Louwerse, M.M. & Mitchell, H.H. (2003). Towards a taxonomy of a set of discourse markers in dialog: a theoretical and computational linguistic account. *Discourse Processes, 35*, 199-239.

Towards a taxonomy of a set of discourse markers in dialog: a theoretical and computational linguistic account.

Max M. Louwerse and Heather Hite Mitchell

University of Memphis

Running head: A taxonomy of discourse markers in dialog

Address for correspondence:

Max M. Louwerse

Department of Psychology

University of Memphis

Psychology Building

Memphis, TN 38152

Phone: (901) 678-2143

Fax: (901) 678-2579

Email: mlouwers@memphis.edu

Abstract

Discourse markers are verbal and non-verbal devices that mark transition points in communication. They presumably facilitate the construction of a mental representation of the events described by the discourse. A taxonomy of these relational markers is one important beginning in investigations of language use. While several taxonomies of coherence relations have been proposed for monolog, only a few have been proposed for dialog. This paper presents a taxonomy of between-turn coherence relations in dialog and discusses several issues that arise out of constructing such a taxonomy. A large number of discourse markers was sampled from the Santa Barbara Corpus of Spoken American English. Two judges substituted each type of these markers for all other markers. This extensive substitution test determined whether hyponymous, hypernymous and synonymous relations existed between the markers from this corpus of dialogs. Evidence is presented for clustering coherence relations into four categories: direction, polarity, acceptance and empathics.

According to transactional theories of language and discourse processing (Clark, 1996; Schiffrin, 1987; Schlegloff, Ochs & Thompson, 1996), language is the act of communication that normally is coordinated between its participants. The speaker or writer of a message needs to coordinate when to say what, what to say to whom, how and why to say it. In writing this is often difficult because the hearer is not simultaneously present in the communicative act[1]. In dialog, speakers have the advantage that hearers are present (Clark, 1996; Goodwin, 1981; Schlegloff, 1996). They know whether they have the hearer's attention, whom they are talking to, when they can start and stop speaking, and what they can say. Hearers generally give clues on each of these aspects by providing feedback. Coordination between speakers and hearers consists of multifaceted tasks between the parties involved. For instance, speakers need to monitor whether hearers are attending to what is said (is the hearer making eye contact?), who they are talking to (is the hearer an authority?), when they are speaking (is there a pause in the conversation which allows the speaker to start speaking?), what to say (how to express a meaningful information?), and whether the speaker needs to follow up on an earlier piece of information (is there anything that is by convention expected from the speaker based on previous pieces of information). This makes dialog a very dynamic act of coordination. Take for instance the following dialog in the Santa Barbara Corpus for Spoken American English (SBSAE)[2], describing a dialog between a girlfriend and boyfriend, Cathy explaining algebra to Nathan:

KATHY: three times ex, minus four.

NATHAN: Three times e=x, minus four?

KATHY: Right.

NATHAN: alright, distribute first, right?

KATHY: mhm.

[1] We use speaker and hearer here instead of sender and receiver, writer and reader, speaker and hearer or addresser and addressee. Also, we are aware that the actions of speaking and listening are usually not mutually exclusive in dialog, but for reasons of clarity we prefer these terms over Participant A and Participant B. At the same time we acknowledge that the border between the roles of speakers and hearers is fuzzy and that they change continuously over the course of the conversation.

[2] For reasons of clarity we have removed all coding information in the examples from the SBSAE, like coding for times, pauses, noises, etc. The only coding information left are extensions of a word (e.g. *al=right*, *rea=lly*) and non-linguistic contextual information (e.g. *THROAT* sound, *HITTING WITH NEWSPAPER*)

NATHAN: Two ex minus= .. ex squa=red, three ex minus tw=el=v=e, you get, WH do that side, so you get .. ex, whoa. I don't want to do that. Negative ex squared, two ex minus, two minus ex, Um, Plus twelve ... Now do you factor this? after you do that?

KATHY: Yeah.

NATHAN: Yeah. oh but first I gotta take out that negative one, don't -- I mean that negative.

KATHY: mhm.

NATHAN: in front of that ex squared so I just, I can multiply that whole side by .. negative one again though X?

KATHY: Yeah. then you flip that sign over.

NATHAN: I have to flip that sign over if I do that?

KATHY: mhm.

NATHAN: see it's little rules like that, that I'm not gonna remember. so if it's a -- if it's less than or equal to, then, and there's a minus=, you have to flip the si=gn. Okay. Are you tired?

KATHY: N=ot really. I mean kind of but,

NATHAN: I'm gonna go home in just a few minutes.

KATHY: why.

NATHAN: cause I can work on this .. at home, and let you get some sleep.

KATHY: (HITS NATHAN WITH PAPER)

The interaction between Nathan and Cathy is a joint activity, where one participant signals the other on the progression of the dialog. When Nathan asks a verification question, Cathy provides the confirming feedback (*right*), which Nathan then acknowledges (*alright*). But Cathy sometimes leaves the answer in the middle (*mhm*), for instance cueing Nathan to find alternative solutions to the math problem. When Nathan asks Cathy whether she is tired, she does not say *yes* or *no*, but *not really* (followed *by kind of*), signaling Nathan that she is tired, but does not want to make this explicit.

We call the linguistic markers like these that bring the interactions between the participants together **discourse markers**. They draw the attention of a participant (*Yeah. **oh** but first*), they serve as turn-taking signals (***Yeah**, oh but first*), to mark agreement with the content expressed by the other participant (***Yeah**. then you flip that sign over*) and they could mark the response to Nathan's proposal to leave. Discourse markers can thus operate at different levels in the communication (e.g. attention, turn-taking, locution, illocution) and one discourse marker could serve multiple functions. Notice that each of these cues can likely be communicated with signals other than verbal markers. Nods and gestures in communication usually serve similar purposes, as do variations in intonation (compare *rea=ly?* with *really!*). Even actions like hitting somebody with a newspaper as in the example above can have such a signaling function. In this study, however, we will focus on linguistic markers, ignoring for reasons of simplification phonetic (like intonation and pauses) and extra-linguistic features (like hand gestures and head nods).

At least 150 different discourse markers can be distinguished in the English language use (Swan, 1996), but others estimate this number to be over 350 (Knott, 1996). The different calculation is related to the question what a discourse marker constitutes. For instance, a discourse marker like *oh* (Schiffrin, 1987) is not included in Swan (1996). Most studies in fact do not include interjections and discourse particles (e.g. Knott, 1996; Sanders, Spooren, Noordman, 1992; 1993). In other words, there seems to be some confusion on the definition of discourse markers. How can this large set of words be characterized and what is their function in discourse? Is the function of these seemingly mundane words important and worth studying? If one discourse marker can have multiple functions in language use, what are these functions? And finally, with such a large set of discourse markers and different functions, can we bring order in a set of seemingly disorganized words? We will try to provide answers to each of these questions.

Discourse markers

Discourse markers mark discourse. They instruct discourse participants how to consider an upcoming utterance, providing a path towards the integration of different components of language use into one coherent discourse (Schiffrin, 1987). In an extensive sociolinguistic study of a subset of twelve discourse markers Schiffrin provides a number of conditions that allow an expression to be used as a marker. For instance, a discourse marker has to be a sequentially dependent element that brackets units of talk, has to be commonly used in its initial position and have a range of prosodic contours. Furthermore, a discourse marker should operate in different discourse planes and operate at both a local and global discourse level. The sequential dependency condition for these discourse relations can also be found in Knott (1996) and Knott & Mellish (1996). Their analysis focuses on sentences rather than units of talk, but the condition seems similar to Schiffrin's. Knott and Mellish argue that any phrase in a naturally occurring context that can be isolated with his host clause and cannot be interpreted without further context is a cue phrase (or discourse marker; Knott, 1996: 40). Examples of cue phrases are *but*, *so*, *or*, *because* and *although*. Phrases like these can usually not be interpreted without context (e.g. ?*Because the streets are wet*). The initial position condition is related to sequential dependency criterion. Discourse markers usually occur in the specifier position of a complementizer phrase (e.g. Pritchett, 1992) and mark a structural boundary (Grosz & Sidner, 1986). They typically start a new structural unit. Related to the initial position condition is Schiffrin's condition of the prosodic contours. In fact, Hirschberg and Litman (1993) showed that the prosodic contour of discourse markers allows for distinguishing between sentential use and discourse use (e.g. *now* as a temporal adverbial versus a return-to-subtopic marker). Although these three conditions are helpful, they are problematic for markers of back-channel response (Yngve, 1970), including words like *uhhuh*, *okay* and *yeah* or acknowledgment tokens (*mhm*, *wow*). They can stand alone, as we have seen in the example before and are therefore not syntactically detached from a sentence. Also, most studies on discourse markers do not include markers like these (Halliday & Hasan, 1976; Knott, 1996; Knott & Mellish, 1996; Mann & Thompson, 1986;

Sanders, Spooren, Noordman, 1992, 1993). However, given that discourse particles and interjections do mark the discourse structure (as Schiffrin for instance shows) they will be included in this study. Alternatively, we would have to use additional terms like project markers (Bangerter & Clark, forthcoming), transition-point markers (Clark, 1996) or turn constructional units (Schegloff, 1996) to encompass all cues that mark the discourse structure.

Whether discourse markers are defined as devices for marking transition points in discourse (Schiffrin, 1987), as devices cueing hearer to a change in discourse structure (Grosz & Sidner, 1986; Knott, 1996), or as devices marking movement between two discourse units (Polanyi & Scha, 1983), they are the conversational glue that participants effectively use to hold the dialog together at different communicative levels. We argue that during any communicative act participants try to build a coherent mental representation of the information being communicated. The coordination between the participants in both monolog and dialog is facilitated by cues to facilitate this process. These cues inform the participants how to build a coherent mental representation. We have called these linguistically marked cues 'cohesion relations', and the representational relationships 'coherence relations'. Cohesion facilitates coherence: it is not necessary, but is weakly sufficient for comprehension (Louwerse & Graesser, in press). Cohesion relations in discourse can be expressed by various linguistic categories, such as anaphora, cataphora, adverbs, connectives and discourse particles. The latter often serve as interclausal relationships in the discourse. In dialog this is not only within turns like in monolog, but also between turns.

Turns can be defined as those discourse units consisting of one or more turn constructional units that govern the allocation of turns. Such a definition leads to circularity, but offers two insights. First, it shows that turns are not related to sentences. Secondly, turns are governed by a local management system that is coordinated by the discourse participants. Generally participants are good at giving and taking turns, at coordinating in the joint project and at allowing transition relevance places (see Clark, 1996; Levinson, 1983).

Furthermore, cohesion relations can macro-manage the discourse by structuring the hierarchy of discourse or micro-manage by relating adjacent text units. Bangerter & Clark (forthcoming) call the relations within a level of a hierarchy 'horizontal transition', those between levels 'vertical transition'. Elsewhere (Graesser & Louwerse, in press; Louwerse, 2002) we have called this local and global cohesion and coherence.

Most studies on discourse markers have primarily looked at one single discourse marker cueing a coherence relation. Recently, studies have looked at multiple discourse markers (Oates, 2000; Webber & Joshi, 1998). The advantage of studying multiple discourse markers is that the investigation takes into account that natural discourse often contains more than one marker. Furthermore, by looking at plausible (*and although it was cloudy, it didn't rain*) and implausible (**and but it was cloudy, it didn't rain*) combinations, we can get an insight into a single marker. For instance, we could hypothesize that *and* is more general than *although*, that general markers should always precede specific markers and that therefore *and although* is an acceptable combination, whereas *and but* is not. The disadvantage of studying multiple discourse markers is the scope. We only know a little about single coherence relations and very little about these relations in dialog. In this study we will therefore focus on single discourse markers. That is, we allow for phrases like *but then again* and *there you have it*, but we will not look at combinations of these cues.

In sum, discourse markers are one type of cohesion relation that cue coherence relations, namely used in dialog. They generally mark a transition point within a sentence, between sentences or between turns and do that at a local or a global level of the discourse. In spoken language they can furthermore be characterized by intonational features. Despite the breadth of the term discourse marker, we will constrain its extent within the scope of this paper and use discourse markers as single constituent relations marking between-turn local transition points in the discourse.

The importance of discourse markers in dialog

Are discourse markers (or coherence relations) as important as we seem to suggest? If one removes markers like *yeah*, *oh* and *mhm* from the dialog between Nathan and Kathy presented earlier, the dialog is still meaningful. Should these markers perhaps be seen as debris in language use?

Cohesion relations can be found in both monolog and dialog. In both language modes the reader/hearer needs to be cued as to how to build the coherent mental representation. Similar cohesion relations can therefore be found in both dialog and monolog. But dialog and monolog also differ in some important respects. Clark and Brennan (1991) and Clark (1996) point out the features of face-to-face conversation, the most basic setting of language use. In dialog participants are copresent, they are visible and audible and their communicative actions are instantaneous. Furthermore, the medium is evanescent, the actions remain unrecorded, and production and reception are simultaneous. Other dialog characteristics are the extemporaneousness of the actions, the self-determination by participants of these actions and self-expression of actions by participants. In many language settings not all of these features are present. In telephone conversations there is some form of dialog, but participants are not co-present and cannot see each other. In instant email messenger services, there is a form of dialog in which medium and control features apply, but the setting is written, instead of spoken. This means that the setting of language use has various dimensions (see also Whittaker, 2002). The communicative act can take place in a written or spoken form, it can take place in the form of a dialog or a monolog, and it can be official or personal. As Clark (1996) points out, if a type of discourse misses one or more of the features in the basic setting of language use, special skills and procedures of language use need to compensate for this deficiency. One would therefore expect significant differences in the linguistic features between monolog and dialog.

