Organization of Pre-Stored Text in Alternative and Augmentative Communication Systems: An Interactive Schema-Based Approach

by

Peter Bryan Vanderheyden

ASEL Technical Report #AAC9501

asel

Applied Science & Engineering Laboratories
University of Delaware/A.I. duPont Institute
1600 Rockland Rd., P.O. Box 269
Wilmington, DE 19803
Phone: (302)651-6830 FAX: (302)651-6895

ACKNOWLEDGMENTS

Work on this project was performed at the University of Delaware Center for Applied Science and Engineering, located at The Alfred I. duPont Institute, in Wilmington, DE, with support from the United States Department of Education, National Institute on Disability and Rehabilitation Research, Grant # H133E30010-94. Additional support was provided by the Nemours Foundation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The field of augmentative and alternative communication (AAC) is concerned with assisting individuals with severe physical and language impairments to communicate more effectively. Existing AAC systems make use of a variety of approaches to accelerate sentence generation, including different selection methods, encoding strategies, and natural language processing. Augmented communicators continue to produce words at a very slow rate, and have difficulty participating actively in conversation.

However, only recently have AAC systems begun to make use of the predictable patterns that occur in conversation. To date, such systems have focussed on either highly constrained and relatively content-free utterances, or on loosely structured, monologue-type text.

This thesis develops an alternative but compatible approach to facilitating conversational participation in AAC which attempts to target a broader range of conversations, representing both their content and structure. Motivated by schema theory, this approach applies schema structures to the domain of conversation. A set of structures is proposed with which text from past conversations can be made available for reuse.

To demonstrate this approach, a prototype is developed and evaluated. The prototype behaves as an interface that augments a user's current AAC system by providing access to conversational schemata created and updated by the user. In the evaluation study, two individuals used the interface while taking part in a series of mock job interviews. Results of the study were encouraging.

Chapter 1

INTRODUCTION

An individual who uses an augmentative communication system gains an alternative voice, one that can augment and complement a natural voice that is difficult to produce or to understand. In order to “speak” with this alternative voice, augmentative communication systems require the individual to physically select symbols representing the words to be spoken, either by hand or using some other motor channel. Dependence on motor abilities that are also impaired, however, means that utterances can take much time and effort to produce.

To reduce this time to speak, systems could make sentences or larger segments of text available as single units. Such “reusable” text could then be spoken, as is, with very little effort. Alternatively, when reusable text is not available, the individual could select fewer items and speak in short, or incomplete, sentences. Although these two strategies might both reduce the time to produce a sentence, speaking with incomplete sentences, or with noticeably “canned” sentences that are not quite context-appropriate, can be interpreted by unfamiliar listeners as signs of cognitive impairment.

The challenge to the designers of augmentative and alternative communication (AAC) systems, therefore, becomes one of increasing the rate of spoken output without compromising the image of the augmented communicator, the individual using the system. Contextual information is very important during conversation for determining both the meaning and the appropriateness of an utterance. Within the proper context, then, reusable text that is made available and selected by the user will not sound canned. However, such precise contextual information is not available, in an automatic fashion, to current AAC systems.

It is, however, available to the individual who is speaking through the system. The individual is aware of the situation in which the conversation is taking place, and of the intended self-image. The challenge for the system becomes one of making context-appropriate reusable text available to the individual in a reasonable amount of time and without excessive cognitive load. This interaction between system and user should involve as little effort as possible during a conversation, so that the individual can concentrate on the topic and on the other participants.

I suggest that there are three requirements for an AAC system to facilitate conversational interaction:

(1) the augmented communicator produces text at some time prior to the conversation, and stores it in the system;

(2) the AAC system supports an organization for stored text that is consistent with observed features and patterns in conversation;

(3) during a conversation, and with very little effort, the augmented communicator is able to retrieve desired and appropriate pre-stored text.

Requirement (1), the pre-storage of text, is already a common feature in many systems. However, few systems offer real support for (2) and (3), structuring and retrieving this text for conversation. Notable exceptions are the systems CHAT and TOPIC that were developed at the University of Dundee (and their realization as a commercial product, Talk:About, manufactured by Don Johnston Inc.). CHAT (Alm et al., 1992) supported quick production of simple utterances, including greetings, small talk, and farewells, applicable in many conversational contexts. TOPIC (Alm et al., 1989) provided a database of reusable text, taken from previous conversations and linked by topic, and was concerned mainly with the monologue-type segments that occur in the body of a conversation.

This thesis discusses an alternative, possibly complementary, approach to organizing and retrieving pre-stored conversational text in an AAC system. This approach is motivated by Schank's (1982) description of schemata, representing the dependence of how we behave and think on how we behaved and thought in similar situations in the past. Situations that are judged similar are grouped together to form the basis for expectations about future instances of similar situations. These can include expectations about what people or things will be involved in a situation, what events will occur, and in what order they will occur. An individual's cognitive system can store experiences more efficiently in this "schematized" form, and can organize new experiences around these schemata.

