INTERNATIONAL JOURNAL OF SPEECH TECHNOLOGY 4, 31–44, 2001

Using Sequence Package Analysis to Improve Natural Language Understanding

AMY NEUSTEIN
Linguistic Technology Systems, 135 East 54th Street, Suite: 7J, New York, NY 10022, USA
lingtec@banet.net

Received December 29, 1999; Revised April 20, 2000

Abstract. Developers of dialogue systems must confront the complexities of natural language. The purpose of this paper is to demonstrate how "sequence package" analysis, as a novel approach, can help to improve natural language understanding. Such an approach would go beyond the standard grammatical formalisms represented in most dialogue systems, to include context-dependent utterance sequences that are shaped by the unfolding talk. What is then comprised in a sequence package is a series of related turn construction units and turns that make up either single or multiple episodes of talk, and sometimes an entire conversation. The author examines help-related dialogue to show how reports of troubles that often appear ambiguous and vague can be better understood by looking at the sequential design of speakers' turn constructions. Subtle features found in troubles-related talk that are important, but often overlooked, may be identified by mapping out the sequence package arrangement of the talk. For example, a caller's need for vital empathic support, before he or she can be ready to receive help, might be hard to detect if the caller only provides hidden, and possibly contradictory, signs of emotional distress. Or a patient might be unclear and somewhat inconsistent when trying to describe his or her chief complaint in the course of a medical interview. Thus, an analysis of sequence packages can potentially uncover crucial information often buried in the talk. In designing dialogue systems that model spontaneous speech, a sequence package analysis might serve as a basic component of natural language systems.

Keywords: speech interface, natural language processing, conversation analysis, human-computer interaction, customer relationship management

1. Introduction

Developers of speech interface programs incorporating Automatic Speech Recognition (ASR) and Natural Language Processing (NLP) commonly face three practical problems: (1) "in depth natural language processing is an expensive endeavor that can strain computational resources" (Riloff and Lehnert, 1994, p. 300); (2) "the human component of naturalness for spoken interaction is still poorly understood" (Markowitz, 1996, p. 260); and (3) spoken language is replete with nonliteral, metaphorical descriptors—idioms, proverbs, aphorisms, colloquialisms, argot and slang (Machowski, 1997; Riloff and Lehnert, 1994). Genuinely interactive speech-based computer interfaces, designed to simulate human dialogue—the true barometer of artificial intelligence—must take into account the rich context-dependent features of discourse, punctuated by ambiguities, ellipses (fragmented, incomplete utterances), metaphors and other vagaries of natural speech, including disfluencies. It is these vagaries—often considered impediments to developing truly intelligent systems—that form the subject domain of sequence package analysis, which is described below. The purpose of this paper is to illustrate how sequence package analysis can help to improve natural language understanding.

Before describing how sequence analysis works, a brief presentation of a few such applications follows.

1.2. Applications

In our service-centered economy, call centers are flooded with daily complaints, but there is a paucity of human agents to handle the large volume of calls. A speech interface design, operating as a kind of "smart bot," may be designed to replace help-line operators. Such a system may be designed to accomplish two important objectives: identifying the source of the trouble; and determining at what point in the talk the caller is actually ready to receive practical help and is finished conveying the emotional distress associated with the complaint. The latter is particularly important because all too frequently when help is offered prematurely, it is rejected. After this occurs, subsequent attempts to offer help may engender resistance. An example provided below, in the section on "Remedy versus Empathy," illustrates the sequence packages appearing in the talk when a person is ready to accept advice and those that appear when a person needs to complain.

Sequence package analysis may also be used for training purposes (Neustein, 1986a, b, 2000). In the section on "Premature Remedy," the author presents a case illustration of specific sequence packages in a medical interview that point to a resident's formulation of a diagnosis before the patient has adequately explained the problem. A program that can analyze actual tape recordings of doctor-patient interaction, pointing out these specific sequence packages where premature diagnostic formulations are found to occur, may help the doctor-in-training to determine "whether his history taking is effective or whether his conversational method automatically distorts the data" (von Raffler-Engel, 1989, p. xxviii). Such training tools may also be of use in call centers where customer service agents, operating under strict time constraints, may preempt callers' reports of trouble via a premature attempt at a solution. Because many of these calls are tape recorded for quality assurance purposes, the data are already available for analytic inspection by a program that would be designed to spot these critical sequence packages and then point them out to the agent.

The current shift to a networked economy creates a need for "intelligent software products that integrate computer interaction and natural language understanding to bring a human-like presence to the points of contact between your company and its customers, partners, suppliers and employees" (Gaudiano, 1999, p. 3). Voice-enabled e-commerce might benefit from improved natural language understanding, permitting customers to speak almost as naturally with a web site as they would with a human customer service agent. Customer Relationship Management (CRM), if fully automated, can allow customer service departments to reserve their human agents for those rare occurrences when a highly complex problem is presented. An automated system might be able to identify relevant sequence packages associated with highly problematic calls and then proceed in those instances to refer the caller to a human operator. Speech interfaces might also be used in health care settings to take initial histories from patients before they are seen by their providers, and can then proceed to triage the cases so as to make greater use of the providers' services.

2. Background

2.1. Defining Sequence Package Analysis

The sequence package approach to the use of natural language for speech interface design applications attempts to provide a structural framework to help analyze free text, which can be characteristically cumbersome and unwieldy. Sequence packaging, as it is proposed here, constitutes a novel concept introduced by this author (Neustein, 1984, 1999).[1] The algorithmic design of a sequence package is based on the following: (1) the structural *location* of a turn construction unit[2] within the turn itself and within the talk at large, rather than based on its pure, context-free linguistic form;[3] and (2) how turn construction units are discretely *packaged* as a sequence in naturally occurring talk. What is then constituted as an actual "sequence package" is a series of related turn construction units and turns that make up either single or multiple episodes of talk, sometimes spanning an entire conversation.

While the field of natural language processing has witnessed the design of generators and parsers that emulate the grammatical formalisms of natural language dialogue (Frolich and Luff, 1990; Gilbert et al., 1990), such systems have been limited to adjacency pair sequences: that is, pairs of utterances in which a given first part of a pair requires a particular second pair part, as in question/answer, summons/response, request/grant-denial, offer/acceptance-rejection, or greeting/return-greeting (Sacks, 1992b; Schegloff and Sacks, 1973).[4]

Sequence package analysis has a much wider scope. Indeed, it is unique in two separate ways. First, the turn construction units that constitute the sequence go well beyond grammatical formalisms of adjacency pair design, both in size and in complexity. For example, as illustrated in the doctor-patient dialogue discussed later on, a sequence package may consist of an alternating pattern of speakers' usage of contrasting qualifiers, such as "equivocality" versus "certainty," found throughout many sequences in the talk, rather than in one particular sequence of the talk (e.g., openings or topic initiation about a chief complaint). Such contrasting pair types (equivocality/certainty) are constituted in situ (within the local, situated context of the unfolding talk), and are seen throughout the progression of the talk. As a result, these contextually relevant, situatedly defined pair types, emerging in the talk as features of the collaborative, interpretive work of two speakers, can only be identified by algorithms designed to spot those pair types (or other sequence phenomena) peculiar to the dialogue. This is why parsers that only recognize standard adjacency pair design, rather than the more loosely organized, complex sequences that are indigenous to the talk, are unable to detect the patterned and orderly use of such features as *contrasting qualifiers*. Idioms, likewise, can be found to elude the recognition of speech interfaces that mainly identify standard grammatical formalisms because idioms, too, constitute indigenous features of talk.[5]