Biber (1988) however did not find a significant difference in linguistic features between monolog and dialog. In an extensive corpus linguistic study he investigated the similarities and differences between various language settings. He compared 67 linguistic features in 481 corpora (23 written and spoken genres) and

identified six underlying dimensions of variation in English: 1) involved versus informational production of discourse, 2) narrative versus non-narrative concerns, 3) explicit versus situated reference, 4) overt expression of persuasion versus lack of opinion and argumentation, 5) abstract versus non-abstract information and 6) on-line information elaboration versus no real-time production constraints. Despite the fact that these underlying dimensions define important and systematic relations among language settings, Biber did not find a dimension that defined an absolute distinction between spoken and written language. This suggests that salient linguistic characteristics in one setting can be overridden by a language user (Biber, 1988: 161). For instance, in dialog participants are limited by real-time production constraints, which would make dialog generally less dense than monolog. But this does not mean that dialog cannot be dense. In Clark's words, special skills and procedures of language use apparently compensate for any differences in the other dimensions of the language setting.

If dynamic coordination between speakers and hearers consist of multifaceted tasks between the parties is involved in dialog and this interaction is absent in monolog, one would expect differences in cohesion relations. Dialog users deal with real-time production constraints, resulting in dialog being more fragmented than monolog. Even if no overall difference in frequency of linguistic features is found, one can expect differences in the use of cohesion relations in different language settings. A corpus analysis is ideal for testing such a hypothesis.

Based on Biber's extensive data analysis, we can compare the frequency of cohesion relations in different settings. For reasons of clarification, we categorized a selection of the 23 genres used by Biber into three dimensions of language use settings: spoken versus written, dialog versus monolog, and formal versus informal. The reason we did not just compute frequencies between dialog versus non-dialog is as follows. Although it can generally be assumed that dialog is an informal spoken text genre, formal letters for instance form an exception. It might therefore be the case that we find frequency differences in cohesion relations, but that they are due to informal language rather than dialog. We selected a subset of the genres that were

analyzed by Biber (1988) on the basis of two criteria. First, genres were selected by fitness in all three categories. That is, those genres were chosen for which enough information was provided to consider them spoken or written, dialog or monolog, and informal or formal. Religion, humor, and skills and hobby genres, for instance, were not used in the analysis because it was not entirely clear which three categories each of these genres fitted. Secondly, the number of texts per genre in one subcategory had to be comparable to the number of texts in a genre fitting the opposite subcategory in order to allow for comparable cell sizes. For instance, the 88 texts in the press genre resulted in largely unequal cell sizes in two of the categories and were therefore not used. An overview of the selected genres is presented in Table 1.

TABLE 1 ABOUT HERE

Having selected the corpora in each of the categories, we next defined the cohesion relations. Because we used Biber's analysis, we were limited to groupings of his linguistic features. The cohesion relations adopted in this study can best be captured by five of Biber's linguistic features. There were four groups of adverbial subordinators that mark informational relations and one group of discourse particles that mark conversational relations. The first three of these groups were the causative adverbial subordinator *because*, the concessive adverbial subordinators *although* and *though,* and the conditional adverbial subordinators *if* and *unless*. According to Biber the group of other adverbial subordinators consisted of relations like *since*, *while*, *whilst*, *whereupon*, *whereas*, *whereby*, *such that*, *so that*, *such that*, *inasmuch as*, *forasmuch as*, *insofar as*, *insomuch as*, *as long as* and *as soon as*. The final group in Biber's classification that resembled our notion of discourse marker consisted of discourse particles (*well*, *now*, *anyway*, *anyhow*, *anyways*) used by participants to maintain conversational coherence.

Biber precisely reports frequency values for the linguistic features in each of the texts, which allowed for a statistical analysis of the number of coherence relations. It must be noted here that to account for differences in corpus size, Biber normalized the frequencies to a text length of a thousand words. Table 2 presents average of the incidence scores (number of occurrences per 1000 words) in the text per category.

TABLE 2 ABOUT HERE

A comparison between the incidence scores of Biber's causative, concessive, conditional relations and discourse particles categories showed that spoken discourse had a higher frequency of relations than written (all *F*s (1, 323) > 13.75, p < .01), dialog more than monolog (all *F*s (1, 323) > 6.04, p < .01) and informal more than formal discourse (All *F*s (1, 323) > 9.02, p < .01). The differences in frequencies were largest for discourse particles. Spoken discourse had more than ten times more discourse particles than written discourse, dialog had more than two times more than monolog, and informal discourse had two times more markers than formal discourse. Contrary to what we expected, in the rest group, containing markers other than *because*, *although*, *though*, *if* and *unless*, written discourse had significantly more markers than monolog (F (1, 323) = 13.75, p < .01).

These results show that cohesion relations are indeed more frequent in dialog than in monolog, in spoken language settings more than written, and in informal more than formal. This is true for causal, concessive and conditional connectives like *because*, *although* and *if*, as well as for discourse particles like *well* and *anyway*. It is not the case for markers like *whereas*, *as soon as* and *insomuch as*, which are more frequent in written than in spoken language. These cues seem to mark interclausal relationships only and tend to be used in formal discourse, which might explain the unexpected pattern.

Because the results show a clear pattern for all three categories with the highest frequencies for dialog, spoken and informal, we can draw the conclusion that a combination of spoken informal dialog has the highest incidence of cohesion relations. Two genres, face-to-face conversation and telephone conversations, qualify for this combined category. Indeed, compared to all other genres, these two genres have the highest incidence of cohesion relations in all five groups, particularly in discourse particles. How can these frequency patterns be explained? According to Clark's features of language settings described earlier face-to-face conversation is the primary setting in language use, and telephone conversations have the same features, except for copresence and visibility. So why are extra cues needed in a setting that belongs to the standard language setting? Recall that those genres other than the standard language setting of face-to-face conversation require special techniques and practices to compensate (Clark, 1996) and that extra markers are therefore not to be expected in the standard language setting of spoken informal dialog. More markers could mean more coherence, but dialog is presumably not easier to understand than monolog. The answer to the high frequency of discourse markers in dialog could lie in the organization of this discourse setting. It can be hypothesized that because spoken informal dialog has a more dynamic and emergent discourse structure than other genres, it has more cues and cues operating at different levels of the interaction. The question then is what these different levels of interaction are and what the role is of discourse markers at these levels.

A Model of Discourse Coordination

In written monolog, the writers can plan messages and use off-line reflection to edit their ideas. Cohesion lies primarily in the connectedness of the verbal information. Dialog has a more dynamic and emergent structure than written monolog and, to a lesser extent, spoken monolog. Participants cannot plan their messages well in advance, because the dialog might take unexpected directions due to the interaction. Instead of off-line reflection, a participant reflects his own idea and that of the other on-line. The participants' construction of a coherent mental representation is a two-way process, in which both the speaker and the hearer contribute to

the emerging structure. The hearer contributes to what the speaker says by back channel feedback, filling in words, interruptions, etc. The discourse participants cannot see the outline of the dialog when they are in the middle of it, only when it is finished (see Schiffrin, 1987: 23). Such an emerging structure potentially requires more cohesion cues to comprehend the information in the communicative act between the discourse participants. Participants need to cue that they are still following the conversation, whether or not they accept the utterance and whether or not they agree with the speaker. Furthermore, the evanescence, recordlessness and immediacy of the medium require more anchors in the communication (see Clark, 1996). One way of compensating for the lack of organized structure is to have a higher density of discourse markers in dialog. Various theories of discourse structure and language use have been proposed (see Gernsbacher, 1994; Graesser, Gernsbacher & Goldman, 2002; Van Dijk, 1997). We will here focus on Schiffrin's (1987) sociolinguistic theory that explains the use of a set of discourse markers and Clark's (1996) cognitive theory of language use.

In her model of discourse, which is also a model of coherence in talk, Schiffrin distinguishes five planes of discourse. The **information state** concerns what the speaker and the hearer know and what they know of each other. This discourse plane is constantly evolving, since the participants' knowledge about the world is distributed through talk. For instance, the discourse marker *you know* marks that the speaker has information available about something. It marks interactive transitions in shared knowledge. A marker like *oh* marks a focus of the speaker's attention, signaling the hearer, and thereby marks a joint focus of attention. The **exchange structure** does not concern the cognitive information states of the speaker and the hearer, but the alternation of roles between them. Units of talk in this discourse plane are turns. A marker like *and* signals that the speaker wants to continue a turn, whereas a marker like *so* completes a turn. Markers in the exchange structure signal negotiation of turns. The exchange structure, like the **action structure**, Schiffrin considers pragmatic, because of the negotiation process of information between the discourse participants. The action structure indicates preceding and intended speech acts. For instance, markers like *because* and *so*

signal motive and motivated action; *but* and *and* mark speaker continuation, regardless of the activity of the hearer. In contrast with the exchange structure and action structure, the **ideational structure** is semantic rather than pragmatic. Instead of being realized by language use, the ideational structure is realized by language. Its units of analysis are propositions or ideas. Connectives like *and*, *but* and *or* coordinate idea units. *But*, for instance, marks a contrast between two propositions. Finally, in the **participation framework** speaker/hearer and speaker/utterance relations are captured. Participation transitions shift responsibilities from the hearer to the speaker and are centered around the accomplishment of an interactional task (p.217). For instance, the marker *I mean* can modify a speaker's prior ideas and intentions. A marker like well signals a digression before an indirect answer is given.

Schiffrin argues that discourse markers have one primary discourse plane (and eventually various secondary planes) in which they function. For example, the primary function of a marker like *oh* is to mark information state transition, but it might also mark certain actions in the action structure. Similarly, according to Schiffrin a marker like *now* has a primary role in the ideational structure, but since it marks the speaker's attention to the upcoming talk, it also plays a role in the participation structure. The marker *because* functions at the plane of information state and action structure but it has a primary role in the ideational structure. Thus markers may seem to be very different in miscellaneous expressions but share functions in the same discourse components (Schiffrin, 1987: 317). Schiffrin concludes that by categorizing markers in planes on which they play primary and secondary functions, some markers that may look very different might share the same function (*now* and *I mean* for instance function both at the participation framework). Similarly, those markers that look the same might have different functions on different planes (only the marker *then* has a function in the action structure, while a similar time deictic marker *now* does not). Schiffrin's model and classification is insightful not only in terms of multiple discourse planes, but it also helps us in understanding the different roles of discourse markers in language and language use. Furthermore, it helps us explain why

spoken dialog has more discourse markers than written monolog: in the latter many of the discourse planes remain implicit.

However, the Schiffrin model leaves some open questions. Schiffrin (1987: 29) argues that coherence is the result from the participants' efforts to integrate knowing, meaning, saying and doing. Different parts of one discourse plane (e.g. ideas, actions) are related to those of another (e.g. ideas and actions of another plane). Alternatively, different planes are related to each other (action structures are related to exchange structures; information states to participation frameworks). Without a further specification of these interactions, it is difficult to see the autonomy of the discourse planes. Also, markers operate on one or more discourse planes, but discourse participants operate on all of these planes. Take for instance the exchange structure. According to Schiffrin markers like *because* and *now* are not used in this discourse structure. However, it is difficult to see how a marker like *but* signals that the hearer wishes to take a turn, but a marker like *now* does not. Similarly, according to Schiffrin's theory a marker like *oh* does not operate in the exchange structure. But if this marker is used as an acknowledgment only, it signals the hearer to continue and hence plays a similar role as *and* in the exchange structure. Also, markers like *but* and *because* are not used in the participation framework, but if a marker like *well* marks responses at an interactional level, why would markers like *but* and *because* not play a role in the participation framework?

Take for instance the following passage from the SBSAE Corpus.

```
MILES:      well,
            if you're HIV positive,
            it's the same difference,
HAROLD:     mhm.
MILES:      since they feel that,
            sooner or later you'll come down with the actual disease.
PETE:       mhm.
JAMIE:      Yeah.
```

```
MILES:      but that's what she sai=d.
            now I don't know if she meant the Bay area or San Francisco,
```

According to Schiffrin the marker *but* does not operate in the information state. Recall that the information state accounts for the organization and management of knowledge and metaknowledge between the participants. *But* in the example below requires that the four discourse participants Miles, Harold, Pete and Jamie share the same knowledge. They are all on the same page on the doubtfulness of the claim the she made that sooner or later somebody comes down with the actual disease. Furthermore, according to Schiffrin *but* does not play a role in the participation framework either. Recall that the participation framework accounts for the relationships between the discourse participants who take stances in performing their speech acts. In the example below this seems to be exactly the case in the use of *but*. The marker contrasts the doubtfulness that is indirectly expressed by Jamie's and Miles's responses.

