collateral signals, *uh* and *um*. At three and four years of age, children did not

distinguish between *uh* and *um* (Hudson Kam & Edwards, 2008), but they did by age 5 or

6, producing shorter pauses after *uh* than after *um* (Van der Wege & Ragatz, 2004).

Indeed, a study of the Turkish signals *şey* (*um*), *yani* (*I mean*), and *işte* (*you know*)

provided evidence that children were still learning how to use discourse markers in adult-

like ways after they were nine years old (Furman & Özyürek, 2007). Adults and children

produced a similar rate of narratives with *şey*, but adults used *yani* more often than

children, and adults and nine-year-olds used *işte* more often than younger children.

Although the rates of *şey* usage were similar, the location was not; unlike adults, children

used it at the beginning of narratives (Furman & Özyürek, 2007).

Second language acquisition of discourse markers can look similar to children's

acquisition, with discourse marker use in second languages increasing with proficiency

and acculturation (Fung & Carter, 2007; Hellerman & Vergun, 2007). Hong Kong

Chinese speakers of English as a second language used English discourse markers less

frequently than native English speakers (Fung & Carter, 2007), preferring markers that

focused on the talk at hand, such as *and*, *but*, *because* and *I think*, over markers that

served interpersonal functions, such as referring to attitudes and mutual knowledge (Fung

& Carter, 2007).

But second language acquisition differs from first language acquisition in that

speakers already have a set of collateral signals that they use in their primary languages.

These signals may have counterparts in the two languages, such as *um*s and *uh*s in

English and *um*s and *uh*s in Dutch, which have been shown to have similar effects on

language comprehension (Fox Tree, 2001). But they may also not have easy translations;

consider the German signal *na* which is "an interjection to bridge different phases of

verbal interaction" (Meng & Schrabback, 1999, p. 1266). Particular discourse markers may be more or less common in one or another language, or follow different conventions. *Diciamo*, "let us say," is common in Italian, but not in English, although it is readily translatable (Waltereit, 2002). And there are other ways cross-linguistic discourse marker use varies, such as in frequencies of co-occurrence with other markers. For example, French-speaking children use more strings of discourse markers than English-speaking or Spanish-speaking children (Andersen, Brizuela, DuPuy, & Gonnerman, 1999).

The presence of discourse markers in the primary language creates translation issues monolingual children do not have. Furthermore, there is some suggestion that discourse marker translation works in a system separate from propositional content translation. Bilinguals can use discourse markers from their primary language in their secondary language and from their secondary in their primary (Matras, 2000). They can use markers with addressees who do not use them, such as a native German speaker using *na* in his English to English speaking addressees (Matras, 2000). This suggests that markers serve production purposes. But they can also correct their choice of discourse marker, replacing a marker from one language with one from another, demonstrating consideration of addressee comprehension (Matras, 2000).

Perhaps because of these translation issues, discourse markers are not typically taught to second language learners (Fung & Carter, 2007; Hellermann & Vergun, 2007). This is particularly problematic because learning discourse markers in a second language can be difficult (Romero Trillo, 2002). Spanish speaking children learning English opted for the marker *listen* when *look* would have been more appropriate because *oye* (*listen*) is more common in Spanish (data are from Spanish spoken in Spain; Romero Trillo, 2002).

Spanish speaking adults learning English used *well* as frequently as an adverb ("how well do you know him?") as as a discourse marker, whereas native speakers used *well* 85% of the time as a discourse marker; in addition, Spanish speaking adults rarely used *I mean* or *you know* (Romero Trillo, 2002). Discourse marker acquisition among Turkish-English bilinguals living in the U.S. for three to four years was idiosyncratic (Demirci & Kleiner, 1997). This could be because speakers choose to use discourse markers from one language or another based on their cognitive load at that moment (Matras, 2000).

Despite not being taught, the acquisition of discourse markers can be important for communicative success. Hearing discourse markers as a group improved lecture learning of material presented in English to a group of native Chinese speakers (Flowerdew & Tauroza, 1995; see also Tyler 1992). Discourse markers can also be used as a means of gaining the cultural clout of a more valued language. French discourse markers replaced Shaba Swahili markers with similar meanings, even in utterances where the main message was expressed in Shaba Swahili (de Rooij, 2000).