The second unique distinction of sequence package analysis is its ability to uncover obscure, yet highly critical, material. For example, as illustrated in the dialogue between a hotline operator and a caller, discussed later on, clues that indicate a caller's readiness to receive practical assistance, rather than empathic support, may be identified by mapping out the patterns of sequence packages associated with such readiness. Often, signs of preparedness to receive assistance are subtle and cloaked. Or a caller might still be locked into a state of ambivalence, and therefore not quite ready to hear advice that could be helpful in redressing the problem that he or she has articulated in the call. Torode (1995) examined some of the subtle, roundabout ways a caller to a consumer help-line, at the Office of Consumer Affairs in Dublin, approached potentially contentious issues. When designing natural language systems, it is, therefore, important to build into the system a structural framework consisting of a sequence package design that can identify the relevant meaning behind utterances that frequently appear obscure and vague.[6]

The difficulties entailed in designing a system to identify such intricate sequences that span broad segments of talk, or even an entire conversation, have understandably caused sociolinguists and software engineers to focus their attention on the narrow, closely ordered sequences of adjacency pairs, which are known to have a primarily standard structural configuration, facilitating easier recognition by chart parsers. This paper aims to illustrate how distinct sequence packages, albeit complex and intricate, might be mapped out over lengthy portions of dialogue that are not confined to standard adjacency pair turn constructions. This might serve as a crucial step in improving natural language understanding.

2.2. Data and Methods

Sequence package analysis is informed by the conversation analytic method of breaking down natural language communication into its primary units of analysis, sequences and turns within sequences (rather than isolated sentences or utterances), to study how participants in a dialogue organize their verbal interactions in both ordinary, mundane conversations and in institutional talk (e.g., news interviews, courtroom examination or doctor-patient interaction). For some thirty years, conversation analysts have been engaged in the detailed study of naturally occurring talk as a socially organized activity. Their goal is to identify and describe how participants in a dialogue systematically accomplish their interactive work (which includes making requests, eliciting information, reporting on troubles or correcting errors and repairing misunderstandings), while they are continually engaged in a process of making sense of the ongoing social activity. This is done by examining how speakers demonstrate, through the design of their speaking turns, their understanding and interpretation of each other's social actions, such as a speaker's noticeable failure to answer a question directed at him or her.

The methods used by conversation analysts to study talk-in-interaction[7] are strictly empirical. They entail recording and transcribing verbal interactions by using highly refined transcription symbols to identify both verbal components and various paralinguistic features, such as stress, pauses, gaps, overlaps, and restarts (Atkinson and Heritage, 1984, pp. ix–xvi). However, the glossary of transcription symbols is not finite; "the notations, widely used by conversation analysts, are a continually evolving set of symbols designed to capture

the interactional qualities and nuances of speech delivered in real-time" (Firth, 1995, p. xiii).

2.3. Syntax vis-à-vis Sequence

In conversation analytic studies, syntax is viewed as being constituted by the sequential organization of talk, rather than consisting of a context-free grammatical structure. For example, in an adjacency pair sequence, such as an offer/acceptance-rejection, a personal declaration—"I'm leaving now"—is not construed as a declarative statement, but is, instead, heard as an "offer" to give someone a lift. Consequently, a personal state disclosure—"I'm all right as I am"—appearing as the second pair part to an offer sequence, can be heard as a "rejection" of an offer, rather than a statement about one's personal state (Coulter, 1979, p. 168). And in some sequences in talk, particularly arguments or heated discussions, questions are not heard as inquisitorial, but as emphatic declarations. Naturally, in such cases, answers are not expected.

Consider the following scenario: At the end of a conversation, a speaker says to his or her co-conversationalist, "Now stop working and get some rest!" Is the speaker uttering an imperative, a command of some sort, to the listener? If one were to examine this utterance for its syntax and prosody (stress and intonation patterns) in a context-free setting, it would appear as a "command." But when looking at this utterance production as part of a closing sequence it becomes clear that at conversation closure points it is not uncommon for one speaker to "invite" the other to end the conversation on a note of finality that in a different sequence, such as a first topic introduction, would constitute a command. "The use of the sheer occurrence of the lexical items, without regard to the placement of utterances in which they occur in the sequential organization of conversation, can be badly misleading..." (Schegloff, 1984, p. 30). Metaphors, likewise, may be found at topic or conversation closure (e.g., "you bet!") to achieve nothing more than the activity of closure. If such an utterance were taken literally—analyzed strictly for its semantic and syntactic properties without a consideration of sequential placement features—one would conclude the speaker is engaging in a discussion about betting. That would be a grossly incorrect reading (Schegloff, 1984).

One of the most telling examples of how syntax is a feature of the sequential arrangements of turn types is found in the production of "assessments," particularly those of a superlative nature. At clearly defined points of demarcation in a conversation, the production of what would ordinarily appear as an assessment can take on a radically different syntactic form. For instance, at topical transitions when a speaker, upon hearing an involved, long-winded complaint, attempts to change the subject, he or she has been found to use purportedly superlative assessments—"something being characterized as 'very, very X'-'something very, very cute happened last night'"—to dramatically shift to a non-troubles-related topic (Jefferson, 1984, p. 195). Or in doctor-patient interaction, physicians have been found to use presumably superlative assessments—"Sounds very good!"—to make a transition from a patient's in-depth reportage of a complaint, or series of complaints, to the next topic, such as treatment protocol (Neustein, 1989).

However, upon closer inspection, it appears that superlative assessments in these instances may not be assessments at all. In the doctor-patient example, if the "assessment" of the patient's progress had been truly meant as laudatory, it would likely have evolved into a sequence of talk about the patient's exceptional progress—but in fact it didn't. Instead, the doctor subsequently closed down the interview. One can see how, in such cases, an apparent superlative assessment actually accomplishes the goal of shifting out of an involved complaint sequence. If such descriptors are taken for their purely syntactical meaning, as superlative assessments, a misreading of the content occurs. While there may be some occasions where a superlative assessment following a lengthy complaint sequence is genuinely superlative, this is uncommon except in those instances where the entire complaint sequence has been punctuated by superlative, exclamatory assessments made by the listener. More typically, the syntax found here, one that is determined by looking at the sequence design format (which consists of a topical transition out of an entrenched complaint sequence), is an "exaggerative declaration" as opposed to a superlative assessment.

In addition, Ochs et al. (1996) point to those instances where a speaker's choice of a syntactic construction is wholly contingent upon the sequential organization of talk. They cite question/answer adjacency pair sequences in which "tag questions" (suffixes appended to questions, such as "isn't that so?" or "don't you?" or "aren't you?") may be likely to occur when the other speaker, at possible transition relevance place, has failed to provide an answer as a second pair part of

a question/answer sequence, and the current speaker is found struggling to obtain one.

2.4. Comparing Conversation Analysis with Speech Act Theory and Discourse Analysis

The conversation analytic method, rooted in a sociological approach to the study of talk-in-interaction, may be distinguished from speech act theory, whose origins are in philosophy, and discourse analysis, which follows in the linguistic tradition. While the linguist focuses on grammatical discourse structure, the conversation analyst focuses on social action (McIlvenny and Raudaskoski, 1992).