Examples like these suggest that markers could operate on discourse planes Schiffrin excludes for these markers. In general, it seems difficult to explain why according to Schiffrin some markers only operate at two discourse planes (*or*), others at three (*oh*, *and*, *but*, *because*, *then*, *I mean*), others at four (*so*, *y'know*) and yet others at all five (*well*).

Earlier we saw that language use operates are different levels and that these levels might be an explanation for the higher frequency of discourse markers in informal spoken dialog. Schiffrin's study of discourse markers proposes five of these levels and shows that one discourse marker can operate at multiple levels. What the study does not show is what the relation is between the different levels. Furthermore, Schiffrin argues for a primary and secondary function of discourse markers at some but not all levels of discourse. As we have seen, it can be argued that discourse markers do play a role at all levels. In consequence, it has become our task to point out what the relation between the different levels are and how discourse markers operate at each of these levels.

A model of language use that does specify the relation between different levels is Clark's (1996). Like Schiffrin (1987: 50), Clark (1996) assumes that communicative actions are joint activities. In addition, however, Clark argues that like most other actions, communicative actions come in hierarchies. His **action ladders** consist of levels of joint actions in which participants are taking their part. Participants must capture the actions in progress to be able to act jointly and be successful in the communication. Cark distinguishes four levels in the action ladder, but doesn't rule out additional levels (p. 149). In the first level a speaker executes a behavior for the hearer who in turn attends to that behavior. An example of this interaction between **execution** and **attention** is attending to a speaker's articulation of a sentence. In a second level speaker present a signal to the hearer who in turn identifies that signal. An example of **presentation** and **identification** is the identification of the speaker's articulated sounds by the hearer. Notice that interaction at this second level can only be established if a behavior is executed and responded to with attention, i.e. completion of the first level. The second level enables a speaker to signal meaningful information to the hearer, who in turn recognizes and understands that information. An example of this exchange of **signaling** and **recognition** (or **meaning** and **understanding**) is the understanding that the speaker says "Run!" The meaning/understanding pair can only be established if the hearer attends to an execution of behavior, and if the hearer recognizes the signal presented by the speaker. Level 3 enables a fourth and final level in which the speaker proposes a joint project to the hearer who in turn considers this joint project. Again, this level of **proposing** a joint project and **considering** implies the other three levels. In our example the hearer might consider the speaker's proposal to run.

These actions in the four levels form an action ladder. Actions in the action ladder are co-temporal, that is, they begin and end together. Furthermore, there is a relation of causality between the actions, that is, the actions are asymmetric, irreflexive and transitive. This leads to two properties in the hierarchical organization of actions. One is upward completion, the property of enabling higher-level actions only by

having completed lower-level actions. Reversely, the property of downward evidence ensures that if one action in the ladder is complete, those in lower levels also have to be complete.

If we are able to map Schiffrin's discourse planes onto Clark's action ladder and can show that discourse markers play a role at each of the levels in the action ladder, we might have a theoretical answer to the question why discourse markers are more frequent in informal spoken dialog. In addition we would be able to structure discourse markers on the basis of their function on the action ladder. We propose a mapping presented in Table 3.

TABLE 3 ABOUT HERE

According to Schiffrin the primary function of the discourse marker *oh* is at the information state, marking the recognition of old information or the receipt of information. In the example below Rebecca signals Rickie that she is still paying attention to Rickie's realization of utterances. Schiffrin's information state thus seems to show similarities with Clark's Level 1. Note that any other signal of attention (markers like *hmm*, *yeah*, nodding, regaining eye contact) to the speaker's utterances would fulfill the function of the marker *oh* in this example.

REBECCA: He went through these doors?
RICKIE: into it unhunh,
into another car.
REBECCA: Oh=.
RICKIE: and then he came back again,
and then,
I guess that was like another stop,
and more people were getting off,

and the=n,

that's when he came .. and s=at b=y me,

sat .. in this seat right here.

Furthermore, Schiffrin states that the information state organizes and manages knowledge and metaknowledge. This seems a prerequisite for any of the other discourse planes. In Clark's Level 2 the speaker presents a signal and the hearer identifies that signal. Identification of the signal is more than just attending to it. The hearer needs to show identification of the speaker's presentation of the utterances at the right moment in the conversation, usually at the end of a speaker's turn. Rebecca signals the identification of the utterance not in the middle of Rickie's sentence, but at the end of it. Note that by using *oh*, instead of for example *by the way*, she simultaneously signals that the speaker can continue. This kind of signaling seems to take place primarily between turns, the units of discourse in Schiffrin's exchange structure. Let's assume that the interaction at the information state enables the interaction at the exchange structure: the receipt of information enables the alternation between of roles between speaker and hearer. According to Clark's theory this enables meaning and understanding. This semantic level is easy to associate with Schiffrin's ideational structure, in which relations between propositional units are coordinated. This Level 3 in Clark's theory enables the fourth level in which the speaker proposes a joint project and the hearer takes it up. For instance, when Kathy tells Nathan to do the calculation himself, Nathan responds with *well*, marking that he hesitates to take up Kathy's proposal. This indirect way of responding, a digression from a direct *yes* or *no*, is what Schiffrin considers a key component in the participation framework.

NATHAN: Oh this is easy. P Ex squared P. Please say this will factor?

Will it?

KATHY: na, you do it.

NATHAN: well, I mean, that's just wasting time. cause if it's not

NATHAN: It does?

KATHY: mhm.

NATHAN: One and negative two-fifths?

KATHY: mhm.

NATHAN: and I can always put those back up into the to=p,

KATHY: THROAT

NATHAN: and, and see if they check. Right? Let me just try the one? Fi=ve, THROAT I got zero equals zero=.

KATHY: tha- that's right.

Four out of the five discourse planes Schiffrin identifies can be mapped onto Clark's action ladder. One discourse plane, the action structure, seems to diverge from the action ladder. This discourse plane deals with situating the speech act in their linear sequence. It seems to operate at what we earlier referred to as a global or macrostructural level of the discourse and what Bangerter & Clark (forthcoming) have called 'vertical transition'. Since we are primarily interested in adjacent actions, the action structure does not fit any of the four levels, but seems to be more likely relate to a fifth level that the other levels enable.

Linking two different models, one of coherence in discourse and autonomous discourse planes, the other of language use and fully dependent hierarchical action levels, is difficult. However, the similarities between the levels seem to allow for such a comparison. With this comparison we can explain the relations between Schiffrin's discourse planes and we can provide further evidence for Clark's action ladder. If we assume that discourse markers function at different levels of the discourse and we assume that these levels are not independent of one another, we can argue that discourse markers cue hearers at all levels of discourse, at least at the four levels identified in this section. With this framework of action levels on which discourse markers operate, we might come closer to an organization of discourse markers. Furthermore, with this framework there is further evidence for the earlier hypothesis on the high frequency of discourse markers in informal spoken dialog. Because of the presence of all four levels in spoken dialog, markers need to guide discourse participants in the conversation.

Classifications of Coherence Relations

Various classifications of coherence relations have been proposed, some focusing on the linguistic realization of these relations (e.g. Halliday & Hasan, 1976; Martin, 1992), others on conceptual relations that do not necessarily have to be marked (Core & Allen, 1997; Hobbs, 1985; Mann & Thompson, 1986); some from a theoretical linguistic perspective (Hallliday & Hasan, 1976) or a sociolinguistic perspective (Schiffrin, 1987), others from a computer science perspective (Marcu, 2000) or psycholinguistic perspective (Louwerse, 2002, Sanders, Spooren & Noordman, 1992); some primarily for monolog (e.g. Hobbs, 1985; Mann & Thompson, 1987), others for dialog (Core & Allen, 1997; Di Eugenio et al., 1998). Although these classifications differ considerably from each other, there are some recurrent categories. We will discuss some of the common categories by first discussing the classifications aimed at relations in monolog, followed by those designed for relations in dialog. The aim here is not to provide a complete and detailed overview of the various proposals, but instead to point out some of the recurrent categories between proposals in order to compare the classifications of monolog and dialog and map them onto the discourse model introduced earlier.

Some classification schemes claim that coherence relations between text units can exist without them being linguistically marked (e.g. Mann & Thompson, 1986). Other studies only use the linguistically marked relations in order to build a classification scheme (Halliday & Hasan, 1976). Louwerse (2002) found four recurrent categories in these proposals, after comparing classifications proposed by Halliday & Hasan, 1976; Hobbs, 1985; Martin, 1992; Mann & Thompson, 1986; Sanders et al. 1992, 1993, Knott & Dale, 1994, Knott & Mellish, 1996;

First, all classifications of coherence relations use a notion of causality, temporality and additivity, marking the semantic relationship between text segments. Examples of these relations are causal relations like those marked by *because* and *although*, temporal relations like those marked by *before*, *after* and *until*, and additive relations marked by *moreover* and *however*. These categories take a slightly different form in some other

proposals, with relations like justification and evidence relations. The general semantics of the relation between text segments is however shared.

Secondly, all taxonomies have some kind of contrastive or adversative relations, in addition to their positive counterparts. 'Positive' relations are for instance marked by *because*, *before*, *moreover,* 'negative' relations by *although*, *until* and *however*. The polarity of the relation marks whether a text segment maps onto the previous text segment.

Thirdly, all classifications accommodate a difference between coherence relation operating between the relations in the world or relations between speech acts. These relations are often called semantic and pragmatic (Sanders et al. 1992, 1993), causal versus diagnostic (Traxler, Traxler, Aked & Moxey, 1997) or external and internal (Halliday & Hasan, 1976, Martin, 1992). In English these relations could be marked by *because* and *since* (as in the sentences *The streets are wet because it rains* versus *It rains since the streets are wet*). In other languages the difference is more apparent, with different conjunctions for the two different readings (French *parce que* and *car*; German *weil* and *denn*; Dutch *omdat* and *want*).

Finally, in all proposals relations are distinguished that mark the order of the clauses that are conjoined. Relations can either be directed forward to the upcoming text segment (*so*, *nevertheless*, *moreover*) or backward to the previous text segment (*because*, *although*, *until*). This category reflects the order of presentation as opposed to the order of the events in the world.

For those classification schemes aimed at dialog a similar comparison can be made. Behind these proposals (implicitly or explicitly) lies the idea of adjacency pairs (Schegloff & Sacks, 1973). These pairs of utterances have the component utterances positioned adjacently, while each part of the pair is produced by a different speaker. Speaker A produces the first part (e.g. question), followed by speaker B producing the second pair part (e.g. answer). Mann (1988) for instance proposes an episodic structure of these adjacency pairs that conforms to a set of conventions. Dialog participants participate in so-called dialogue games, conventions of interactive goal pursuit in order to pursue individual as well as collaborative goals. A more

general proposal to annotate a set of task-oriented dialogs comes from Carletta, Isard, A., Isard, S., Kowtko and Doherty-Sneddon (1996). They also build onto the idea of adjacency pairs (for instance by dividing between initiation and response, question and yes-reply, statement and command acknowledgment, etc). Similarly, Core and Allen (1997) propose a dialog annotation tool that is used for their task-oriented dialogs. Core and Allen distinguish between different layers in their dialog act mark up scheme. Although these layers can also be deduced from Carletta et al.'s proposal, they are well specified in Core & Allen's annotation scheme that has multiple layers. There is direction category, that takes a forward or backward looking function, depending on whether the current utterance is related to the previous or future discourse. There are utterance features with information about the form and content of an utterance. The proposal has been very influential on other proposals that extend or revise it (Di Eugenio et al. 1998; Poesio & Traum, 1998). As with other proposals, the scheme conforms to the notion of adjacency pairs with relation pairs like Info-Request and Answer, Offer, Accept, etc. Least specified are so-called Informational Relations that describe how the information in one utterance relates to the information in another.

When these different proposals are compared, the following recurrent categories emerge. First, all proposals accommodate relations concerning whether or not the speaker understood (i.e. heard) an utterance. Good examples of utterances where the speaker signals that the antecedent is not understood are metacommunicative expressions like "Can you repeat that?". Monolog obviously does not have such a category, because of the lack of instantaneity, evanescence, recordlessness, and simultaneity of the utterances in this language setting (Clark & Brennan, 1991).

A second category recurrent category concerns whether the participant agrees to what the speaker said (for instance, whether the topic of negotiation is being accepted or rejected). Most taxonomies assume at least four features for this category: the speaker fully accepts or fully rejects, or partially accepts or partially rejects. This category shows strong similarities with the 'positive' and 'negative' relations found in monolog classifications.

Finally, as in monolog, the direction of the communicative functions can be either forward or backward. That is, current utterances can either relate to the subsequent dialog or interaction or relate to previous discourse. A request for information would be an example of forward relations, an answer to a question an example of a backward relation.

Obviously, the classifications of coherence relations in dialog are much more elaborate than presented here. However these are the most apparent recurrent categories between proposals. Moreover, these categories show similarities with those recurrent categories found for classifications in monolog. The presentation of the relation (forward or backward) can be found in the classifications for both language settings, as well as the polarity of the relation (positive, neutral, negative). The metacommunicative category for classifications of relations in dialog cannot be found in monolog, most likely because monolog does not allow for marking signal detection.