Not much is known about how discourse markers change in both primary and secondary languages when a new language is learned. One hypothesis is that markers that serve a function in a primary language but do not have a second language equivalent will be carried over into a speaker's second language. For example, *diciamo* ("let us say" in Italian) is used to give the subsequent utterance the appearance of being joint, indirectly requesting confirmation of this stance (Waltereit, 2002). This suggests that Italian speakers learning English would feel an absence of the function of *diciamo*, and therefore be likely to use the expression "let us say" in English more often than a native speaker would. But this prediction is not born out by studies of German-English bilinguals over

time. By the third generation, English markers replaced German markers in German-American speech, even when the German markers did not exist in English (Salmons, 1990). For example, the German marker of contrast *doch* was not preserved in the German of German-English speakers (Salmons, 1990).

Another hypothesis is that markers are inherently language-specific, and thus a new system of marking needs to be learned with each new language. For example, Hebrew has been proposed to have two versions of *I mean* (Maschler, 1994). The literal translation has a negative connotation associated with not having stated things clearly the first time, but the other version does not: Literally translated "as to say," the Hebrew discourse marker *klomar* suggests that the words themselves need expansion. The marker davka, "a marker introducing a side comment contrary to the immediately preceding discourse," does not even have an approximate translation into English (Mashler, 1994, p. 338). As another example, second language learners with less contact with native speakers use markers differently from those with more contact. German speakers with less contact used *well* four times as often to begin unqualified answers to questions, used *you know* more often before an explanation, and used *you know* less often before an implication than those with more contact (Müller, 2005). The type of English German speakers encountered also influenced discourse marker use. German speakers with American contact used more *like*s than those with British contact (Müller, 2005).

Observations like these fit well with the proposal that discourse marking systems are learned separately from the regular words and grammar of a language (Goss & Salmons, 2000). Like words, discourse markers may also be borrowed from one language to another. One proposal is that the borrowing of discourse markers is related to

pragmatic detachability, which is the extent to which a marker has lexical content and is tied to propositional information (Fuller, 2001). Those that are more detachable, such as *well* ("not easily analyzed in terms of lexical meaning"), will be borrowed sooner than those that are less detachable, such as *you know* ("highly lexical," Fuller, 2001, p. 355). This can lead to a situation where some discourse markers have replaced others (such as *well* replacing *naja* in Pennsylvania German), but others are used interchangeably (*but* and *aber*, Fuller, 2001; see Boas & Weilbacher, 2007, 2009 for contrasting data).

Yet another approach to how discourse markers change in language contact situations is to focus on the functions of the markers. A function that is accomplished by one kind of collateral signal in one language may be accomplished by a different kind of collateral signal in another language. In indicating upcoming delay, English-weighted bilingual speakers prefer to use *um*s and *uh*s, but Spanish-weighted bilingual speakers prefer to use prolongation (Dunn & Fox Tree, 2009). It remains to be seen whether highly fluent speakers of both languages will adopts *um*s and *uh*s in their English but prolongations in their Spanish. That is, whether speakers treat collateral signals as systems that need to be translated, similar to the propositional content of talk, or whether they use whatever system is most accessible or useful. For a Spanish-English bilingual, prolongation would be more useful as it would work for both languages; *um*s and *uh*s are markedly English.

Indeed, there is evidence that it is the function of the discourse marker that determines its relative frequency of use, even across speakers, settings, and languages. *You know* and its translation *znaš* are used similarly in the Croatian of Croatian-Australian bilinguals, but *like* is used much less frequently than the Croatian equivalent

kao which has a wider range of use than *like* (Hlavac, 2006). Similarly, English *like* may be used more often by teenagers because they lack the linguistic skills and confidence to speak directly (G. Andersen, 2000). As another example, *You know* and *I mean* should be infrequent when inviting addressee inferences is undesireable, such as when presenting oneself as having fully formed ideas in advance (Fox Tree & Schrock, 2002).