Speech act theory, introduced by Austin (1962), views dialogue as a sequence of speech acts, uttered by each party to achieve specific goals and to do things which serve some social function like promising, affirming, or commanding. Grice (1975) expounded on rules speakers must abide by as a precondition for meaningful communication. Allen (1983), Cohen and Perrault (1979), Heeman and Allen (1999), and Perrault and Allen (1980) further developed the speech act model of dialogue into a technique for inferring the plans of a speaker so that the notion of speech acts corresponds with the characterization of actions in AI planners: individual actions are defined in terms of preconditions and effects, and sequences of actions are described as plans to achieve a desired goal. These researchers have developed algorithms for inferring the plans behind what a speaker is saying.

Such theories presuppose speakers' utterances are "always perfectly formed and fully in accordance with Gricean rules" (Hirst, 1991, p. 211). Do decisions to violate rules, such as *not* providing an answer as an appropriate second pair part to a question, result in impaired dialogue or a total communication breakdown? Apparently not. In contrast to speech act theorists, conversation analysts consider the "rules" of interaction as cultural conventions rather than prescriptions that must be followed. That is, the rules themselves are studied as interpretive resources that are invoked by speakers within the context of their use. Schiffrin (1990) posits that "grammatical structures and patterns (and rules) may be emergent from specific instances of communication" (p. 149).

In computational linguistics, a specialized area of discourse analysis, several techniques centered on discourse grammars have been developed to understand or generate natural language utterances in context. For example, Reichman (1985) introduces the concept of "context space" for describing the current set of things being talked about in any given dialogue and defines various "conversational moves," sometimes marked by clue words such as "so," "anyway" and "now," that can effect *transitions* between these spaces. Sidner (1983) and Grosz and Sidner (1986) use various focus registers and spaces to represent the focus of attention in a dialogue where shifts in focus are made by speakers making utterances that either add or subtract items from the current focus registers. McTear (1987) has adapted some of these techniques for constructing natural language interfaces. Lastly, Grosz and Hirschberg (1992), in an effort to perfect speech synthesis, examined the correlation between intonational variation (e.g., pitch, prosody and pauses) and discourse structure, building upon earlier work in this area.

Linguistics has contributed to the field of conversation analysis by offering its knowledge of phonetics. Yet, the study of phonetics demands a careful examination of the ways in which speakers actually deploy the phonetic resources of their language for *interactive* purposes, as is provided by those engaged in analyzing the systematic organization of conversational interaction (Local and Kelly, 1986). For example, the computational linguist's isolated tracking of clue words "so," "anyway," and "now" (and their intonational features) as indications of transitions to another "space" (topic) may produce misleading findings if such occurrences are analyzed without examining the turn-taking features of talk.

Here is an example of the linguist's clue words of "so," "anyway," and "now," that do *not* perform topic transition but, in contrast, attempt to produce continued talk on the *same* topic: A speaker is engaged in the reporting of a trouble, but encounters the other speaker's repeated failure to begin to speak, notwithstanding the many opportunities he or she is given to seize the next turn at the countless possible transition relevance places. The other speaker's failure to begin a turn might connote his or her disinterest or plain apathy. In such a case, the current speaker might then append to his or her description of the trouble the lexical item "so," followed by an immediate "trail off" silence—the kind of silence that provides a clear transition relevance place for the other speaker to begin to speak, which is in direct contrast to a silence that implies a current speaker's claim to the turn, known as a "holding" silence (Jefferson, 1983, 1986). Had the speaker intended, however, to "transition" to another

topic, he or she might have readily produced a "holding" silence, rather than a "trail off" silence, after the clue word "so," serving to preserve his or her right to the speaking turn, and in so doing make use of the current turn to introduce an entirely new topic.

What this example shows is that discourse structures cannot be ascribed rigid definitions, because talk, as an activity that is collaboratively produced by the alternation of speaking turns, is shaped by a locally managed, context-dependent orientation to language structures. The location in the talk where meaning is negotiated, interpreted and "subject to dispute and retrospective interpretation" (McIlvenny and Raudaskoski, 1992, p. 271) is in the sequential arrangement of turns. A sequence package approach to understanding natural language is designed to spot those precise locations in the unfolding talk where meanings are constituted and assembled.

3. Demonstration

3.1. Formal Settings

The sequence packages identified below are not endemic to medical or psychotherapeutic discourse, from which these case examples are derived. Instead, they are generic to talk-in-interaction occurring in "formal settings." In the conversation analytic literature formal settings refer to "'institutional interaction'... that (is) work or task-oriented and 'non-conversational'"(Drew and Heritage, 1992, p. 59). These sequence packages are also not entirely restricted to institutional interactions, because it is the same basic apparatus for the production of talk-in-interaction—the sequential organization of turn-taking activity (Sacks et al., 1974)—that appears in "formal settings" as in "informal" ones with, of course, varying degrees of constraints placed upon turn-type construction. For example, doctors ask by far more questions than patients (Frankel, 1990), and their questions are appreciably more topic-directing or topic-initiating than those few questions posed by patients (Maynard, 1991b; ten Have, 1991).

Two case illustrations are cited below: (1) a call between a suicidal woman and a clinical psychology graduate student manning the nighttime phones at a suicide prevention center (Sacks, 1992b, pp. 376–409); and (2) an interview between a medical resident and an emergency room patient (Neustein, 1989, pp. 68–72).[8] The first demonstrates the graduate student's eagerness to provide "remedy", which appears to be at variance with the caller's display of a clear need for empathy until she, herself, is ready for redress. The second demonstrates the resident's premature attempt to form a diagnosis before the patient has explained the source of her problem. Each illustration, below, is based on empirical data rather than on hypothetical constructions of dialogue.

3.2. Remedy versus Empathy

In the sequence below the psychology graduate student repeatedly offers the suicidal caller a remedy: an invitation to come down to the emergency psychiatric clinic—which operates this suicide hotline—during day hours when the clinic is open.[9] The caller, apparently not ready to receive a solution to her problems, such as counseling offered by the clinic, continually rejects his advice. And it is this advice rejection that, at times, escalates to her confrontational challenge of the hotline operator who is found to engage in extensive interactive work to regain the confidence of the caller. This takes some time, and it is only afterwards that the caller shows a readiness to receive help (e.g., she asks a question about how to obtain the counseling services from the clinic). Had the hotline operator waited patiently for the caller, herself, to display a readiness for practical help, her persistent rejections and her confrontational challenges might very well have been avoided.[10]

Caller: **I can't go through with it-----I can't go through with the evening----I can't (sniffle)**
------(deleted dialogue)----------

I just can't last---and I just can't go it another minute---I just can't
----

Hotline: **This proves your need for psychological help, doesn't it?**
(caller rejects this advice and becomes challenging)

Caller: **Doesn't it bore you to talk to people like me?**

Caller: **Doesn't it bore you on New Year's Eve when you want to go out and get to your party?**

Hotline: Have you tried a clinic?
Caller: I don't want to go to a clinic
Hotline: Why wouldn't you be willing to go to a clinic?
Caller: Cause I don't want to identify with poor people
Hotline: If you're not willing to work at this and help yourself in a clinic, or somewhere, how can I help you over the telephone?

There is an interesting pattern emerging in this segment of talk: the caller begins with the use of an object pronoun ("I can't go through with **it**"), which is only secondarily explained by the caller as referring to the "evening" ("I can't go through with the **evening**"). Sacks (1992b) notes that the caller's initial use of an object pronoun, without specifying the noun to which it refers, reflects the speaker's feelings or emotional state that is only known to the speaker at this point. The caller subsequently amplifies or elaborates on her emotional state by referring to the source of the trouble, thus making it now known to the listener. After this amplification, however, the caller recycles her first utterance component, referencing her emotional state ("**I can't**"). The caller repeats this same sequence pattern in a subsequent utterance: "I just can't last"(emotional state) "I just can't go it another minute" (amplification on emotional state) "I just can't" (recycle of emotional state).