The question is how the recurrent categories for monolog relate to those in dialog and how both relate to Schiffrin's discourse model and Clark's action ladder? Existing classifications of coherence relations allow us for some rudimentary comparison that shows that there are similarities between the recurrent categories. In general it seems that the use of a cohesion relation in itself solicits the hearer's (or reader's) attention. Given that cohesion relations are weakly sufficient but not necessary for coherence, their presence functions as a cue. By articulating the relation, the speaker (or writer) notifies the hearer (reader) to pay attention to the relation. The presence of the relation seems to apply to the first discourse level of information framework. This cue for attending to the relation enables the presentation and identification of the signal. Does the speaker want to mark what was said before or what is coming? Does the speaker want to say more (e.g. in the case of *and*) or not (e.g. in the case of *oops*). These backward and forward relations seem to operate at the second discourse level, the exchange structure. The marking of a relation and the cue in which direction the relation is to be made enables the meaning and understanding of the relation. Both in monolog and dialog we find positive and negative (non-adversative and adversative) relations. In monolog the semantics are

further specified with causal, temporal and additive relations. In monolog we also find a difference between relations in the world and relations between speech acts. Relations between speech acts seem to be built onto the relations of the world, see Traxler et al.). That is, processing the speech act relation *It rains because the streets are wet* is possible by means of understanding the relation *The streets are wet because it rains*. The semantic category (Level 3) would enable a speech act category (Level 4).

As we pointed out at the beginning of this section, it has not been our aim to present a straightforward comparison of classifications and map them onto discourse models. We are very well aware that such a comparison – if possible at all – is far more complicated. We have however tried to show that some relations that are similar in nature can be found in various proposed classifications and that these relations could be mapped onto the action-ladder discourse model. This framework of recurrent categories and discourse model might help us in building a theoretically constrained taxonomy of discourse markers grounded in a corpus.

Toward a Taxonomy of Discourse Markers in Dialog

Why do we want to classify the discourse markers? The evident answer would be scientific: to find order at all points (Jefferson, 1984). An alternative related answer is to provide insight in cognitive processes. It seems unlikely that language users use different cognitive processes for the many different discourse markers that are available. Clustering of linguistic markers might help us understand underlying processes; cohesion can tell us something about coherence (Louwerse, 2002; Sanders et al. 1992, 1993). Although an understanding in the cognitive processes of understanding and generating discourse markers is certainly one of our aims in developing a taxonomy of discourse markers, there is another related answer. From a computational linguistic perspective it is helpful to have an underlying structure behind a group of linguistic features, for instance when one develops an intelligent conversational system. One such system is AutoTutor, a conversational agent that assists students in actively constructing knowledge by holding a conversation in

natural language (Graesser, K. Wiemer-Hastings, P. Wiemer-Hastings, Kreuz, & the Tutoring Research Group, 2000). The question is how to build a taxonomy?

Earlier we pointed out that we will focus on between-turn single discourse markers in dialog. Our aim is to build a theoretically constrained taxonomy of these discourse markers in dialog that is grounded in a corpus, by determing the semantic relationships between the markers. By defining synonymous, hypernymous, hyponymous and exclusive relationships between markers we might be able to form a network of relations. By finding the underlying features that constitute these relations, we can next reduce the network into a taxonomy of discourse markers. Ideally, the categories in the taxonomy should show some resemblance to existing classifications as well as to the action-ladder discourse model outlined in the previous section.

Method

After selecting large numbers of text segments from a corpus, each containing a discourse marker, we substitute the discourse marker in each segment with each discourse marker from the whole set of discourse markers that can be found in that corpus. Knott (1996; Knott & Dale, 1994; Knott & Mellish, 1996) used such a substitution test before. The method shows similarities with Martin's (1992) substitution tests for explicit and implicit relations and paratactic and hypotactic relations, and Mann & Thompson's (1987) claim of independendency of text spans. After selecting large numbers of text segments from a corpus, each containing a discourse marker, the discourse marker in each segment is replaced by each discourse marker from the whole set of discourse markers that can be found in that corpus. If a marker can be replaced without significantly changing the meaning of the initial relation, we can argue that there is a semantic relationship between these markers. Four kinds of relations can be distinguished. Two markers are considered synonymous if they are inter-substitutable. If one marker can be replaced by the other, but not the other way around, the latter is a hyponym of the former (and the former a hypernym of the latter). If two

markers cannot be substituted in any given context they are exclusive. Figure 1[3] presents examples of the four relations for potential discourse markers in dialog (where the arrow means "replaces").

FIGURE 1 ABOUT HERE

Substituting *n* discourse markers results in a large web of (*n* x *n* - *n*) interrelations between these markers. Accordingly, a synonymous, hypernymous, hyponymous or exclusive relation can be defined for each discourse marker, excluding the reflexive relations.

Corpus

The Santa Barbara Corpus of Spoken American English (CSAE) was chosen as the corpus for the study because of its naturalness and diversity. The CSAE is part of The International Corpus of English (ICE) that started in 1990 with the primary aim of collecting material for comparative studies of English worldwide. The CSAE corpus consists of speech and transcript files of 14 texts, each between 15 and 30 minutes. The total corpus contains 66436 words in 7646 turns. The conversations in the corpus are primarily face-to-face between American English speakers. The corpus has various advantages over other corpora. First of all, the corpus is heterogeneous in the sense that various regional accents and dialects are represented: a total of 52 different dialects of English from Midwest to East coast. The corpus is also heterogeneous in sex, race, age and level of education of speakers. Of the 50 speakers in total, 28 are female, and speakers range in age between 17 and 90 years. Seven different ethnicities can be found among the speakers, with the majority (37%) Caucasian. Highest level of education for most speakers is college level (BA, MA, PhD or postdoc),

[3] Obviously, the hypernyms and hyponyms show identical relations, because every hypernymous relation between *p* and *q* is a hyponymous relation between *q* and *p*.

but for about 25% high school is the highest level. We have seen that discourse markers are likely to be most frequent in informal spoken dialog, which is the setting of all dialogs in this corpus. At the same time, it includes a range of different topics (tutoring, discussions, gossip). These features make the CSAE ideal for the purpose of determining relations between discourse markers in natural dialog.

Procedure

In addition to the theoretical definition of discourse markers, we used a practical working definition to select the markers in the corpus. Discourse markers are those linguistic items defined as cue phrases (Knott, 1996), coherence relations (Sanders, Spooren & Noordman, 1992; 1993), conjunctions (Halliday & Hasan, 1976; Martin, 1992) and discourse markers (Schiffrin, 1987). In addition, we used any linguistic item labeled as discourse particle from the Penn Treebank project and the CELEX database. This total set of over 400 discourse markers was used to identify the between-turn discourse markers in the Santa Barbara corpus. The resulting set had 196 markers that were used in this study.

The 196 discourse markers marked the start of a total of 3817 turns in the Santa Barbara corpus. We next selected five lines before and after the discourse marker, to provide a context for the substitution test. The set of materials consisted of 744315 substitutions to be made ((196 types x 3817 tokens) – 3817 reflexive relations). However, many of these 744315 substitutions are rather redundant, since substitutions for samples of the total frequency of a discourse marker can be reliably generalized over that total frequency. For instance, if *because* appears to be a hyponym of *and* 20 times in a row, we can reliably predict this will be the case for all 331 occurrences of *and*. To reduce the number of redundant substitutions a normalization procedure for the frequent discourse markers was used. Ideally we would like to have at least five substitutions per discourse marker. However, we also want to do justice to the marker frequency, as it can be argued that more frequent markers are more ambiguous. Therefore, five tokens of one marker type were selected, and 5 percent of the total frequency of markers was added, resulting in a total of 747 marker tokens and 145665 substitutions.

Two graduate students in psychology were trained to perform the substitution test. They studied the methodological section in Knott (1996) and went through a series of exercises before they made the substitutions of the 196 markers types. In a matrix of 196 columns and rows they recorded for each text segment whether or not a substitution was possible. Based on the frequencies in rows versus columns it could then be determined whether a marker was a synonym, hypernym, hyponym or exclusive of another marker.

An example of two text segments is given in the following dialog between Sharon and Carolyn, where *unbelievable* can be replaced by *man* in the first text fragment. In the second fragment *man* can be replaced by *unbelievable*. If for every occurrence in the corpus a discourse marker *p* in context *c1* can be replaced by discourse marker *q* and discourse marker q in context *c2* can be replaced by discourse marker *p*, markers *p* and *q* are considered synonymous.

Text fragment 1.

SHARON: another month and a half,to e- have any lunch,because you can't access, you know,her form?

CAROLYN: . . .

SHARON: What's the deal.

KATHY: **unbelievable.**

CAROLYN: They're just giving -- I think,it sounds,like,to me,they're giving you a lot of shit for no reason.

Text fragment 2.

KATHY: They fired him?

SHARON: oh, Mister Samuel,

CAROLYN: He -- k- -- sh- [God,

SHARON: He was the most --

CAROLYN: **man**, he threatened little kids.

SHANE: . . .

CAROLYN: We'd all be sitting in a circle, and it'd be current events time, we'd all be sitting there,

SHARON: He was harsh.

Similarly, if in all text segments *unbelievable* could be replaced by *man*, but not the other way around, *unbelievable* would be a hypernym of the marker *man*. If *unbelievable* could not be substituted by *man* and not the other way around, the two discourse markers would be considered exclusive.

Analysis

After the two judges recorded a total of 145665 times whether or not a substitution was possible, the results from both judges were added together. Contrary to Knott's (1996) study, this study was conducted by two judges instead of one and used more than one text sample with a particular discourse marker. This allowed for two levels of analyses, one relating to the agreement between the two judges, the other relating to the proportion of substitutions.

In the first level of analysis, related to the agreement between the judges both an **independent** and **dependent** judgment analysis could be performed. The results from both judges could be added together, regardless of the interrater reliability (independent judgment). The advantage is a large number of **potential** substitutions that can support a taxonomy. The disadvantage is that the substitution is a subjective artifact since the judges do not necessarily agree on the substitution. Alternatively, only those substitutions are used in which the two judges fully agree (dependent judgment). That is, if one substitutes a marker 26 times and the second only 25 times the substitutions are excluded from the analysis. The advantage of this more conservative analysis is that only the most apparent substitutions are used for further analysis.

The second level of analysis relates to the proportion of substitutions. Since this study uses more than one text sample in which a marker can be substituted, two conditions can be distinguished, one in

which a discourse marker can always be substituted by another, and one where this is only sometimes the case. We will call the first (conservative) condition **strong** relations, the second (liberal) **weak** relations. To determine whether a relation is strong or weak, one obviously needs to account for the frequency of the discourse markers in the corpus. Furthermore, one needs to allow for mistaken oversights and inserts by the judges. A margin was therefore applied: relations that could only be substituted in less than 50% of the cases were considered erroneous. If relations were substituted between 50-95% of the time, it was considered a weak relation. If substituted above 95% of the cases, the relation was considered a strong relation.

The frequencies of substitutions in a matrix of 196 x 196 cells could now be analyzed using independent and dependent judgments and weak and strong relations. The independent judgments using the weak relations would yield the most dynamic results (potential substitutions by at least one rater). The dependent judgments using the strong relations would yield the most secure results (apparent substitutions by two judges). The number of substitutions of markers presented in the column of the table with markers in the row of the table could now be compared with the number of substitutions of row-markers with column-markers. If marker x could replaced in more than 95% of the cases with marker y, but marker y could not be replaced in more than 50% of the cases with marker x, marker y was considered a strong hypernym of marker x. This method would work nicely for equal distributions of discourse markers. However, some discourse markers are more frequent than others. We therefore also considered the frequency of each marker in the SBSAE Corpus, by normalizing the frequency of discourse markers per 1000 words. A computer program was written that compared the number of substitutions between discourse markers relative to their frequency in the corpus. The result was a long list of markers, their hypernymous, hyponymous and synonomous relations with other markers, whether this relation was weak or strong, and based on a dependent or independent judgment.

The interrater reliability between the two judges was high (Kappa = .71). This result was found in the independent judgment category using both strong and weak relations. This high interrater reliability suggests

that the substitution method is not entirely subjective, particularly because we used *both* weak and strong relations rather than only the least controversial strong relations. Furthermore, we used independent judgments rather than the dependent judgments which would per definition lead to high interrater reliability.

The complete network of synonymous, hypernymous and hyponymous strong and weak relations between all discourse markers in the independent judgment conditions is far too large to be presented in a single table. Because the whole network cannot be presented in printed form, Figure 2 shows a subset of the network of relations.

FIGURE 2 ABOUT HERE

The dependent-judge strong-relation graph shows that according to substitutions of discourse markers in the SBSAE Corpus markers like *yep*, *mhm*, and *oh yeah* are synonymous. In more than 95% of the cases when a speaker uses *yep* this marker can be replaced by *oh yeah* or *mhm* without considerably changing the semantics and syntax of the relation. On the other hand, a marker like *oh yeah* can almost always replace the marker *that's fine*, but not vice versa. This means that *oh yeah* is a hyponym of *that's fine*. Notice the complexity of the network. Whereas *oh yeah* and *mhm* are synonymous, *that's fine* is a hypernym of *oh yeah*, but a synonym of *mhm*. These counterintuitive relations could be explained by the error margins we allowed (50% and 95%), but are more likely to say something about the dynamics of discourse.