<div align="center">Novel Settings</div>

The same functions that make discourse markers frequent and useful in spoken dialogue make them useful in non-spoken spontaneous settings. In both, discourse markers comment on a primary message with the main goal of achieving *grounding* (Clark, 1996). Grounding is the process by which communicators check for understanding to make sure that what they have communicated and understood is sufficient for and appropriate to the current purposes of the communication. Grounding is essential for all forms of communication. Voice recognition software that responds immediately to a command can be frustrating (for example, a GPS system that responds to the command *go home* by turning on the fan). Many systems now check that the proper command has been received, with messages such as "I heard you say 'go home.' Is that correct?" This models the way that people routinely communicate, although people have many more ways of checking for understanding. Discourse markers are one of the principal ways to achieve grounding in the domain of spontaneous, unrehearsed communication.

The way people achieve grounding is influenced by various *grounding constraints* (Clark & Brennan, 1991). For example, when people can see and hear each other (the constraints of *visibility* and *audibilitiy*), communicators can make use of a variety of cues

to ensure that they have been understood, such as head nods, quizzical expressions, or the words "uh huh." Grounding constraints vary across communicative settings. Face-to-face talk is both visible and audible, but telephone talk is only audible. This means that any information ordinarily provided through visual information needs to be conveyed in some other way. Each communicative medium requires learning how to accommodate to the constraints of that medium. Two year olds who do not understand the constraints of a telephone do incongruous things, such as responding to the question "How old are you?" by holding up two fingers, or saying "look at the boo-boo I got today!" while pointing at the injury.

Spontaneous communication is particularly affected by the constraints of reviewability and reviseability. Because communicators cannot privately review or revise their communication, they need to create a discourse record that it easy to remember as well as a system for correcting errors publicly. Discourse markers can help with both these activities by sign-posting how talk should be interpreted; for example, *oh* can be used to indicate that upcoming information belongs earlier in the discourse record and *I mean* can be used to indicate that upcoming information corrects a preceding error (Fox Tree, 2000). In contrast, prepared talk should contain few if any discourse markers because neither reviewing nor revising is necessary. Predictably, inaugural speeches contain no *you know*s, *I mean*s, or *oh*s (Kowal, O'Connell, Forbush, Higgins, Clarke, & D'Anna, 1997).

But even among spontaneous talk there should be some with more and some with fewer discourse markers. One communicative variation is face to face versus monologue. In monologue spontaneous talk, there is no negotiating with addressees, which results in

the production of fewer *like*s, *oh*s, *well*s, *I mean*, and *you know*s (Fox Tree, 1999) as

well as other discourse markers (Stenström, 1990a). Furthermore, because there is no

feedback, speakers must come up with the hypotheses for how they are being understood

on their own. This should lead them to provide too much or too little information, and

this is the case: the range in the number of words needed to describe abstract shapes was

twice as large in monologues as opposed to dialogues, with monologues containing both

far fewer words and far more words (Fox Tree, 2000).

In the past, written communication was almost always carefully planned, and

spoken almost always not planned. Modern communicative technologies have blurred

this distinction, with some written forms of communication becoming increasingly

spontaneous. As written communication becomes more spontaneous, elements of

spontaneous talk useful for grounding should be adopted. For example, *oh* is useful in

speaking spontaneously to indicate when information has been remembered that was

forgotten earlier; this function would not be necessary in drafting a formal speech,

because the missing information would simply be inserted where it belonged. But in

writing spontaneously, with someone reading the typing as it is being written, the need

for *oh* resurfaces.

Some discourse markers, of course, have long had a presence in writing (*and*, *but*,

so), although the argument can be made that the way the word is pronounced may

distinguish between written conjunctive functions and spoken marker functions (Fox

Tree, 1995). But even traditionally spoken markers have been documented in written

sources. Historical approaches to the study of discourse markers have made use of

personal letters (Goss & Salmons, 2000) and medieval documents (Brinton, 1990).

Modern *you know* has been traced back to old English *hwoet* (Brinton, 1990). The recent proliferation of novel communicative technologies (email, instant messaging, texting) has brought forth a wealth of written discourse marker use. Written communication produced by these new technologies is much closer to speaking than earlier forms of writing. Spontaneous writing shares some aspects of traditional writing in that there is a public, accessible record. But it also has an effervescent quality similar to talking.