Sacks shows a discrete grammatical sequence operating in both these instances: (a) an initial elliptical or allusive statement, using object pronouns that don't specify the source of the trouble; (b) an elaboration or amplification of the source of the trouble; and (c) an emphatic end, consisting of a recycle of the initial allusive description (of an emotional state), but with the added feature of a declaratory resolve. What this demonstrates is that even in the most ambiguous, convoluted dialogue, utterances are, nevertheless, sequentially ordered with remarkable tightness and precision. It is this orderly design of talk that lends itself to the formulation of algorithms, required for building fully interactive, speech-based interfaces that understand spontaneous speech.

In analyzing the sequence packages appearing in this segment of talk, one might question what the caller is attempting to convey by presenting her problem as first and foremost how it appears to her. While the caller's amplification of her initially allusive descriptor displays her orientation to basic conversational procedures (specifically, those procedures requiring a speaker to make himself or herself understood by the other speaker), the caller's recycle of her emotional state descriptor, after she references the source of the trouble, places the primary emphasis on her emotional state. As a result, any reference to the source of the trouble is overshadowed by the caller's repeated description of her emotional condition. What this demonstrates is that redress is not this caller's most pressing concern, in that her talk reflects an evident need for empathy. Thus, it is not surprising she rejects the advice (going for counseling) offered by the hotline operator.

Much later on in the conversation this same caller displays openness to remedial assistance (counseling at the clinic overseeing this suicide hotline), but not until extensive repair work had been done.[11] Her final receptivity to the graduate student's repeated offers of counseling is displayed here in the call closing sequence:

Hotline: If you need us again you'll call back, won't you?
Caller: I will and I just want to thank you so much. I think I can go wash my face now and put on my pretty dress
Hotline: Good and remember Wilshire Methodist (counseling clinic)
Caller: Wilshire Methodist, yeh I will
Caller: What do I do just go in and ask about it, huh?
Hotline: Yes I believe I'd give em a telephone call and ask em how to do it
Caller: Yeah
Hotline: I think you'll find help there
Caller: I sure thank you

Hotline: Good bye
Caller: Good bye

In saying "**I think I can go wash my face and put on my pretty dress**" the caller is back-referencing to early topical material about her emotional state at the time the call was made: she had been crying and felt she couldn't go through with her evening plans to have a male friend come over to take her out on a date. Studies on closing sequences show that back-referencing, unlike other sorts of closings,[12] is not closing implicative. That is, back references "do not project the relevancy of reinitiating closings...(but) project that development

(of a topic) is sequentially relevant...." (Button, 1987, pp. 111, 112). As we've seen, the consistent topic of this talk is the offer of practical help to the caller. Her choice of a non-closing implicative turn type, as opposed to one that is closure implicative, serves to reopen the talk to discussion about her problem and possible remedy.

When the hotline operator repeats his offer of counseling services at the Wilshire clinic, this time the caller concedes: "**yeh, I will**." But had the caller made this concession to seek counseling within the context of a closing sequence that lacked this critical feature of back referencing, a feature that effectively "undoes" the closing, such a concession could have been viewed merely as a graceful way to end the conversation, rather than a sincere gesture. The caller, in fact, shows that her intent is indeed sincere by making a specific inquiry about the clinic: "**What do I do just go in and ask about it, huh?**"

In sum, the cues that indicate readiness to receive help are couched in neat, ordered utterance sequences. Thus, a sequence package analysis may be applied to talk associated with "advice acceptance," just as this type of analysis may be applied to talk associated with "advice rejection," as shown earlier.

3.3. Premature Remedy

In the sequence below, the medical resident is confronted with an emergency room patient. He must diagnose the problem, determine if admission to the hospital is warranted, and proceed with recommending the proper treatment protocol. But in his attempt to identify the cause of the patient's complaints, he prematurely forms a diagnosis while the patient is still trying to articulate her major symptoms. The resident concludes the problem is of a cardiac nature. However, later in the interview, while he is performing a rather perfunctory organ system review, asking her routine questions about neurological, digestive, musculo-skeletal, and other organ systems, the patient reveals that the problem necessitating her trip to the hospital was not heart-related at all, but rather digestive in origin.

Doctor: **Was it a dull achy pain?**
Patient: **Perhaps, maybe**
Doctor: **But it wasn't a sharp pain or a squeezing pain, was it?**
Patient: **No**
Doctor: **Just like a dull achy pain. And as you stated previously, you said that when you lied down it came again and when you got up it went away?**
Patient: **Yes**

Doctor: **Do you ever um have heartburn, indigestion?**
Patient: **I had trouble with my stomach. I ate something today and I think it didn't agree with me, so I had pain?**[13] **uh I mean just heaviness...so I came here**

When the resident asks the patient if her chief complaint, prompting her rush to the hospital emergency room, can be described as "**a dull achy pain**," she responds with uncertainty: "**Perhaps, maybe?**" This is not a suitable response for the doctor, who is obliged to make critical, time-sensitive decisions concerning patient admission and course of treatment. In the data sample above, the resident changes the patient's responses to fit a model of certainty, rather than vagueness. In the short run, he achieves his goal, but in the long run he undermines his own efforts by basing a diagnosis on distorted history data.

A program designed to apply a sequence package analysis to doctor-patient dialogue might be able to locate at precisely what point in the medical interview the resident began to formulate an errant diagnostic impression. This would be achieved by identifying the patterned features found in the resident's dialogue with the patient that alter the patient's equivocal symptom descriptions, giving them the appearance of greater certainty. In this example, the resident begins to alter the patient's symptom description when, after the patient produces an equivocal description of her symptoms, he immediately poses a sharply contrastive description (**"But it wasn't a sharp pain or a squeezing pain, was it?"**), successfully eliciting the patient's confirmation. Once the resident reaches an agreement between himself and the patient, at least about ruling out the contrastive description of the patient's chief complaint, the resident then uses his next speaking turn to compress into one turn both the prior equivocal descriptor ("**dull achy pain**") and a piece of unambiguous information that had been offered previously by the patient: **"Just like a dull achy pain. And as you previously stated, you said that when you lied down it came again and when you got up it went away?"**

Compression devices in talk-in-interaction serve a crucial function. When a speaking turn encompasses both an "equivocal" fact and one of "certainty," it is not uncommon for the listener to utter a singular confirmation, rather than a compound answer that would affirm one fact and deny the other—unless the two items are in such blatant contrast with one another that it would be inconceivable to do so (Neustein, 1989). However, the use of such a compression device can be likely to result in misinterpretations about critical facts that do not always lend themselves to clarification later on.

In this instance, however, at a subsequent point in the interview, during a routine organ system review, wherein questions were posed to the patient about her general state of health, the patient revises this critical symptom descriptor, constituting it as a "heaviness" (rather than "pain"), caused by a food-related problem: **"I had trouble with my stomach. I ate something today and I think it didn't agree with me, so I had pain? uh I mean just heaviness"**. The patient self-corrects the symptom descriptor ("pain") by first questioning its accuracy—**"so I had pain?"**—in an attempt to determine whether it was indeed pain that she had experienced. She then immediately offers a revised version of her symptoms, presented in the form of a commonly used repair type, known as a self-initiated repair: "uh I mean" (Schegloff et al., 1977).