It is worth pointing out that it is feasible to find counterexamples for the relations presented here. We can think of examples where *hell* might not be replaced by *since* or *basically*, whereas according to Figure 2 this is always possible. We can also think of examples where discourse markers like *hell* and *man* can be replaced, although these markers are mutually exclusive according to Figure 2. In the corpus data we have

used, the raters did not encounter these relations. In other words, the network of relations between discourse markers is constrained by the corpus we have used.

Interpretation of the Taxonomy

The overall network of relations nicely illustrates the many possible relations between discourse markers in dialog. However, such a visual representation is not very meaningful if the semantics and pragmatics of these relations is not specified. To speak of a **taxonomy** of coherence relations, the various relations in the network need to be reduced into meaningful categories. The question is how to reduce this large set of relations between the markers. Because we are dealing with a large number of observed variables and want to identify underlying variables, reduction techniques like factor analyses are useful at least for exploratory purposes.

Factor analyses help to discover the nature of relationships between variables. However, we only have the number of substitutions to work with, with generally low numbers and small variances. Nevertheless, we entered the scores in a factor analytic model, using a principal components factoring procedure that attempt to account for all the variance in the data. To compensate for the fact that the first factors account for the greatest portion of the variance of the markers and do not leave much for subsequent factors, factors will be rotated. That is, after rotation factors will be more equal in size than the initially extracted factors. To allow for correlations between the factors, we used the Promax rotation (Biber, 1988; Comrey & Lee, 1992).

Absolute values less than .30 were excluded as unimportant. The first ten factors appeared to account for 76.39% of the shared variance, based on the results in the scree plot. The scree plot is a graph of the eigenvalues against all the factors. It helps to decide how many factors to retain. The point where additional factors do not contribute to the overall analysis is where the curve starts to flatten.

A total of 114 discourse markers were retained in the factor analysis. Those markers that were not included in these ten factors were the basis for the first category in the taxonomy. One group consisted of

very infrequent markers (*doo-doo-doo*, *ni*, *ola*) in which they were dispensed with since there were not enough substitutions in the corpus to create an informative relation. The other group, however, consisted of markers like *earlier*, *before*, *excuse*, *after*, *assumed*, *ready*, *as long as*, *thanks*, many of which were temporal deictic markers. These markers did not directly follow up on the speaker's utterance, but instead started a turn in which lexical information following the marker cued a relation to the previous utterance. They should therefore be considered within-turn relations that happen to be at the start of the turn. They will not be further classified, but marked as **FORWARD** relations.

The remaining 114 discourse markers included in the factors will be called **BACKWARD** relations. They respond to the previous turn. To derive the categories of the BACKWARD relations included in the factors, the positive features were compared with the negative features per factor, which created two distinct groups. Interestingly, the factor itself did not allow for a label, but the various groups across factors did. To further discriminate between the two groups per factor, the features for the positive scores on the one hand and the negative on the other were divided into two equal groups. Next, the values were ordered by factor score. This strictly objective method resulted in four taxonomy categories. The category direction consisted of markers that had either backward or forward relations. A **polarity** category marked the continuity of the turn with either positive, neutral or negative relations. In an **acceptance** category each of these relations could be accepted or partially accepted. Finally, a category of **emphatics** marked how strongly a speaker felt. This factorial combination of DIRECTION (2) AND FOR THE RETURN RELATIONS CONTINUITY (3), ACCEPTANCE (2), and EMPHATICS (2) offered 12 kinds of connective relations, realized in common discourse markers.

The INDEPENDENT JUDGMENT analysis and the DEPENDENT JUDGMENT analysis resulted in identical factors with identical categories, obviously with a smaller number of markers in the DEPENDENT JUDGMENT. Similarly, including both WEAK and STRONG relations resulted in the same categories as including only the STRONG. To maximize the number of relations included in the taxonomy, we therefore selected INDEPENDENT judgments and both STRONG and WEAK relations. Examples for each of the relations from

the Santa Barbara Corpus of Spoken American English are given following the relations. An overview of the discourse markers in each of the categories is given in Appendix 1.

DIRECTION

1. BACKWARD relations relate to the implicature of the previous utterance expressed by the speaker and do not require any further communicative act;
2. FORWARD relations relate to the implicature of the previous utterance expressed by the speaker, but also require a further communicative act in addition to the relation.

The principal component analysis filtered the FORWARD relations from the BACKWARD relations. Although an analysis of the FORWARD relations is desirable, the current analysis does not provide information on additional categories for these relations. An additional principal component analysis on the forward relations only did not generate any underlying categories. However, since the forward relations like *because*, *although* and *before* can also be used as within-turn relations it seems likely they can be categorized in existing taxonomies of interclausal relationships in monolog.

A further analysis of the backward relations generated three additional categories in the taxonomy: POLARITY, ACKNOWLEDGMENTS and EMPHATICS. By contrasting the positive factor scores with the negative scores across factors, a second category emerged, which we label **polarity**.

POLARITY

1. POSITIVE relations follow up the response expected by the speaker.
2. NEGATIVE relations violate the response expected by the speaker.
3. NEUTRAL relations do not explicitly follow up on or violate the response expected by the speaker.

Examples for each of these features in the category are given below.

POSITIVE

JAMIE: I just read an article on him.

.. You .. you probably read the same Examiner article

Yeah,

probably,

yeah.

MILES: talking about how Gregory Hines said,

he doesn't realize a human being can't tap that fast?

JAMIE: **Right.**

Right.

PETE: . . .

JAMIE: **Yeah.**

NEUTRAL

DARRYL: That's right.

It'd be a different personality.

PAMELA: So .. I, .. I, ... I think it's, very fascinating,

DARRYL: Of course that's a hypothetical, how do you know that.

PAMELA: it's very interesting

DARRYL: **Maybe, may-,**

PAMELA: . . .

DARRYL: **maybe,**

PAMELA:	. . .
DARRYL:	**maybe**,
	maybe the spacesuit has something to do,
	with,
	.. with who's inside of it.
	.. I mean you don't know that.

NEGATIVE

LENORE:	Did they train you-
LYNNE:	yeah.
LENORE:	Did they train you that -
	. . .
LYNNE:	yeah.
	... yeah.
LENORE:	.. So you have your own equipment,
LYNNE:	…
LENORE:	but,
LYNNE:	**No**.
	I don't have my own equipment at all.
	... Dad, ... you know, has done some of it.

In positive relations the hearer follows up on the utterance expressed by the speaker, marking agreement. In the example above the markers *yeah* and *right* cue the hearer that the speaker agrees with the content of the utterance. Miles' utterance that Gregory Hines doesn't realize a human being can't tap that fast, Jamie replies

"*Right"*, acknowledging that he does not realize this. In negative relations the opposite is the case. Markers like *no* mark that the hearer disagrees with speaker. In the above example, when Lenore assumes that Lynne has her own equipment, Lynne replies with no and reiterates this reply by saying that she does not have her own equipment at all. In neutral relations the hearer neither agrees nor disagrees, but leaves the response in the middle. In the above example *maybe* marks that the hearer does not fully agree with the speaker, but he does not disagree either (he leaves the option open to either agree or disagree).

Note that this category and its members are formed by the principal component analysis. This means that we can do not rule out situations where positive relations in which the hearer agrees with the speaker are marked by *no*. Take for instance the following sentence pair (not derived from the SBAE corpus).

Speaker A. She has not done any homework, has she?

Speaker B. No.

Despite the fact that Speaker B uses the marker *no*, he seems to agree with Speaker A that she has not done any homework. Similarly, context and pragmatics can bias the use of a discourse marker. For instance, *maybe* in the above example could be considered a negative relation that is used to avoid a face threatening situation. Our analysis however shows that these three categories apply to the majority of the SBSAE markers in their context.

In addition to the polarity category, the clustering of markers shows an additional category, which we will call **acknowledgment**. This third category emerged by splitting each factor into two equal groups.

ACKNOWLEDGMENT

1. ACKNOWLEDGMENT relations fully acknowledge the implicature of the utterance expressed by the speaker.

2. PART-ACKNOWLEDGE relations acknowledge the general implicature of the speech act expressed by the speaker, but add some comments to it

ACCEPT

REBECCA: um, they're all different ages, and,
.. um, they, ... you know, for the most part,
they were probably very nervous when they came in just for jury duty.

RICKIE: **Yeah**.

REBECCA: So,

RICKIE: . . .

REBECCA: um,

RICKIE: . . .

REBECCA: .. they just want to hear your story, and they're not .. judging you.

RICKIE: **Okay**.

REBECCA: .. Um,
.. let me .. show you

PARTIAL ACCEPT

MILES: ... But I'm sure those guys get a lot of attention from women. They travel all over the world. .. I'm sure a lot of women throw themselves at em, so that's what they expect from women.

HAROLD: ... A lot of groupies.

MILES: .. Yeah ... Fringe benefits.

PETE: Hm.

MILES: ..

JAMIE: .. The only fringe benefit of being a dancer, probably.

HAROLD: .. Well in San Francisco,

MILES: **But then again,**

HAROLD: .. you never know if it's a benefit or not,

either.

JAMIE: That's true.

PETE: . . .

In the above example *yeah* and *okay* fully acknowledge the implicature of Rebecca. In the second example on the other hand, Miles *but then again* marks that he does not fully acknowledge Jamie's claim that lots of groupies are the only fringe benefit of being a dancer.

Markers in the acknowledgment category determine whether or not the speaker has some reservations with what has been said. It is no surprise that the ACKNOWLEDGEMENT category is strongly linked to the POLARITY category. In fact, most existing classifications of coherence relations have the polarity and acknowledgment categories amalgamated into one category. The factor analysis however gives strong evidence discriminating between these categories.

The final category that can be derived from the factor analysis is **emphatics**. This category was formed by ordering the factor scores, with highest scores forming emphatic and lowest scores forming non-emphatic relations.

EMPHATICS

1. EMPHATIC relations emphasize the animation toward the implicature of the utterance expressed by the speaker;
2. NON-EMPHATIC relations are dispassionate about the implicature of the utterance expressed by the speaker.

EMPHATIC

MONTOYA: I mean uh, if .. one looks at what,

uh, Jesse Jackson is doing, vis-a-vis .. who.

... The major league?... Baseball teams and all that?

.. Football and all that? …What's his underlying argument.

.. What's his criticism.... That there're not sufficient numbers of what?

FRANK: ... Blacks.

MONTOYA: .. Well he says minorities... He's smart. He's talks about minorities.

But he's really talking about African-Americans.

FRANK: ... In the uh,.. managerial process of,

MONTOYA: Right.

FRANK: .. of .. professional sports.

MONTOYA: **Exactly**.

... Managerial,... owners,... managers, et cetera. right?

.. Alright. Have you heard of a Latino say this?

... Or would Jesse Jackson have said that about Latinos?

... Hm? ... No.... Unh-unh.

NON-EMPHATIC

MONTOYA: ... But why is it important,
... in a democracy for the people to participate.
... Just deal with the question... Not the methods,
CAROLYN: .. You mean because .. if ... the people don't vote,
.. it will be, ... there will be an elite group,
it'll be .. more of a .. monarchy,... that has control,
... it won't be a democracy.
MONTOYA: **Okay,**
what's a monarchy.
RUBEN: ... Kings,
FRANK: One ruler,

The emphatics category marks to which extent the speaker is passionate about the previous utterance. The hearer can emphatically mark agreement with the speaker. For instance in the above example Montoya does not just agree with Frank, but wants to emphasize this agreement by using an emphatic marker. On the other hand, in the second example Montoya marks agreement, but does that in a rather non-passionate way.
The emphatics category has been presented here as a binary category. However, the analysis instead showed a gradation of emphatics, with some markers being more emphatic than others.

The four categories that emerged from the principal component analysis generate a total of twelve combinatory categories (see Table 4). Except for two categories these relations are all realized in English discourse markers. NEUTRAL PART-ACKNOWLEDGE EMPHATIC relations and NEGATIVE PART-ACKNOWLEDGE EMPHATIC relations are not. An easy pragmatic explanation can be given for the absence of

markers in these categories: When a participant in general accepts the implicature, but is not passionate about it and remains ambiguous about whether or not to follow up on this implicature, the dialog participants do not have to be cued. Similarly, a participant who disagrees will express this in an emphatic way and fully rejects the implicature.

TABLE 4 ABOUT HERE

By substituting the different discourse markers in the SBSAE corpus, we constructed a large network of relations. By running a principal component analysis over the relations between the markers we found four categories that emerged: direction, polarity, acknowledgment and emphatics.

Mapping the taxonomy onto the discourse model

The remaining question is how the categories of the taxonomy can be related to the discourse model that was introduced earlier. That model consisted of an action ladder with four levels: the first level coordinates attention in an information framework, the second signal identification in the exchange structure, the third comprehension coordination in the ideational structure and the fourth project coordination in the participation framework. Important in this amalgamation of Clark's and Schiffrin's model are the notions of upward completion and downward evidence. Though all four levels are co-temporal, lower levels enable higher levels, while higher levels provide evidence for the completion of lower levels. If the categories that emerged from the principal component analysis are to fit in the discourse model, they should be depedent on each other. The comparison of the discourse model and the proposed taxonomy is presented in Table 5.