By considering how grounding constraints change across media, predictions can be made about how communication changes when moving from face-to-face talk to the written domain. With face-to-face, communicators can make use of a variety of cues to help them deliver timely messages in a format addressees can understand; they can look for head nods, listen for *uh huh*s, and infer understanding from the timing of turns (Clark & Brennan, 1991). Discourse markers can further help accommodate to the pressures of speaking in an unprepared, unrehearsed fashion (Fox Tree, 2000). Because not everything is going to be said in order, *oh*s can be used to indicate that upcoming information does not follow from prior information. When mistakes are made, *I mean*s can help fix them. Because traditional letter writing involves greater planning, there should be fewer occasions of out of order information or errors needing repair, and consequently fewer discourse markers. But this is not true of more recent forms of written communication.

Instant messaging can be thought of as spontaneous written dialogues. Correspondingly, discourse markers are common in instant messaging exchanges, and the more so the more expertise the communicators have with the domain (Fox Tree, Mayer, & Betts, 2009). This is because experienced users shift from treating instant messaging as a more formal communicative medium towards treating it like a conversation (Fox Tree,

Mayer, & Betts, 2009). More experienced users who treat instant messaging like a conversation care less about perfecting their information in advance, but then end up requiring more adjustments on the fly. They use less formal language and more discourse markers (Fox Tree, Mayer, & Betts, 2009). The discourse markers found in instant messaging overlap with those found in dialogue, including *I mean*, *you know*, *well*, *oh*, *I dunno*, and *like* (Fox Tree & Mayer, 2008; Fox Tree, Mayer, & Betts, 2009; Schourup, 1985). But the rates of use differ. In instant messaging, the mean rate of discourse markers was about 1%, upping to 2.5% with *yeah*s included as markers (Fox Tree, Mayer, & Betts, 2009). In conversations, the mean rate of discourse markers not including *I dunno* but including *yeah* was almost twice as high, 4.8% (Fuller, 2003). This difference can be seen as resulting from the relative costs of typing a discourse marker versus saying one, and from the minimal amount of private reviewing and revision time afforded by the text preparation window in instant messaging.

Summary

Discourse markers are conventionalized, learned expressions that provide information about how the propositional content of messages should be interpreted. They vary by individual, setting, and community. They further vary in how they are used across different languages, a comparison which is complicated by the fact that a discourse marker's function may be manifested as an insert in one language but as a concomitant in another. Discourse markers may even vary by the type of technology used for communication, although details on that dimension are only beginning to emerge. What discourse markers share is a central role in helping people achieve grounding in unplanned, unrehearsed communication.

Despite the fact that researchers do not agree on what counts as a discourse marker nor what functions discourse markers serve in language production and comprehension, the study of discourse markers has proliferated. Arenas of exploration include (1) tracking how children learn to use them, how second language learners acquire them, and how bilinguals treat them, (2) analyzing how their usage changed over time, (3) documenting the variety of implications discourse markers convey, (4) distilling implications to basic meanings, (5) testing understanding with both on-line and off-line comprehension tests, and (6) comparing production across settings, such as public versus private or natural versus computer-mediated. Because discourse markers can be seen through multiple lenses, their study will prove fruitful for years to come.

References

Aijmer, K. (1984). 'Sort of' and 'kind of' in English conversation. *Studia Linguistica, 38*, 118-128.

Aijmer, K. (1988). "Now may we have a word on this": The use of "now" as a discourse particle. In M. Kytö, O. Ihalainen & M. Rissanen (Eds.), *Corpus Linguistics, Hard and Soft: Proceedings of the 8th International Conference on English Language Research on Computerized Corpora* (pp. 15-34). Amsterdam: Rodopi.

Andersen, E. (2000). Exploring register knowledge: The value of "controlled improvisation". In L. Menn & N. Bernstein Ratner (Eds.), *Methods for Studying Language Production* (pp. 225-248). New Jersey: Lawrence Erlbaum.

Andersen, E., Brizuela, M., DuPuy, B., & Gonnerman, L. (1999). Cross-linguistic evidence for the early acquisition of discourse markers as register variables. *Journal of Pragmatics, 31*, 1339-1351.