The distinctive identifying feature of this sequence package is its alternation between "equivocality" and "certainty" across several speaking turns, followed by a recycling of this alternating "equivocality/certainty" contrastive pair type later on in the dialogue. The sequence package has the following structural design: (a) patient produces an "equivocal" symptom descriptor; (b) physician follows with a markedly "contrastive" descriptor; (c) physician compresses into one utterance the "equivocal" descriptor with one of greater "certainty," which was derived from an earlier portion of the talk; and (d) patient recycles her initially "equivocal" descriptor and then, immediately, produces a "certainty" descriptor, in the form of a self-initiated repair.

Notice how the patient herself preserves the indigenous features of this particular sequence package design—the alternating equivocality/certainty contrastive pair type—when she revises the description of her chief complaint. She could have conceivably offered a correction punctuated by certainty (albeit in the characteristically timid way patients are found to make assertions to their physicians), without prefacing it with a recycle of her equivocal descriptor: **so I had pain?** Instead, she couches her correction in a compound descriptor consisting of the dual features of equivocality and certainty, which strongly characterize the earlier portions of the medical interview. This example demonstrates how participants in a dialogue are found to collaboratively produce sequence packages that mold and shape the talk, and in so doing constrain the subsequent production of utterances. It is this sort of systematic and orderly feature of talk that is suited for algorithmic design.

4. Discussion

Modeling the sequence package properties of human dialogue can be a complex and costly project. It is true that sequence package analysis is not appropriate for every natural language program, but it should be reserved for user interfaces requiring sophisticated, advanced natural language understanding. Otherwise, it would be impractical to engineer a costly system to perform perfunctory tasks, easily achieved by a simpler program. But judging from the currents of an industry that aspires to the pervasive use of automatic speech recognition and natural language processing, particularly in complex environments, applying resources now to develop these kinds of programs can help to prepare for the next wave of speech technology.

A speech-based interface that can effectively process sequence package material requires a sophisticated dialogue manager that can interact with parsers, message generators and the database, while managing the conversation as a whole by keeping a detailed history of the dialogue. Because syntactic form and semantic properties are contingent on sequence data, building a sensitive dialogue manager that keeps track of the sequential development of talk is sine qua non for interactive systems. Since talk-in-interaction is inherently digressive and recursive, and characteristically replete with ellipses (omissions), deixes (indirect referents), and idioms, only a highly developed natural language interface can potentially identify where speakers are sequentially located in the talk.

Moreover, placement within the sequential arrangement of talk may dramatically change at the very next speaker turn. A conversational closing sequence, for example, may contain topical expansion features to reopen the talk, thus moving away from conversation closure (Button, 1987); or a help-oriented sequence, for that matter, can suddenly transform into an argument sequence in which the recipient of the help challenges

and/or rejects assistance (Sacks, 1992b; Jefferson and Lee, 1981). It is the characteristic extemporaneity of talk, and its multiple possibilities for sequence development, which makes it imperative for algorithms to be based on probabilities rather than certainties, all of them simultaneously active at all times. Fuzzy set theory may be applicable here.

In spite of the manifold difficulties mentioned above, there are those who contend "it is clear conversation analysis must have a role in Natural Language Understanding, because there is a sense in which...(it) is just a small sub field of artificial intelligence" (Hirst, 1991, p. 225). They point to the impoverished methods of those who design interactive systems without a full appreciation of conversational analytic findings: "there has been an unfortunate tendency to discuss aspects of conversational organization...in the abstract, removed from empirical materials" (Hutchby and Wooffitt, 1998, pp. 244–245).[14] It is further believed "that in order to design computer systems which either simulate, or more ambitiously reproduce the nature of human communication, it is necessary to know about the ways in which everyday (conversational) interaction is organized" (Hutchby and Wooffitt, 1998, p. 241).

However, within the field of conversation analysis itself, strong views have been expressed by some, though not all, of its well-known researchers against deriving programming rules from conversation analytic findings.[15] Button et al. (1995) argue, "whilst having a finite set of rules might seem to be a nice convenience for computational linguistics, the fact (is) that in practice the inferential possibilities of a sentence...(are) not constrained in that way...A myth is being peddled within some quarters of computational linguistics that it is possible exhaustively to specify these inferential possibilities" (p. 176). In *Computers and Conversation* (1990), Button contends that the rules operating in conversation are not givens, nor are they finite: they are not "codifiable" or "reducible to an algorithm" but are, instead, "resources" (p. 84) for speakers to discover as their talk becomes, in situ, that of an achieved orderly form of social activity. In *The Social and Interactional Dimensions of Human-Computer Interfaces* (1995) Button and Sharrock argue that "the prospect of constructing a simulation of ordinary conversation is going to be lacking in procedures for achieving [the] essential feature of projecting turn completion, and thus the management of turn transition will not be arranged in the way that is in conversation" (p. 122). Such views are based upon the understanding that "possible [turn] completion is something projected continuously (and potentially shifting) by the developing course and structure of the talk" (Schegloff, 1992).[16] But as shown above in the example of "trail off" versus "holding" silences, locating the sequence package features that signal a speaker's intent to either hold or give up a speaking turn can help to accurately gauge transition relevance place, thus permitting the design of simulacra that model human dialogue.

Gilbert et al. (1990) analyze yet another one of the central arguments made by conversation analysts against building computational models of conversation, referred to as "indexicality": "the meaning of specific terms or expressions is not fixed, as in a dictionary definition, nor computable using simple rules of deduction, but dependent on the context in which the item is embedded. The hearer has to work actively to find a meaning for the term which makes sense within that context" (p. 254). However, the authors show that just as how, in human-to-human interactions, speakers overcome the problem of context-dependent meanings by treating new material as an instance of a presupposed underlying pattern against which new material can be interpreted, in computational modeling "the grammar a chart parser operates on will have alternative 'patterns' against which the input can be matched" (pp. 255–256). As the system begins to build a history of the dialogue, the apparent ambiguity, caused by several different possible interpretations of the same utterance, would be drastically reduced by the cumulative effect of other data shaping what then becomes the "context" of the talk. In keeping with this argument, a sequence package analysis, by virtue of its capacity to map out the orderly sequences that emerge as indigenous to the talk, can therefore be viewed as one way of providing a dialogue system with a clear, unambiguous schematic design that makes up the context of the talk.

In the final analysis, those who are wedded to the belief that natural language systems cannot possibly simulate human dialogue have, nevertheless, been found to strongly encourage system designers in the building of speech interfaces to incorporate critical features of the turn-taking model (e.g., turn transition relevance), which they refer to as "the development of functional equivalents to the organizational activities...engaged in by speakers and hearers" (Button and Sharrock, 1995, p. 122). In a paper presented at the Thirteenth Scandinavian Conference on Linguistics in 1992, McIlvenny and Raudaskoski argue "that interactional and linguistic

concerns will have to be mutually addressed in computational linguistics. Interactional demands simply cannot be ignored in spoken language artifact design...If we understand computational linguistics in the broad sense of modeling language use and structure using computers as a tool and with language technology as a product, then it should be clear that interactional concerns are crucial" (*Proceedings of the Thirteenth Scandinavian Conference on Linguistics*, 1992, p. 274).