TABLE 5 ABOUT HERE

Earlier we argued that the presence of a discourse marker in itself executes the speaker's behavior to point out a transitional relation. This would be a function at the first level in Clark's action ladder and Schiffrin's information framework. At this level interactive transitions in shared knowledge are coordinated. The speaker marks that she has certain information available, or marks a focus of the speaker's attention. In addition to the presence of the discourse marker, the category of emphatics seems to play a role at this level. The cognitive information states and speaker's stance is coordinated at this level. Does the speaker feel strongly about the topic or not? Given that emphatics is not only communicated through semantics but also (and perhaps primarily) through intonation, the tone of the speaker's voice without any further decoding from the side of the hearer further suggests that the emphatic action is completed prior to the other three levels.

When the hearer attends to the speaker's information turn-taking has to be coordinated. The speaker's signal needs to be identified by the hearer and the speaker should make clear what she plans to do in the exchange. Schiffrin's exchange structure maps onto the second level of Clark's action ladder where the presentation and identification of the signal takes place. In the alternation of speaker and hearer roles the speaker needs to point out whether he wants to take the turn, wants to keep the turn, wants to follows up on something said earlier, wants to initiate something new or a combination of these options. The direction category we found in our analysis applies to this level. By marking the turn with a backward or forward relation, the direction of the conversation is signaled. In the choice of the discourse marker the speaker does not only mark that he wants to take the turn, but also marks whether he wants to keep the turn (as necessarily the case in the use of forward relations).

Once attention is communicated and speaker and hearer know who is taking a turn and want to keep the turn, meaning can be coordinated at the third level in the ideational structure. At this level the relations between ideas of speaker and hearer is communicated in language. The category that emerged from our analysis is polarity in which the speaker agrees, disagrees or is neutral about the previous utterance. Polarity should certainly not be considered the only feasible category. It seems likely that at least relations like causal, temporal and additive, as found in other taxonomies, play a role. However, only the category of polarity emerged from the PCA analysis.

The fourth and final level in the action ladder matches Schiffrin's participation framework. At this level the accomplishment of the interactional task plays a central role. Speaker's prior ideas and intentions can be altered and digressions can be signaled (e.g. through indirect answers). At this level reservations the speaker has is coordinated. The category of acknowledgement can therefore be corresponds to this level. The speaker can either fully acknowledge the previous utterance or express reservations, depending on the flow of the conversation (does the conversation allow for hesitations?) and the status of speaker and hearer (one does not want to fully reject a previous utterance from one's employer).

The overview in Table 5 shows the relations between the four discourse planes and the categories in the proposed taxonomy. The use of discourse markers and the emphatics of their use result in the hearer's attention in the information framework. The speaker needs to ask the question "Do I need the hearer's attention" in order to decide on the use of a discourse marker and its emphatics. If a marker is used, particularly if an emphatic marker is used, the hearer is cued to pay attention. This attention is needed to communicate the direction of the relation that is marked in the exchange structure. The speaker needs to consider the question "Do I want to say more?" in order to decide on backward or forward relations. The hearer is cued whether the speaker wants to take and keep the turn. Coordination of attention and direction enable the coordination of the polarity of the discourse marker in the ideational structure. The speaker needs to consider whether he accepts or rejects the previous utterance or takes a neutral stance. In this ideational

structure, the hearer is informed about whether the speaker is positive, negative or neutral. Based on the speaker's role with respect to the hearer and the conversation, the speaker can express reservations in accepting or rejecting the previous utterance. Based on the speaker's role in the acknowledgment the hearer can reconsider the participation roles in the conversation. As Clark (1996: 154) points out, these may not be the only levels, but according to our analysis they turn out to be the major ones.

Discussion

Discourse markers are more frequent in dialog than in monolog because of the various levels they operate on and can be coordinated at. Finding order in the large set of markers is difficult. We have conducted a substitution test to determine the relations between the markers and used a principal component analysis to derive categories from these relations. We thereby focused on single, local, between-turn, primarily backward discourse markers. While constraining the study in order to make it manageable, this focus also brings caveats.

A first problem with the current study is the definition of discourse markers. We have included words and phrases that other researchers would not consider a discourse marker, but instead an acknowledgment token, interjection or conjunction. We have purposely tried to include a large range of words that potentially mark the discourse under the assumption that the substitution test and subsequent analyses would filter any debris. No evidence came forward that showed that clustered a category like conjunctions from other categories, suggesting that substitutions can be made across syntactic and morphological categories.

We have not looked at relations between discourse units larger than turns (e.g. groups of turns). The consequence of this is that the substitutions we have made are at a local level only. Bangerter & Clark (forthcoming) have pointed out that some markers are often used in specific points of a dialog. For instance, the marker *okay* closes a joint subproject, a marker like *hmm* instead marks the process of such a subproject,

but does not close it. It is therefore important to point out that our substitutions do not necessarily carry over to global relations. That is, while two discourse markers might be mutually substitutable at a local level, they might not be when the flow of the conversation is taken into consideration. For instance, *right* and *mhm* are synonymous according to the substitution test, but it seems plausible that *mhm* marks a series of exchanges on a subtopic that is closed with the marker *right*. Marking the interactive process with *right* and closing the subtopic with *mhm* seems less intuitive.

In this study we have focused on between-turn markers occurring at the beginning of a turn. Markers that might have been used to relate two turns, but did not occur at the beginning of the turn have been excluded from the analysis. In a similar vein, within-turn markers that occur at the beginning of the turn have been included in the analysis. This led to a second selection, one that distinguished between PUSH and RETURN relations. Our taxonomy has primarily focused on RETURN relations. The principal component analysis excluded discourse markers like greetings (*hi*, *goodbye*) and thanks (*thanks*) were excluded from the further analysis. Similarly, markers like *because* and *before*, despite that they occurred at the start of the turn were excluded. Clearly, this does not mean that forward relations should not be considered. Instead it means that return relations were considered because they turned out to have the largest shared variance. The categories in the proposed taxonomy are therefore primarily relevant for the subset of return relations.

Despite the fact that the substation test proves to be fruitful in building a network of relations between discourse markers and even to derive major categories from this network, it eliminates some valuable information that is particularly important for dialog. For instance, no information is included about intonation. This has some important consequences. The intonation of a marker like *yes* could for instance determine whether it would be interpreted as CONTINUOUS or DISCONTINUOUS. Instead the substitution test has merely looked at semantic relationships. Furthermore, the semantic relations are investigated with a limited context. Elaborate context can override the meaning of discourse markers, for instance in the example we referred to earlier where the negative marker no can actually mark a positive relation.

Despite the limitations of scope (single, local, between-turn, primarily return relations), context (limited context with local discourse markers that cue adjacent turns) and linguistics (semantics only, without considering phonetic, pragmatic, and sociological aspects), a taxonomy has emerged that seems plausible. Its plausibility lies in the various relations between the markers, the agreement among raters on these relations, the categories emerging from these relations, the dependency between the categories, and the mapping of these categories onto a discourse model.

Conclusion

This study aimed at developing a theoretically constrained taxonomy of discourse markers grounded in a corpus. Discourse markers are the conversational glue and therefore play a very prominent role in discourse. A corpus linguistic study we conducted showed that they are particularly frequent in informal spoken dialog. The explanation is that discourse markers in this language setting operate at various different levels in the conversation. At least four of these levels can be distinguished and there is evidence that these levels can be hierarchically ordered, lower levels enabling higher levels. The taxonomy we have proposed can be mapped onto each of these levels. A substitution analysis we conducted generated a network of relations between discourse markers. The four categories that emerged from a data reduction analysis on this network mapped onto the four discourse levels. The categories show the interdependency between them and the coordination that takes place between speaker and hearer. They also illustrate how discourse markers operate at different levels in the conversation. Most importantly, they provide us with a further insight in language and its uses.

Acknowledgements

This research was supported by grants from DoD Multidisciplinary University Research Initiative (MURI) program administered by the Office of Naval Research (N00014-00-1-0600) and the National Science

Foundation (SBR 9720314 and REC0106965). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of ONR or NSF.

References

Bangerter, A. & Clark, H. H. (forthcoming). Navigating joint projects with dialogue.

Biber, D. (1988). *Variation across speech and writing*. Cambridge, etc. Cambridge University Press.

Carletta, J. , Isard, A. , Isard, S. , Kowtko, J. , Doherty-Sneddon, G. (1996). HCRC dialogue structure coding manual. Technical Report HCRC/TR-82.

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

Comrey, A. L. & Lee, H. B. (1992). A first course in factor analysis. Hillsdale, NJ, Lawrence Erlbaum.

Core, M.G. & Allen, J. (1997). *Draft of DAMSL: Dialog Act Markup in Several Layers*. Unpublished Manuscript.

Di Eugenio, B., Jordan, P.W. & Pylkkänen, L. (1998). *The COCONUT project: dialogue annotation manual*, ISP Technical Report 98-1.

Gernsbacher, M.A. (1994). *Handbook of Psycholinguistics*, New York, etc.: Academic Press.

Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York, Academic Press.

Graesser, A.C., Gernsbacher, M.A. and Goldman, S.A. (Eds.) (in press). *Handbook of Discourse Processes*. Hillsdale, NJ: Erlbaum.

Grosz, B. J. & Sidner, C. L. (1986). Attention, intentions and the structure of discourse. *Computational Linguistics 12* (3), 175-204.

Halliday, M.A.K. and Hasan, R. (1976). *Cohesion in English*. London, Longman.

Hirschberg, J. & Litman, D. (1993). Empirical studies on the disambiguation of cue phrases. *Computational Linguistics, 19* (3), 501-530.

Hobbs, J.R. (1985). *On the coherence and structure of discourse*. Technical Report CSLI-85-37, Center for the Study of Language and Information, Stanford University.

Jefferson, G. (1984). Notes on a systematic deployment of the acknowledgement tokens "yeah" and "mm hm". *Papers in Linguistics*, 17, 197-216.

Knott, A. (1996). *A data-driven methodology for motivating a set of coherence relations*. Unpublished PhD thesis, University of Edinburgh.

Knott, A. and Dale, R. (1994). Using linguistic phenomena to motivate a set of coherence relations. *Discourse Processes* 18(1): 35-62.

Knott, A. & Mellish , C. (1996). A feature-based account of the relations signalled by sentence and clause connectives. *Language and Speech 39* (2-3), 143-183.

Levinson (1983). *Pragmatics*. Cambridge, etc. Cambridge University Press.

Louwerse, M.M. & Graesser, A.C. (in press). Coherence in discourse. In Strazny, P. (ed.), *Encyclopedia of Linguistics*. Chicago, Fitzroy Dearborn.

Louwerse, M. M. (2002a). An analytic and cognitive parameterization of coherence relations. *Cognitive Linguistics*, pp. 291-315.

Mann, W.C. (1988). Dialogue games: Conventions of human interaction. *Argumentation, 2*, 511-532.

Mann, W.C. & Thompson, S.A. (1987*). Rhetorical structure theory: a theory of text organization*. Technical Report ISI/RS-87-190, Information Science Institute, University of Southern California.

Marcu, D. (2000). *The Theory and Practice of Discourse Parsing and Summarization.* Cambridge, MA, MIT Press.

Martin, J.R. (1992). *English text. System and structure*. Amsterdam/Philadelphia, John Benjamins.

Oates, S. L. (2000). Multiple Discourse Marker Occurrence: Creating Hierarchies for Natural Language Generation. Proceedings of the ANLP-NAACL Student Research Workshop, ANLP-NAACL 2000

Poesio, M. and Traum, D. (1998) Towards an Axiomatisation of Dialogue Acts. In J. Hulstijn and A. Nijholt (eds.), *Proceedings of the Twente Workshop on the Formal Semantics and Pragmatics of Dialogues*, May 1998.

Polanyi, L. and Scha, R. (1983). The syntax of discourse. In: J. Hobbs (ed.), *Formal Methods in Discourse Analysis. Text 3* (3), pp. 261-270.

Pritchett, B. L. (1992). *Grammatical competence and parsing performance*. Chicago/London, University of Chicago Press.

Sanders, T.J., Spooren, W.P.M. and L.G.M. Noordman (1992). Toward a taxonomy of coherence relations. *Discourse Processes 15*, 1-35.

Sanders, T.J., W.P.M. Spooren and L.G.M. Noordman (1993) Coherence relations in a cognitive theory of discourse representation. *Cognitive Linguistics 4* (2), 93-133.

Schegloff, E. A. (1996). Turn organization: one intersection of grammar and interaction. In E. Ochs, E. A. Schlegloff & S. A. Thompson, *Interaction and grammar* (pp. 52-133). Cambridge, etc. Cambridge University Press.

Schiffrin, D. (1987). *Discourse markers*. Cambridge, etc. Cambridge University Press.

Schlegloff, E. A., Ochs, E. & Thompson, S. A. (1996). Introduction. In E. Ochs, E. A. Schlegloff & S. A. Thompson, *Interaction and grammar* (pp. 1-51). Cambridge, etc. Cambirdge University Press.

Schlegloff, E.A. & Sacks, E.A. (1973). Opening up closings. *Semiotica, 8*, 289-327.

Swan, M. (1996). *Practical English usuage*. Oxford, etc., Oxford University Press.