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.

Andersen, G. (2000). The role of the pragmatic marker like in utterance interpretation. In G. Andersen & T. Fretheim (Eds.), *Pragmatic Markers and Propositional Attitude* (pp. 17-38). Philadelphia: John Benjamins Publishing Company.

Arnold, J., Hudson Kam, C. L., & Tanenhaus, M. K. (2007). If you say thee uh you are describing something hard: The on-line attribution of disfluency during reference comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(5), 914–930.

Bazzanella, C. (1990). Phatic connectives as interactional cues in contemporary spoken Italian. *Journal of Pragmatics, 14*, 629-647.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999.) *Longman Grammar of Spoken and Written English*. London: Longman.

Blakemore, D. (2002). *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers*. Cambridge: Cambridge University Press.

Boas, H. C. & Weilbacher, H. (2009). The unexpected survival of German discourse markers in Texas German. *Proceedings from the Annual meeting of the Chicago Linguistic Society, 42*(1), 1-15.

Boas, H. C. & Weilbacher, H. (2007). How universal is the pragmatic detachability scale? Evidence from Texas German discourse markers. In F. Hoyt, N. Seifert, A. Teodorescu, & J. White (Eds.), *Texas Linguistics Society 9: Morphosyntax of Underrepresented Languages* (pp. 33-58). Stanford: CSLI Publications.

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

Broen, P. A., & Siegel, G. M. (1972). Variations in normal speech disfluencies. *Language and Speech*, 5, 219-231.

Clark, H. H. (1996). *Using Language*. New York: Cambridge University Press.

Clark, H. H. (2004). Pragmatics of language performance. In L. R. Horn & G. Ward (Eds.), *Handbook of Pragmatics* (pp. 365-382). Oxford: Blackwell.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine & S. D. Teasley (Eds.), *Perspectives on Socially Shared Cognition*

(pp. 127-149). Washington, D. C.: American Psychological Association.

Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition, 84*, 73-111.

Clarke, D. D., & Argyle, M. (1982). Conversation sequences. In C. Fraser & K. R. Scherer (Eds.), *Advances in the Social Psychology of Language* (pp. 159-204). Cambridge: Cambridge University Press.

de Rooij, V. A. (2000). French discourse markers in Shaba Swahili conversations. The International Journal of Bilingualism, 4(4), 447-467.

Demirci, M., & Kleiner, B. (1997). Discourse markers in second language research. *Journal of Intensive English Studies, 11*, 131-142.

Duncan, S., Jr. (1972). Some signals and rules for taking speaking turns in conversation. *Journal of Personality and Social Psychology, 23*(2), 283-292.

Duncan, S., Jr. (1974). On the structure of speaker-auditor interaction during speaking turns. *Language in Society, 3*(2), 161-180.

Dunn, A. L., & Fox Tree, J. E. (2009). A quick, gradient Bilingual Dominance Scale. *Bilingualism: Language and Cognition, 12*(3), 273-289.

Erman, B. (1987). *Pragmatic Expressions in English : A Study of You Know, You See, and I Mean in Face-to-Face Conversation*. Stockholm, Sweden: Almqvist & Wiksell International.

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

Flowerdew, J., & Tauroza, S. (1995). The effect of discourse markers on second language lecture comprehension. *Studies in Second Language Acquisition, 17*, 435-458.

Fox Tree, J. E. (1995). The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language, 34*, 709-738.

Fox Tree, J. E. (1999). Listening in on monologues and dialogues. *Discourse Processes, 27*, 35-53.

Fox Tree, J. E. (2000). Coordinating spontaneous talk. In L. R. Wheeldon (Ed.), *Aspects of Language Production* (pp. 375-406). Philadelphia: Psychology Press.

Fox Tree, J. E. (2001). Listeners' uses of um and uh in speech comprehension. *Memory and Cognition, 29*(2), 320-326.

Fox Tree, J. E. (2006). Placing like in telling stories. *Discourse Studies, 8*(6), 723-743.

Fox Tree, J. E. (2007). Folk notions of um and uh, like, and you know. *Text & Talk, 27-3*(297-314).