In building natural language systems, neural networks, or connectionist models, may be viewed as the natural choice for investigating the patterns underlying the orderliness of talk, as they require no model, but rather deduce the correct model from the data presented to them.[17] In fact, neural networks are equipped to handle the ambiguities of natural language because of their capacity, when confronted with incomplete or somewhat conflicting information, to produce a fuzzy set—a group of candidate patterns, each with a known likelihood of being the actual pattern for the representation of the portions so far given to it.[18] In short, artificial neural networks, which allow for flexible pattern recognition and co-existing probabilities, might be best suited to developing programs performing a sequence package analysis.

Predictions about how readily natural language systems can be integrated into everyday life, from consumer activities to health care delivery, are difficult to make. A user's acclimation to speech-based computer interfaces depends on complex human factors design issues requiring thoughtful planning and consideration (Lindeman, 1993a, b). Nevertheless, human-computer interaction deserves nothing short of a broad multidisciplinary approach, combining computational linguistics, conversation analysis, and human factors design principles. The use of the sequence packages approach to natural language understanding is an attempt to integrate these diverse disciplines for the purpose of designing speech interfaces (and training tools) that are sensitive to commonly occurring mishaps in communication.

There is, however, a daunting side to this that warrants some consideration. Although this will take time, the potential development of a fully interactive, conversational, speech-based interface, designed to identify when a caller requires vital empathic support, rather than sheer remedial assistance, might actually give the user cathartic relief. And that thought may give one pause. Could machines show greater empathy than humans? What a sad commentary on human service delivery if machines could be more adept at being "human."

Notes

1. There is some discussion of "packaging" in the conversation analytic literature, although not in the sense introduced here. For example, Pomerantz (1997) makes reference to "packaging" as "the ways in which speakers form up and deliver actions" (p. 72) and points to the conversation analytic research on prefatory utterance types that lead up to invitations (Drew, 1984) or to the making of inquiries (Schegloff, 1980). Whereas, as explained by Pomerantz, this notion is related to "the packaging of a given action to understand its consequentiality" (p. 73), sequence package analysis, as demonstrated in this paper, extends well beyond the production of a given action, and the small, closely ordered sequence in which a given action is produced, so as to include the long range, and possibly winding, paths speakers take over large episodes of talk.
2. These are syntactically bounded lexical, phrasal, clausal or sentential constructions found in a speaking turn whereupon at their completion a transition relevance place emerges, allowing the next speaker to begin a turn, provided the current speaker exercises the option of releasing his or her speakership rights to the turn. The current speaker then has the option to continue his or her turn by beginning a new turn construction unit.
3. An emphasis on the structural location of turn constructional units within sequences of talk does not presuppose that syntactic, semantic, prosodic or pragmatic data be ignored. What is proposed is to include critical sequence data along with syntax and other data in the analysis of natural language. Actually, looking at the structural location of lexical items is not a foreign concept to speech interface design engineers. For example, Heins et al. (1997), in their analysis of barge-in use in Spoken Language Systems (SLS), point out the advantage of developing "a system capable of predicting what was said, based on when it was said" by looking at the key words uttered by the system immediately preceding the user's interactional work of beginning his speaking turn at transition relevance place a place in the dialogue where the prior speaker has possibly completed his or her speaking turn (p. 164).
4. Adjacency pairs lend themselves to algorithmic formulation by virtue of their normative character. That is, the production of the first part of a given pair raises the expectation that the relevant second part will follow. This occurrence is defined in the conversation analytic literature as "conditional relevance": on the provision of the first part of an adjacency pair, an appropriate second part is conditionally relevant (Schegloff, 1968, 1972). A request, for example, makes the subsequent granting/rejection hearable as conditionally relevant to the request. If, however, the request were to be misheard by the other speaker as something other than a request, such as a complaint, and that speaker were to reply, in his next utterance, with an argument or an insult, the prior speaker would now have the opportunity to perform the necessary repair work on his initial utterance in the next speaker slot, known as "third position repair" or a "third turn option" (Heritage, 1984).

5. A sequence package analysis may be used to interpret complex idiomatic expressions, not by their literal meaning, but by the sequence design format in which idioms are found to occur (Neustein, 1999).
6. As a precondition for the design of such systems, it is important to understand whether humans display the same conversational patterns when interacting with computers as they do when communicating with other humans. The results of the Surrey WOZ (Wizard of Oz) bionic wizard simulation study of human-computer interaction showed that humans who communicated with computers (which were actually humans simulating computers, although the subjects thought they were interacting with real computers) remarkably displayed some of the same communicative competencies that they commonly use in human-to-human communication. For example, subjects were found to initiate repairs of misunderstandings (on the part of the computer) at the first available transition relevance place. Because there is this systematic carryover of conversational competencies when subjects interact with machines, the sequential patterns underlying spontaneous speech can be anticipated in human-computer interactions.
7. "Talk-in-interaction" is an umbrella term appearing in the conversation analytic literature to describe both mundane conversational interaction, as between two friends, and institutional talk, such as between doctor and patient, attorney and witness, or news reporter and interviewee (e.g., Clayman, 1988; Drew and Heritage, 1992; Frankel, 1990; Heritage and Greatbatch, 1991; Maynard, 1991a).
8. The suicide hotline data was collected in 1963 by Harvey Sacks, who became a fellow at the Los Angeles Suicide Prevention Center. Audio recordings and transcriptions of telephone calls to their clinic hotline were routinely made. The medical interview data was collected in 1983 by the author as part of a research project on doctor-patient communication. The study was conducted at Downstate Medical Center of the State University of New York, under a grant from the American Arthritis Association.
9. It must be noted that this case is not perceived by the hotline as a life-threatening situation; if it were, other actions would be suggested, such as an immediate call to 911 to arrange transport to a hospital.
10. A future study might entail measuring the actual difference between the time allotted to help-line calls that attempt hastily to provide solutions—engendering rejection and possible challenge—and those that patiently await receipt of cues from the caller displaying his or her preparedness to receive help.
11. Some of the repair features occurring in this suicide hotline call pivot on the speakers' collaborative production of idioms (Sacks, 1992b). The general warrant for idiomatic usage is manifold, but what it achieves is the construction of indefeasible arguments (Drew and Holt, 1988; Pomerantz, 1986; Torode, 1995) because the idiom, itself a product of a culturally established "stable body of knowledge," is not subject to challenge. At most, it is the application of the idiom that may be open to challenge. (Sacks (1992a) first made these analytical findings in his examination of the use of proverbs, which may be considered a subset of idiomatic expressions.) When there is a dispute, idioms can be used to forge consensus. Pomerantz (1984) examines dialogue wherein one of the speakers reverses her position on a sensitive matter by supporting "the newly affirmed position with...(an) aphorism" (p. 161). Thus, the speaker's use of an idiom brought about an agreement over a matter than had hitherto been in dispute.
12. A classic example of closing implicative (nonexpansive) closing sequences is an "arrangement sequence": "Bye, make sure to bring the business plan when I see you Wednesday." Such types of sequences show that "although the arrangement projects that the next turn responds to the arrangement, the closing implicative nature of arrangements also projects that following the next turn responses to the arrangement, the subsequent turn may be occupied with a reinitiation of closings" (Button, 1987, p. 109).
13. The descriptor–"pain"–is followed by a question mark, a transcription symbol indicating a rising inflection that does not necessarily indicate a question (Atkinson and Heritage, 1984). Here, the speaker's use of a rising inflection gives the descriptor a slightly inquisitive or uncertain quality.
14. An outspoken critic of natural language software developers, Lucy Suchman, argues in *Plans and Situated Order* (1987) that user models employed in computational linguistics depend on users starting with and sticking as closely as possible to a plan, when in fact user interaction with machines and with other humans is a characteristically ad hoc, situated achievement that does not lend itself to an a priori designation of plans and goals. "Suchman's work has had an important impact on the field of system design. Not only did it propose a strong critique of the user as plan-following and goal-seeking, but it introduced the significance of conversation analysis...to a community of system developers" (Hutchby and Wooffitt, 1998, p. 243).
15. These views may derive in part from the fact that conversation analytic research findings, themselves, cannot be readily quantified (Schegloff, 1993). While studies of conversation uncover the systematic and orderly features of talk, these features are not a "count" in the quantitative sense, but rather convey "the characteristic shape of things that investigators have observed in their materials" (p. 119).
16. Sacks et al. (1974), in their seminal paper on the architecture of converation, showed how transition relevance place may be projected, rather than predicted in conversation as only a possible, rather than a definitive transition relevance place.
17. While Hidden Markov Models, rather than neural networks, represent the predominant approach to building dialogue systems because of the advantage of experience with this sort of design, neural networks are potentially capable of processing input more rapidly (Wooffitt, Frazer, Gilbert, and McGlashan, 1997).
18. Research findings already exist in a related area that may be used as the basis for applying neural network design to the study of talk. For example, in DISCERN, which is a subsymbolic neural network model of script-based story understanding, "schemas are based on statistical properties of the training examples, extracted automatically during training. The resulting knowledge structures do not have explicit representation...There is no all-or-nothing instantiation of a particular knowledge structure. The strongest, most probable correlations will dominate, depending on how well they match the input, but all of them are simultaneously active at all times...(allowing for) inferencing that is intuitive, immediate,...as script-based inference in humans" (Miikkulainen, 2000, pp. 905–919).