Traxler, M. J., A. J. Sanford, J. P. Aked, and L. M. Moxey (1997). Processing causal and diagnostic statements in discourse. *Journal of Experimental Psychology 23* (1), 88-101.

Van Dijk, T.A. (1997). *Discourse as structure and process. Discourse studies 1. A multidisciplinary introduction*. London, etc.: Sage.

Webber, B. L. & Joshi, A. K. (1998). Anchoring a lexicalized tree-adjoining grammar for discourse. In M. Stede, L. Warner and E. Hovy (Eds.). *Discourse relations and discourse markers. Proceedings from the Workshop. COLING-ACL'98* (pp. 86-92). Montreal.

Whittaker, S. (in press). Mediated communication. In A.C. Graesser, M.A. Gernsbacher, and S.A. Goldman, in press.

Yngve, V. H. (1970). On getting a word in edgewise. *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society*, 6, 657-677.

Table 1

Three categories of six meta-genres derived from genres in Biber (1988).

1	**Spoken**	Spontaneous speeches (16), planned speeches (14), face-to-face conversation (44), telephone conversation (27), public conversations, debates and interviews (22), broadcasts (18)
	Written	academic prose (80), official documents (14), fiction (74), personal letters (6), professional letters (10)
2	**Dialog**	face-to-face conversation (44), telephone conversation (27), public conversations, debates and interviews (22), broadcasts (18), personal letters (6), professional letters (10)
	Monolog	Spontaneous speeches (16), Planned speeches (14), academic prose (80), official documents (14)
3	**Informal**	Spontaneous speeches (16), face-to-face conversation (44), telephone conversation (27), fiction (74), personal letters (6)
	Formal	Planned speeches (14), public conversations, debates and interviews (22), broadcasts (18), fiction (74), academic prose (80), official documents (14), professional letters (10)

Number of texts in genre between brackets.

Table 2

Means and standard deviations of the five linguistic features in six meta-genres. Derived from Biber (1988: 247-269).

Genre	**No. texts**	**causal**	**concessive**	**conditional**	**rest group**	**particles**	**Mean cohesion relations**
Dialog	198	2.47 (1.87)	0.58 (0.92)	3.15 (2.47)	0.93 (1.20)	2.83 (1.97)	1.99 (1.68)
monolog	127	1.08 (1.16)	0.37 (0.62)	2.27 (1.69)	1.04 (1.15)	1.24 (1.08)	1.20 (1.14)
Spoken	141	2.32 (1.67)	0.25 (0.53)	3.15 (2.08)	0.68 (0.93)	3.60 (2.40)	2.00 (1.52)
Written	184	1.26 (1.40)	0.77 (1.08)	2.27 (2.15)	1.34 (1.47)	0.32 (0.56)	1.19 (1.33)
Informal	167	2.44 (1.54)	0.63 (0.78)	3.57 (2.69)	0.90 (1.15)	3.10 (2.02)	2.13 (1.64)
Formal	158	1.33 (1.55)	0.37 (0.78)	2.07 (1.63)	1.05 (1.20)	1.28 (1.18)	1.22 (1.27)
total/mean	975	1.82 (1.53)	0.49 (0.79)	2.75 (2.12)	0.99 (1.18)	2.06 (1.53)	

causal = *because*

concessive = *although*, *though*

conditional = *if*, *unless*

rest group = *Since*, *while*, *whilst*, *whereupon*, *whereas*, *whereby*, *such that*, *so that*, *such that*, *inasmuch as, forasmuch as, insofar* as*, insomuch as, as long as*, *as soon as*

particles = *well*, *now*, *anyway*, *anyhow*, *anyways*

Table 3.

A comparison of Clark's action ladder and Schiffrin's discourse model

Clark	**Schiffrin**
LEVEL 4: PROPOSAL and CONSIDERATION	PARTICIPATION FRAMEWORK
LEVEL 3: LEVEL MEANING and UNDERSTANDING	IDEATIONAL STRUCTURE
LEVEL 2: PRESENTATION & IDENTIFICATION	EXCHANGE STRUCTURE
LEVEL 1: EXECUTION and ATTENTION	INFORMATION FRAMEWORK

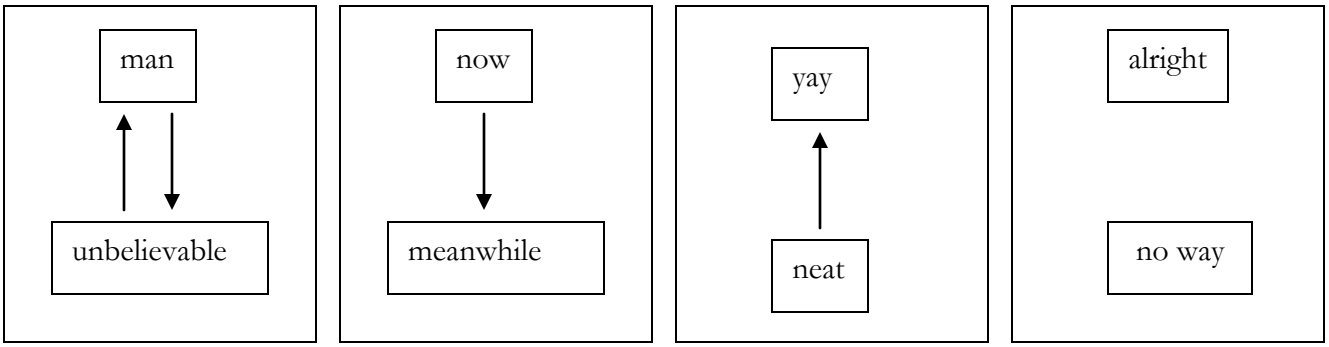
Table 4. Combinatory categories

Turn	**Direction**	**Continuity**	**Acknowledgment**	**Emphatics**	**Realization**	**Example**
Within-turns	Forward	Positive	n/a	n/a	+	
		Negative	n/a	n/a	+	
	Backward	Positive	n/a	n/a	+	
		Negative	n/a	n/a	+	
Between-turns	Forward	n/a	n/a	n/a	+	
	Backward	Positive	Acknowledge	Emphatic	+	
				Non-Emphatic	+	
			Part-Acknowledge	Emphatic	+	
				Non-Emphatic	+	
		Neutral	Acknowledge	Emphatic	-	
				Non-Emphatic	+	
			Part-Acknowledge	Emphatic	+	
				Non-Emphatic	+	
		Negative	Acknowledge	Emphatic	-	
				Non-Emphatic	+	
			Part-Acknowledge	Emphatic	+	
				Non-Emphatic	+	

Table 5.

A comparison of Clark's action ladder and Schiffrin's discourse model

Clark	**Schiffrin**	**Taxonomy**
LEVEL 4: PROPOSAL & CONSIDERATION	PARTICIPATION FRAMEWORK	ACKNOWLEDGMENT
LEVEL 3: MEANING & UNDERSTANDING	IDEATIONAL STRUCTURE	POLARITY
LEVEL 2: PRESENTATION & IDENTIFICATION	EXCHANGE STRUCTURE	DIRECTION
LEVEL 1: EXECUTION & ATTENTION	INFORMATION FRAMEWORK	PRESENCE OF MARKER + EMPHATICS



Synonym Hyponym Hypernym Exclusive

Figure 1. Relations in the substitution test.

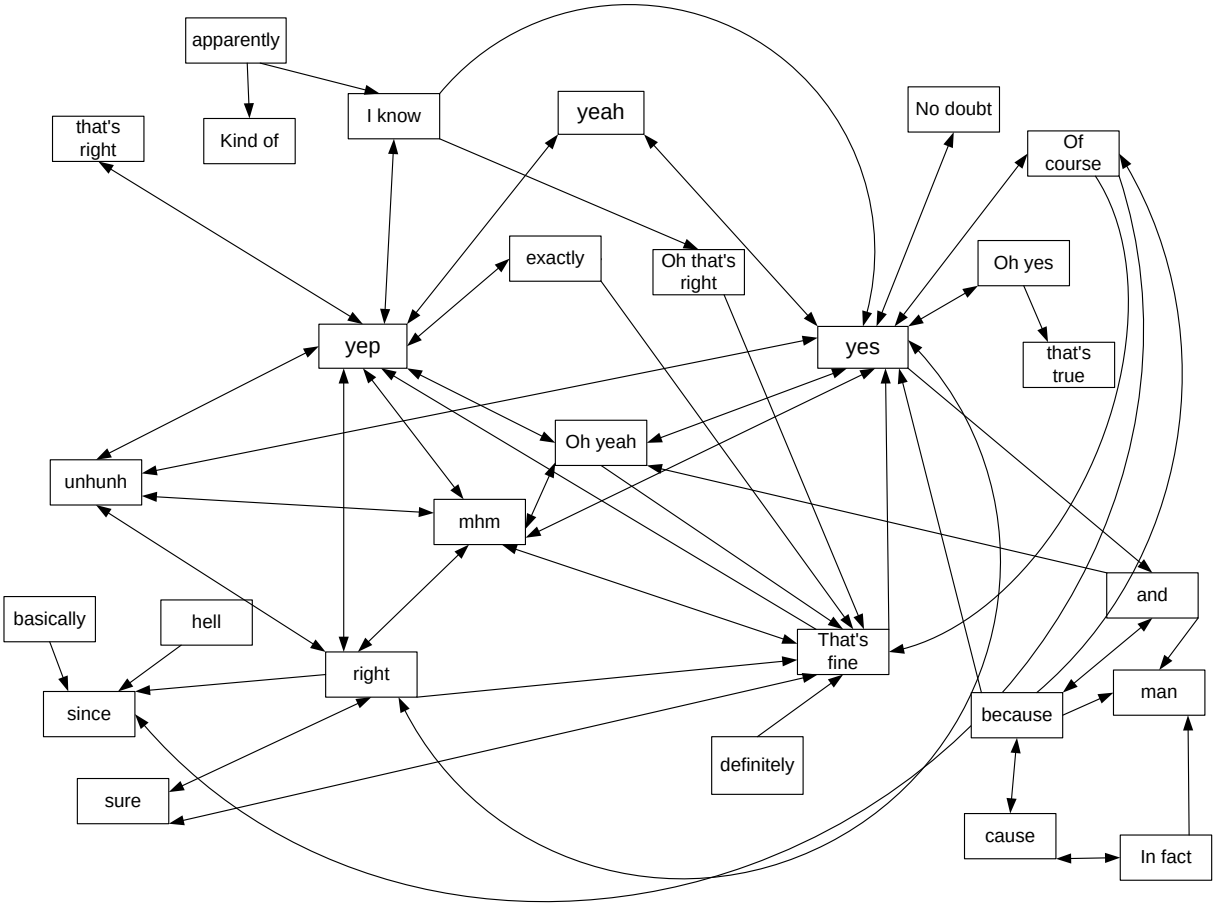


Figure 2. Sample overview of network of discourse markers

Appendix 1

Results PCA analysis

Discourse marker	CONTINUITY	ACKNOWLEDGMENT	EMPHATICS
alright	CONTINUOUS	ACKNOWLEDGE	EMPHATIC
and	CONTINUOUS	PART-ACKNOWLEDGE	NON-EMPHATIC
anyway	CONTINUOUS	PART-ACKNOWLEDGE	NON-EMPHATIC
apparently	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
as	CONTINUOUS	PART-ACKNOWLEDGE	NON-EMPHATIC
aw	DISCONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
aye	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
boo	DISCONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
but	CONTINUOUS	PART-ACKNOWLEDGE	NON-EMPHATIC
but then	NEUTRAL	PART-ACKNOWLEDGE	EMPHATIC
but then again	CONTINUOUS	PART-ACKNOWLEDGE	NON-EMPHATIC
by virtue of	CONTINUOUS	PART-ACKNOWLEDGE	NON-EMPHATIC
cause	CONTINUOUS	PART-ACKNOWLEDGE	EMPHATIC
cool	CONTINUOUS	PART-ACKNOWLEDGE	NON-EMPHATIC
course	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
dang	DISCONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
definitely	CONTINUOUS	ACKNOWLEDGE	EMPHATIC
disgusting	CONTINUOUS	ACKNOWLEDGE	EMPHATIC
eek	DISCONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC

every time	NEUTRAL	PART-ACKNOWLEDGE	EMPHATIC
ew	DISCONTINUOUS	ACKNOWLEDGE	EMPHATIC
exactly	CONTINUOUS	ACKNOWLEDGE	EMPHATIC
except	CONTINUOUS	PART-ACKNOWLEDGE	EMPHATIC
first	CONTINUOUS	PART-ACKNOWLEDGE	NON-EMPHATIC
gee	DISCONTINUOUS	ACKNOWLEDGE	EMPHATIC
good	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
guess	CONTINUOUS	PART-ACKNOWLEDGE	NON-EMPHATIC
hm	CONTINUOUS	PART-ACKNOWLEDGE	NON-EMPHATIC
how bizarre	DISCONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
hunh	DISCONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
hunh-unh	DISCONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
i bet	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
i guess	NEUTRAL	ACKNOWLEDGE	NON-EMPHATIC
i know	CONTINUOUS	ACKNOWLEDGE	EMPHATIC
i mean	CONTINUOUS	PART-ACKNOWLEDGE	NON-EMPHATIC
i'll bet	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
in order to	CONTINUOUS	PART-ACKNOWLEDGE	NON-EMPHATIC
jeez	DISCONTINUOUS	ACKNOWLEDGE	EMPHATIC
just	NEUTRAL	ACKNOWLEDGE	NON-EMPHATIC
kay	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
kind of	NEUTRAL	ACKNOWLEDGE	EMPHATIC
let me see	NEUTRAL	ACKNOWLEDGE	NON-EMPHATIC
like	CONTINUOUS	PART-ACKNOWLEDGE	NON-EMPHATIC