Fox Tree, J. E., & Clark, H. H. (1997). Pronouncing "the" as "thee" to signal problems in speaking. *Cognition, 62*(2), 151-167.

Fox Tree, J. E., Mayer, S. A., & Betts, T. E. (2009, July 26-28). *At the crossroads of speaking and writing*. Paper presented at the Society for Text and Discourse, Rotterdam, the Netherlands.

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.

Fox Tree, J. E., & Schrock, J. C. (2002). Basic meanings of you know and I mean. *Journal of Pragmatics, 34*, 727-747.

Fuller, J. M. (2001). The Principle of Pragmatic Detachability in Borrowing: English-origin discourse markers in Pennsylvania German. *Linguistics, 39*(2), 351-369.

Fuller, J. M. (2003). The influence of speaker roles on discourse marker use. *Journal of Pragmatics, 35*, 23-45.

Fung, L., & Carter, R. (2007). Discourse markers and spoken English: Native and learner use in pedagogic settings. *Appplied Linguistics, 28*(3), 410-439.

Furman, R., & Özyürek, A. (2007). Development of interactional discourse markers: Insights from Turshish children's and adults' oral narratives. *Journal of Pragmatics, 39*, 1742-1757.

Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in Spontaneous Speech*. New York: Academic Press.

Goss, E. L., & Salmons, J. C. (2000). The evolution of a bilingual discourse marking system: Modal particles and English markers in German-American dialects. *International Journal of Bilingualism, 4*(4), 469-484.

Hellerman, J., & Vergun, A. (2007). Language which is not taught: The discourse marker use of beginning adult learners of English. *Journal of Pragmatics, 39*(157-179).

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.

Heritage, J. (1998). Oh-prefaced responses to inquiry. *Language in Society, 27*, 291-334.

Hlavac, J. (2006). Bilingual discourse markers: Evidence from Croatian-English code-switching. *Journal of Pragmatics, 38*, 1870-1900.

Holmes, J. (1990). Hedges and boosters in women's and men's speech. *Language and Communication, 10*(3), 185-205.

Holtgraves, T. (2000). Preference organization and reply comprehension. *Discourse Processes, 30*(2), 87-106.

Hosman, L. A., & Wright, J. W., II (1987). The effects of hedges and hesitations on impression formation in a simulated courtroom context. *Western Journal of Speech Communication, 51*, 173-188.

Hudson Kam, C. L., & Edwards, N. A. (2008). The use of uh and um by 3- and 4-year old native English-speaking children: Not quite right but not completely wrong. *First Language, 28*(3), 313-327.

James, D. (1972). Some aspects of the syntax and semantics of interjections. Papers from the eighth regional meeting of the Chicago Linguistic Society (pp. 162–172). Chicago, IL: Chicago Linguistic Institute.

Jucker, A. H. (1993). The discourse marker 'well': A relevance theoretical account. *Journal of Pragmatics, 19*(5), 435-452.

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. pp. 171-201). Amsterdam: John Benjamins.

Jucker, A. H., Smith, S. W., & Lüdge, T. (2003). Interactive aspects of vagueness in conversation. *Journal of Pragmatics, 35*, 1737-1769.

Jucker, A. H., & Ziv, Y. (1998). Discourse markers: Introduction. In A. H. Jucker & Y. Ziv (Eds.), *Discourse Markers: Descriptions and Theory* (1-12). Amsterdam: John Benjamins.

Kowal, S., O'Connell, D. C., Forbush, K., Higgins, M., Clarke, L., & D'Anna, K. (1997). Interplay of literacy and orality in inaugural rhetoric. *Journal of Psycholinguistic Research, 26*(1), 1-31.

Lalljee, M. G., & Cook, M. (1975). Anxiety and ritualized speech. *British Journal of Psychology, 66*(3), 299-306.

Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word, 75*, 19-44.

Mahl, G. F. (1956). Disturbances and silences in the patient's speech in psychotherapy. *The Journal of Abnormal and Social Psychology, 53*(1), 1-15.

Maschler, Y. (1994). Metalanguaging and discourse markers in bilingual conversation. *Language in Society, 23*, 325-366.