References

Allen, J. (1983). Reorganizing intentions from natural language utterances. In M. Brady and R.C. Berwick (Eds.), *Computational Models of Discourse*. Cambridge, Mass.: MIT Press, pp. 107–166.

Atkinson, J.M. and Heritage, J. (1984). Transcript notation. In J.M. Atkinson and J. Heritage (Eds.), *Structures of Social Action: Studies in Conversation Analysis*. Cambridge: Cambridge University Press, pp. ix–xvi.

Austin, J.L. (1962). *How to do Things with Words*. Oxford: Oxford University Press.

Button, G. (1987). Moving out of closings. In G. Button and J.R.E. Lee (Eds.), *Talk and Social Organization*. Clevedon: Multilingual Matters, pp. 101–151.

Button, G. (1990). Going up a blind alley: Conflating conversation analysis and computational modeling. In P. Luff, N. Gilbert, and D.M. Frolich (Eds.), *Computers and Conversation*. London: Academic Press, pp. 67–90.

Button, G., Coulter, J., Lee, J.R.E., and Sharrock, W. (1995). *Computers, Minds and Conduct*. Cambridge: Polity Press.

Button, G. and Sharrock, W. (1995). On simulacrums of conversation: Toward a clarification of the relevance of conversation analysis for human-computer interaction. In P.J. Thomas (Ed.), *The Social and Interactional Dimensions of Human-Computer Interfaces*. Cambridge: Cambridge University Press, pp. 107–125.

Clayman, S.E. (1988). Displaying neutrality in television news interviews. *Social Problems*, *35*:474–492.

Cohen, P.R. and Perrault, C.R. (1979). Elements of a plan-based theory of speech acts. *Cognitive Sciences*, *3*:177–212.

Coulter, J. (1979). Beliefs and practical understanding. In G. Psathas (Ed.), *Everyday Language: Studies in Ethnomethodology*. New York: Irvington Publishers, Inc., pp. 163–186.

Drew, P. (1984). Speakers' reporting in invitation sequences. In J.M. Atkinson, and J. Heritage (Eds.), *Structures of Social Action: Studies in Conversation Analysis*. Cambridge: Cambridge University Press, pp. 129–151.

Drew, P. and Heritage, J. (Eds.). (1992). *Talk at Work: Interaction in Institutional Settings*. Cambridge: Cambridge University Press.

Drew, P. and Holt, E. (1988). Complainable matters: The use of idiomatic expressions in making complaints. *Social Problems*, *35*(4):398–417.

Firth, A. (1995). Transcript notatiton. In A. Firth (Ed.), *The Discourse of Negotiation: Studies of Language in the Workplace*. Oxford: Pergamon, pp. xiii–xv.

Frankel, R. (1990). Talking in interviews: A dispreference for patient-initiated questions in physician-patient encounters. In G. Psathas (Ed.), *Interaction Competence*. Washington, D.C.: University Press of America, pp. 231–262.

Frolich, D.M. and Luff, P. (1990). Applying the technology of conversation to the technology for conversation. In P. Luff, N. Gilbert, and D.M. Frolich (Eds.), *Computers and Conversation*. London: Academic Press, pp. 187–220.

Gaudiano, P. (1999). *Smart Bots: Solutions for the Networked Economy* (White Paper). New York: Artificial Life, Inc.

Gilbert, G.N., Wooffitt, R.C., and Frazer, N. (1990). Organizing computer talk. In P. Luff, N. Gilbert, and D.M. Frohlich (Eds.), *Computers and Conversation*. London: Academic Press, pp. 235–257.

Grice, P. (1975). Logic and conversation. In P. Cole and J. Morgan (Eds.), *Syntax and Semantics*. London: Academic Press, pp. 41–58.

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

Grosz, B.J. and Hirschberg, J. (1992). Some characteristics of discourse structure. *Proceedings of the International Conference on Spoken Language Processing*, *1*:429–432.

Heeman, P.A. and Allen, J.F. (1999). Speech repairs, intonational phrases and discourse markers: Modeling speakers' utterances in spoken dialog. *Computational Linguistics*, *25*(4):222–256.

Heins, R., Franzke, M., Durian, M., and Bayya, A. (1997). Turn-taking as a design principle for barge-in in spoken language systems. *International Journal of Speech Technology*, *2*:155–164.

Heritage, J. (1984). *Garfinkel and Ethnomethodology*. Cambridge: Polity Press.

Heritage, J. and Greatbatch, D. (1991). On the institutional character of institutional talk: The case of news interviews. In D. Boden and D.H. Zimmerman (Eds.), *Talk and Social Structure: Studies in Ethnomethodology and Conversation Analysis*. Cambridge: Polity Press, pp. 93–137.

Hirst, G. (1991). Does conversation analysis have a role in computational linguistics? *Computational Linguistics*, *17*(2):211–227.

Hutchby, I. and Wooffitt, R. (1998). *Conversation Analysis: Principles, Practices and Applications*. Cambridge: Polity Press.

Jefferson, G. (1983). On a failed hypothesis: 'Conjuctionals' as overlap vulnerable. *Tilburg Papers in Language and Literature*, *28*:29–33.

Jefferson, G. (1984). On stepwise transition from talk about a trouble to inappropriately next-positioned matters. In J.M. Atkinson and J. Heritage (Eds.), *Structures of Social Action: Studies in Conversation Analysis*. Cambridge: Cambridge University Press, pp. 191–222.

Jefferson, G. (1986). Notes on 'latency' in overlap onset. *Human Studies*, *9*(2/3):153–183.

Jefferson, G. and Lee, J.R.E. (1981). The rejection of advice: Managing the problematic convergence of troubles-telling and a service encounter. *Journal of Pragmatics*, *5*:399–422.

Lindeman, M.J. (1993a October). How to design "caller friendly" applications: Seven steps to effective caller interface design (Caller Interface, Part I). *Voice Processing Magazine*, Oct. 1993, 1–4.