In this thesis, I explore the notion of storing reusable text for schematized situations in a manner similar to that described by Schank's schemata. The intuition is that an AAC system which represents conversation similarly to our own cognitive system should be able to offer the user access to conversational text in a way that is both intuitive and efficient. In pursuit of this goal, a prototype interface, SchemaTalk, has been developed that adds schematic organization to a text-based AAC system, and enables the user to access that information. The effectiveness of this configuration was investigated in a study in which two participants, involved in mock job interviews, communicated using the interface and sentences they had organized into schemata.

Chapter 2

AUGMENTATIVE AND ALTERNATIVE COMMUNICATION

In North America, there are over two million people unable to speak adequately to meet their communication needs (American Speech-Language-Hearing Association, 1991; cited in Beukelman & Mirenda, 1992, p. 4). The field of augmentative and alternative communication is concerned with developing methods and devices, tuned to the abilities of each individual, to facilitate active and effective participation in conversation and other forms of communication (e.g., writing). In this thesis, I will focus on augmenting spoken conversation, and on electronic AAC systems with speech synthesis capabilities.

2.1 AAC Users

The American Speech-Language-Hearing Association (ASHA) gives the following definition for the population of individuals who might use AAC systems:

> Individuals with severe communication disorders are those who may benefit from [AAC] — those for whom gestural, speech, and/or written communication is temporarily or permanently inadequate to meet all of their communication needs. (American Speech-Language-Hearing Association, 1991, p. 10; quoted in Beukelman & Mirenda, 1992, p. 4)

Emphasis is placed on the individual's natural modes of communication not being adequate to meet all of their needs. In some situations, and with some conversational partners, individuals may prefer to communicate with their natural voice or gestures, and may find it more effective to do so.

Communication may be severely impaired as a consequence of a congenital neurologic dysfunction, such as cerebral palsy, mental retardation, autism, developmental verbal apraxia, and specific language disorders (Mirenda & Mathy-Laikko, 1989, p. 3). Severe communication impairment may also be acquired as a result of amyotrophic lateral sclerosis (ALS), multiple sclerosis (MS), brain injury, stroke, or spinal cord injury (SCI) (Beukelman & Yorkston, 1989, pp. 42-47).

These same neurological conditions may impair non-language motor abilities and perception, as well. Motor deficits, such as hypertonic muscle tone, often accompany cerebral palsy, as do visual and hearing deficits (Mirenda & Mathy-Laikko, 1989, p. 4). Acquired brain or spinal injuries may result in limited mobility or sensory losses. The ability to control a communication device that relies on motor input will, in many cases, be affected.

2.2 Communicative Competence

The goal of AAC is to assist the user in becoming a competent communicator. Light defines communicative competence as the ability "to initiate and maintain daily interactions within the natural environment" (1989, pp. 138) adequately to meet daily needs. This presupposes knowledge, judgment, and skill in four areas (Light, 1989, pp. 139):

(1) linguistic competence, using the rules of the language code (phonology, morphology, syntax, and semantics);

(2) operational competence, using the AAC system itself (e.g., controlling the volume, retrieving and producing a word);

(3) social competence, interacting with others (e.g., initiating a conversation, reacting to what another person says);

(4) strategic competence, adapting to a situation and compensating for any difficulties that may arise (e.g., rephrasing an utterance if the listener did not understand it, rather than simply repeating it).

As well, communicative competence is relative, not absolute. An individual may be competent interacting with one partner but not with another, in one situation but not another, or at one stage of the conversation but not another.

People interact for a variety of reasons: to communicate their wants or needs, to convey or receive other information, to increase social closeness, and to fulfil the requirements of social etiquette (Light, 1988, p. 76). Interactions with different goals may differ in many ways. Social etiquette and expression of wants and needs, for example, may be characterized by highly predictable interactions in which communication rate is very important. Communication rate may also be important when the goal of the interaction is to convey or receive information (Light, 1988, p. 76).

AAC systems need to recognize these varying demands in order to support communication more effectively.