look like	NEUTRAL	ACKNOWLEDGE	NON-EMPHATIC
looks like	NEUTRAL	ACKNOWLEDGE	NON-EMPHATIC
man	DISCONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
maybe	NEUTRAL	ACKNOWLEDGE	EMPHATIC
meanwhile	CONTINUOUS	PART-ACKNOWLEDGE	NON-EMPHATIC
mhm	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
n	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
na	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
nah	DISCONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
neat	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
nkay	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
no	DISCONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
no doubt	CONTINUOUS	ACKNOWLEDGE	EMPHATIC
nope	DISCONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
now	CONTINUOUS	PART-ACKNOWLEDGE	NON-EMPHATIC
of course	CONTINUOUS	ACKNOWLEDGE	EMPHATIC
oh alright	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
oh boy	DISCONTINUOUS	ACKNOWLEDGE	EMPHATIC
oh dear	DISCONTINUOUS	ACKNOWLEDGE	EMPHATIC
oh god	DISCONTINUOUS	ACKNOWLEDGE	EMPHATIC
oh good	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
oh gosh	DISCONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
oh great	DISCONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
oh gross	DISCONTINUOUS	ACKNOWLEDGE	EMPHATIC

oh my	DISCONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
oh no	DISCONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
oh okay	CONTINUOUS	PART-ACKNOWLEDGE	NON-EMPHATIC
oh shit	DISCONTINUOUS	ACKNOWLEDGE	EMPHATIC
oh that's right	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
oh yeah	CONTINUOUS	ACKNOWLEDGE	EMPHATIC
oh yes	CONTINUOUS	ACKNOWLEDGE	EMPHATIC
okay	CONTINUOUS	PART-ACKNOWLEDGE	NON-EMPHATIC
on the other hand	NEUTRAL	PART-ACKNOWLEDGE	EMPHATIC
oo	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
ooo	DISCONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
possibly	NEUTRAL	ACKNOWLEDGE	EMPHATIC
probably	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
really	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
right	CONTINUOUS	ACKNOWLEDGE	EMPHATIC
shit	DISCONTINUOUS	ACKNOWLEDGE	EMPHATIC
so	CONTINUOUS	PART-ACKNOWLEDGE	NON-EMPHATIC
sort of	NEUTRAL	ACKNOWLEDGE	NON-EMPHATIC
sure	CONTINUOUS	ACKNOWLEDGE	EMPHATIC
that too	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
that’s fine	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
that’s right	CONTINUOUS	ACKNOWLEDGE	EMPHATIC
there you have it	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
too bad	DISCONTINUOUS	ACKNOWLEDGE	EMPHATIC

tough	DISCONTINUOUS	ACKNOWLEDGE	EMPHATIC
ugh	DISCONTINUOUS	ACKNOWLEDGE	EMPHATIC
um	CONTINUOUS	PART-ACKNOWLEDGE	NON-EMPHATIC
unbelievable	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
unhunh	CONTINUOUS	ACKNOWLEDGE	EMPHATIC
well	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
whoa	DISCONTINUOUS	ACKNOWLEDGE	EMPHATIC
woo	DISCONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
wow	CONTINUOUS	ACKNOWLEDGE	EMPHATIC
yeah	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
yep	CONTINUOUS	ACKNOWLEDGE	EMPHATIC
yes	CONTINUOUS	ACKNOWLEDGE	EMPHATIC
you bet	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC
you know	CONTINUOUS	PART-ACKNOWLEDGE	NON-EMPHATIC
yuck	DISCONTINUOUS	ACKNOWLEDGE	EMPHATIC
yup	CONTINUOUS	ACKNOWLEDGE	NON-EMPHATIC

Appendix 2

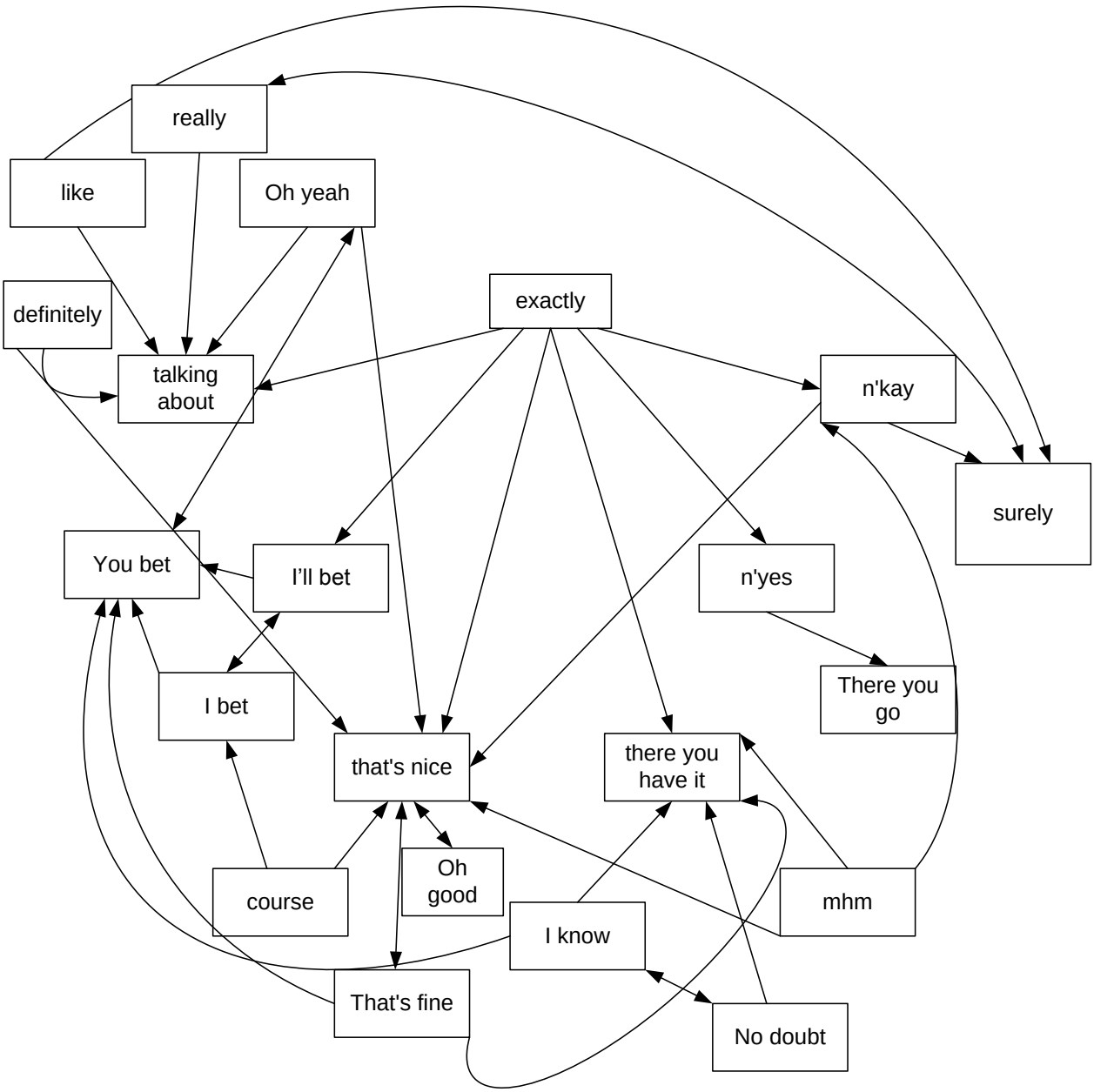


Fig.2. Continuous relations

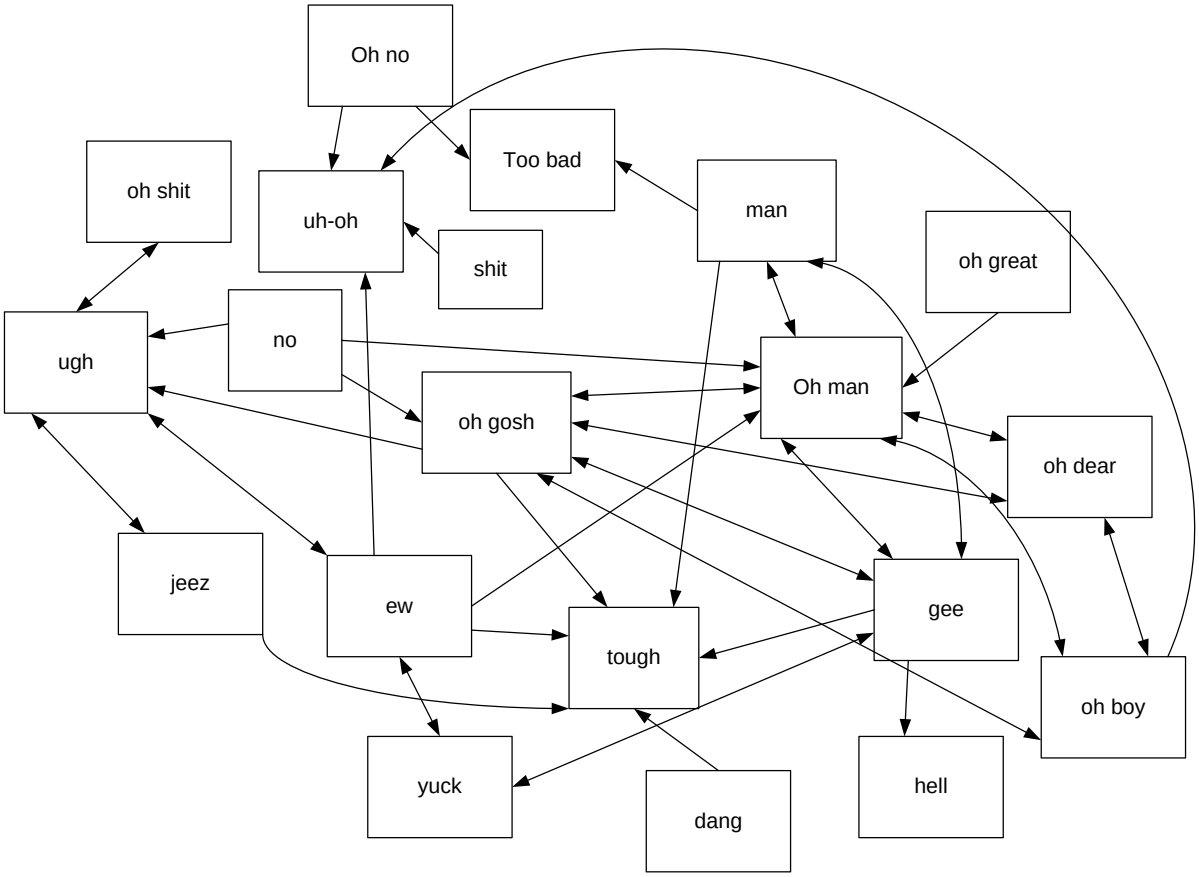


Fig.3. Discontinuous relations

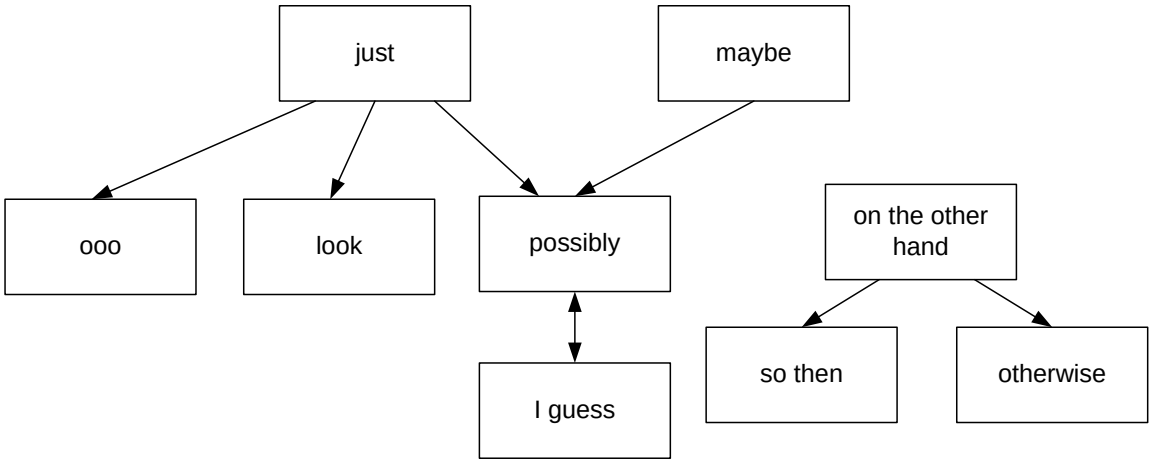


Fig. 4. Neutral relations