Lindeman, M.J. (1993b, November). Encyclopedia of human factors: Human factors issues to consider during design (Caller Interface, Part II). *Voice Processing Magazine*, Oct. 1993, 1–5.

Local, J. and Kelly, J. (1986). Projections and 'silences': Notes on phonetic and conversational structure. *Human Studies*, *9*(2/3):185–204.

Machowski, M. (1997). Speech recognition and natural language processing as a highly effective means of human-computer interaction. Paper written for User Interface class. Denver: Computer Science Department, University of Colorado.

Markowitz, J.A. (1996). *Using Speech Recognition*. New Jersey: Prentice Hall.

Maynard, D.W. (1991a). The perspective-display series and the delivery and receipt of diagnostic news. In D. Boden and D.H. Zimmerman (Eds.), *Talk and Social Structure: Studies in Ethnomethodology and Conversation Analysis*. Cambridge: Polity Press, pp. 164–192.

Maynard, D.W. (1991b). Interaction and asymmetry in clinical discourse. *American Journal of Sociology*, *97*:448–495.

McIlvenny, P. and Raudaskoski, P. (1992). The mutual relevance of conversation analysis and linguistics: A discussion in reference to

interactive discourse. In L. Heltoft and H. Haberland (Eds.), *Proceedings of the Thirteenth Scandinavian Conference on Linguistics*. Roskilde, Denmark: Department of Languages and Culture, Roskilde University, pp. 263–277.

McTear, M.F. (1987). *The Articulate Computer*. Oxford: Blackwell.

Miikkulainen, R. (2000). Text and discourse understanding: The DISCERN system. In R. Dale, H. Moisl, and H. Somers (Eds.), *A Handbook on Natural Language Processing: Techniques and Applications for the Processing of Language as Text*, Vol. II. New York: Marcel Dekker, pp. 905–919.

Neustein, A. (1984). Linguistic technology and artificial intelligence in medical history-taking. *Update: Computers in Medicine*, *2*(5):56–60.

Neustein, A. (1986a). Computer-aided instruction for improving history-taking skills (Part I). *Physicians and Computers*, *4*(6):32–35.

Neustein, A. (1986b). Computer-aided instruction for improving history-taking skills (Part II). *Physicians and Computers*, *4*(7):33–37.

Neustein, A. (1989). Medical history-taking as an interactive event. In W. von Raffler-Engel (Ed.), *Doctor-Patient Interaction* Amsterdam/Philadelphia: John Benjamins Publishing Co, pp. 61–76. Pragmatics and Beyond New Series, Vol. IV.

Neustein, A. (1999). How "sequence packages" can aid language understanding. *Speech Technology*, *4*(4):36–37.

Neustein, A. (2000). Designing the virtual agent: Some theoretical and practical considerations. *Call Center News Service*, 4(24):1–4.

Ochs, E, Schegloff, E.A., and Thompson, S.A. (Eds.). (1996). *Interaction and Grammar*. Cambridge: Cambridge University Press.

Perrault, C.R. and Allen, J. (1980). A plan-based analysis of indirect speech acts. *American Journal of Computational Linguistics*, *6*(3/4):167–182.

Pomerantz, A. (1984). Pursuing a response. In J.M. Atkinson and J. Heritage (Eds.), *Structure of Social Action: Studies in Conversation Analysis*. Cambridge: Cambridge University Press, pp. 152–163.

Pomerantz, A. (1986). Extreme case formulations: A way of legitimizing claims. *Human Studies*, *9*(2/3):219–229.

Pomerantz, A. and Fehr, B.J. (1997). Conversation analysis: An approach to the study of social action as sense making practices. In T.A. van Dijk (Ed.), *Discourse as Social Interaction*. Thousand Oaks: Sage Publications, pp. 64–91. Discourse Studies: A Multidisciplinary Introduction, Vol. 2.

Reichman, R. (1985). *Getting Computers to Talk Like You and Me*. Cambridge, Mass.: MIT Press.

Riloff, E. and Lehnert, W. (1994). Information extraction as a basis for high-precision text classification. *ACM Transactions on Information Systems*, *12*(3):296–333.

Sacks, H. (1992a). In G. Jefferson (Ed.), *Lectures on Conversation*, Vol. I. Oxford: Blackwell, pp. ix–818.

Sacks, H. (1992b). In G. Jefferson (Ed.), *Lectures on Conversation*, Vol. II. Oxford: Blackwell, pp. ix–580.

Sacks, H., Schegloff, E.A., and Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*:696–735.

Schegloff, E.A. (1968). Sequencing in conversational openings. *American Anthropologist*, *70*:1075–1095.

Schegloff, E.A. (1972). Notes on a conversational practice: Formulating place. In D. Sudnow (Ed.), *Studies in Social Interaction*. New York: Free Press, pp. 75–119.

Schegloff, E.A. (1980). Preliminaries to preliminaries: "Can I ask you a question?" *Sociological Inquiry*, *50*(3/4):104–152.

Schegloff, E.A. (1984). On some questions and ambiguities in conversation. In J.M. Atkinson and J. Heritage (Eds.), *Structures of Social Action: Studies in Conversation Analysis*. Cambridge: Cambridge University Press, pp. 28–52.

Schegloff, E.A. (1992). To Searle on conversation: A note in return. In J. Verschueven (Ed.), *Searle on Conversation, Pragmatics and Beyond New Series*. Amsterdam/Philadelphla: John Benjamins Publishing Co., Vol. 21, pp. 113–128.

Schegloff, E.A. (1993). Reflections on quantification in the study of conversation. *Research on Language and Social Interaction*, *26*(1):99–128.

Schegloff, E.A., Jefferson, G., and Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, *53*:361–382.

Schegloff, E.A. and Sacks, H. (1973). Opening up closings. *Semiotica*, *7*:289–327.

Schiffrin, D. (1990). The principle of intersubjectivity in communication and conversation. *Semiotica*, *80*(1/2):121–151.

Sidner, C.L. (1983). Focusing in the comprehension of definite anaphora. In M. Brady and R.C. Berwick (Eds.), *Computational Models of Discourse*. Cambridge, Mass.: MIT Press, pp. 267–330.

Suchmann, L. (1987). *Plans and Situated Actions*. Cambridge: Cambridge University Press.

ten Have, P. (1991). Talk and institution: A reconsideration of the 'asymmetry' of doctor-patient interaction. In D. Boden and D. Zimmerman (Eds.), *Talk and Social Structure*. Cambridge: Polity Press, pp. 138–163.

Torode, B. (1995). Negotiating 'advice' in a call to a consumer helpline. In A. Firth (Ed.), *The Discourse of Negotiation: Studies of Language in the Workplace*. Oxford: Pergamon, pp. 345–372.

von Raffler-Engel, W. (1989). Introduction. In W. von Raffler-Engel (Ed.), *Doctor-Patient Interaction.* Amsterdam/Philadelphia: John Benjamins Publishing Co., pp. xvii–xxxviii. Pragmatics and Beyond New Series, Vol. IV.

Wooffitt, R., Frazer, N.M., Gilbert, N., and McGlashan, S. (1997). *Humans, Computers and Wizards: Analyzing Human (Simulated) Computer Interaction*. London: Routledge.

Wooffitt, R. and MacDermid, C. (1995). Wizards and social control. In P.J. Thomas (Ed.), *The Social and Interactional Dimensions of Human-Computer Interfaces*. Cambridge: Cambridge University Press, pp. 126–141.