Matras, Y. (2000). Fusion and the cognitive basis for bilingual discourse markers. *International Journal of Bilingualism, 4*(4), 505– 528.

Meng, K., & Schrabback, S. (1999). Interjections in adult-child discourse: The cases of German HM and NA. *Journal of Pragmatics, 31*, 1263-1287.

Müller, S. (2005). *Discourse Markers in Native and Non-native English Discourse*. Philadelphia: John Benjamins.

Montes, R. G. (1999). The development of discourse markers in Spanish: Interjections. *Journal of Pragmatics, 31*, 1289-1319.

O'Donnell, W. R., & Todd, L. (1991). *Variety in Contemporary English* (2 ed.). New York: Harper Collins Academic.

Östman, J.-O. (1981). *You Know: A Discourse Functional Approach*. Amsterdam: John Benjamins.

Östman, J.-O. (1995). Pragmatic particles twenty years after. In B. Wårvik, S.-K. Tanskanen & R. Hiltunen (Eds.), *Organization in Discourse: Proceedings from the Turku Conference* (Vol. 14, pp. 95-108). Turku, Finland: Anglicana Turkuensia.

Redeker, G. (1990). Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics, 14*, 367-381.

Redeker, G. (1991). Linguistic markers of discourse structure. *Linguistics, 29*, 1139-1172.

Romaine, S., & Lange, D. (1991). The use of *like* as a marker of reported speech and thought: A case of grammaticalization in progress. *American Speech, 66*(3), 227-279.

Romero Trillo, J. (2002). The pragmatic fossilization of discourse markers in non-native speakers of English. *Journal of Pragmatics, 34*, 769-784.

Russell, B., Perkins, J., & Grinnell, H. (2008). Interviewees' overuse of the word "like" and hesitations: Effects on simulated hiring decisions. *Psychological Reports, 102*(111-118).

Salmons, J. (1990). Bilingual discourse marking: code switching, borrowing, and convergence in some German-American dialects. *Linguistics, 28*, 453-480.

Schegloff, E. A. (1987). Recycled turn beginnings: A precise repair mechanism in conversation's turn-taking organisation. In G. Button & J. R. E. Lee. (Eds.), *Talk and Social Organisation* (Vol. 1, pp. 70-85). Clevedon, England: Multilingual Matters, Ltd.

Schiffrin, D. (1987). *Discourse Markers*. New York: Cambridge University Press.

Schourup, L. C. (1985). *Common Discourse Particles in English Conversation*. New York: Garland.

Sprott, R. A. (1992). Children's use of discourse markers in disputes: Form-

function relations and discourse in child language. *Discourse Processes, 15*, 423-439.

Stenström, A.-B. (1990a). Lexical items peculiar to spoken discourse. In J. Svartvik (Ed.), *The London-Lund Corpus of Spoken English: Description and Research* (pp. 137-175): Lund University Press.

Stenström, A.-B. (1990b). Pauses in monologue and dialogue. In J. Svartvik (Ed.), *The London-Lund Corpus of Spoken English: Description and Research* (pp. 211-252). Lund: Lund University Press.

Stubbe, M., & Holmes, J. (1995). You know, eh and other 'exasperating expressions': An analysis of social and stylistic variation in the use of pragmatic devices in a sample of New Zealand English. *Language and Communication, 15*(1), 63-88.

Tomlinson, J., Jr., & Fox Tree, J. E. (2009, July 30 – August 1). *Talking it up: How the function of rising declaratives depends on prolongations and listeners expectations*. Paper presented at the 31st Annual Conference of the Cognitive Science Society, Amsterdam, The Netherlands.

Tyler, A. (1992). Discourse structure and the perception of incoherence in international teaching assistants' spoken discourse. *TESOL Quarterly, 26*(4), 713-729.

Van Der Wege, M. M., & Ragatz, E. C. (2004). *Learning to be fluently disfluent*. Paper presented at the The Cognitive Science Society.

Waltereit, R. (2002). Imperatives, interruption in conversation, and the rise of discourse markers: A study of Italian guarda. *Linguistics, 40*(5), 987-1010.

Wilkins, D. P. (1992). Interjections as deictics. *Journal of Pragmatics, 18*(119-158).