2.3 Components of an AAC System

An AAC system (Figure 2.1) can be described in terms of its language model, and

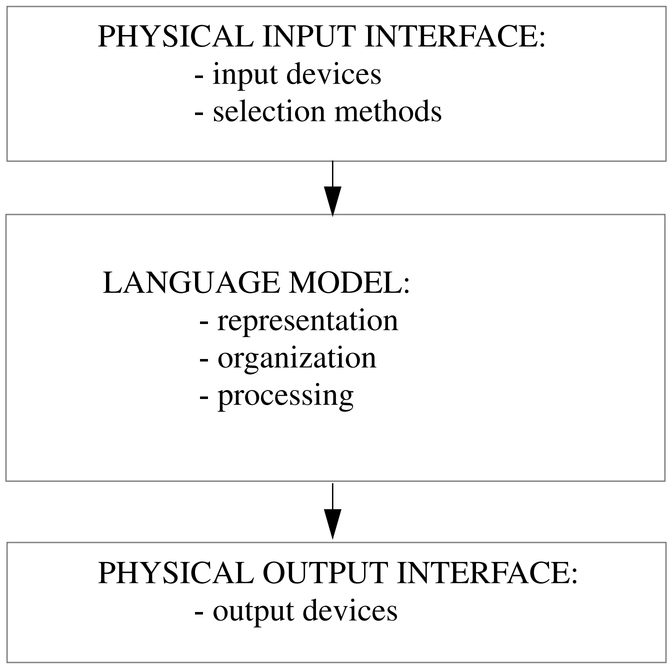

Figure 2.1: Components of an AAC system

its input and output interfaces (Demasco & Mineo, 1995). The input interface provides the user with a method for selecting symbols (letters, words, or icons) represented in the system. How symbols are represented, organized and processed is specified in the language model. The user's message is presented by the output interface.

Currently, AAC systems can accept input from a wide variety of devices. Keys on a keyboard can be selected with the user's hands, or with a stick fastened to a head-mount or held in the user's mouth. For users with more limited motion, switch devices can be activated by movement of the hand, foot, or eyebrow. A beam of light, emitted from a head-mounted source and detected by receivers on the AAC system, can be used to make selections on a switch or a keyboard. There is even work in progress to detect and follow the user's eye gaze (Sandler, 1994).

Selection methods can be generally classified as either direct or scanning. With direct selection, the user indicates the desired item from a set of items (Beukelman & Mirenda, 1992, p. 58). Spelling words on a computer keyboard is an example of direct selection. Each key represents a letter in the alphabet and is selected by the user via a key stroke. Using a scanning method, items in the set are displayed in some predetermined order by the system, or by a conversational partner or facilitator, and the user indicates when the desired item has been presented (Beukelman & Mirenda, 1992, p. 62). In row-column scanning, for example, symbols are organized into rows and columns and the system highlights rows until the user indicates the row containing the desired symbol. The system then highlights columns until the user indicates the column containing the desired symbol. The system selects the symbol located at the point where that row and column intersect.

Scanning input can be slower than direct selection, because the user must wait while the system traverses undesired items. However, direct selection of even a relatively small number of items requires a fair amount of motor dexterity. With scanning, an individual with severely limited motor abilities is potentially able to select any symbol represented in the system using only a single key or switch.

The words and messages of the language model that are available for selection on an AAC system can be represented in a variety of symbol sets. The most appropriate set for a specific user will depend on that user's age, cognitive and language abilities, and perceptual abilities (see Beukelman & Mirenda, 1992, pp. 21-27, for a discussion on "representational symbols"). Letters, pictures, abstracted icon sets, or combinations of all three are used by different systems.

The symbols available in a system must be organized and displayed to the user in some fashion. For instance, in a letter-based system performing row-column scanning, decisions must be made about how many rows and columns to use, and in which order the letters should occur in the columns. A system may contain more symbols than it can display at one time. Symbols must then be organized to provide the user with a consistent method for accessing them. This organization includes how symbols are physically distributed on the system display, as well as how hidden symbols are reached.

In addition to providing this vocabulary of symbols, an AAC system may provide processing capabilities to facilitate its use. For instance, a vocabulary symbol may have a phrase or sentence associated with it which is retrieved by the system when the symbol is selected. Several strategies of processing to increase communication rate are discussed in Section 2.4.

Both visual and spoken output are available on many AAC systems.Visual output methods include displaying symbols on a computer screen, or printing symbols out on paper. Spoken output is achieved with either digitized or synthesized speech. Digitized speech is pre-recorded and fixed, while synthesized speech is generated at the time of message production and is not constrained in terms of message content. However, the voice quality of digitized speech is far more natural than that of synthesized speech in current AAC systems.

2.4 Strategies for Accelerating Message Production

One of the great challenges to AAC has been to increase the rate of message production by augmented communicators. With AAC systems, production rates of between 2 and 20 words per minute are not uncommon, an order of magnitude slower than for unaugmented speakers using their natural voices (Kraat, 1987, p. 63). This slow production rate may affect how augmented communicators participate in conversations, and how they are viewed by their conversational partners.

Two acceleration strategies used in existing AAC systems are message encoding and message prediction. Compansion, a third acceleration strategy, in which the user inputs a sentence in “telegraphic” style and the system completes the sentence, is discussed in Section 2.5.1.

Message encoding refers to retrieving a word, or a longer unit of text, by using an associated code that requires fewer selections. Codes can consist of letters, numbers, icons, or some combination of these. For example, “Please open the door” may be associated with the code “OD”, the initial letters of the salient words “open” and “door” (salient letter codes, Light & Lindsay, 1992, p. 35). An alternate encoding might be “DD”, because the sentence is giving “directions” regarding the “door” (letter category codes, Light & Lindsay, 1992, p. 35).

Flexible abbreviation expansion (Stum & Demasco, 1992) is an alternative letter coding system for abbreviating words. In a standard coding or abbreviation scheme, there is a fixed set of codes which is stored by the user ahead of time. The processing done by the system is a simple table look-up. In contrast, flexible abbreviation expansion has no fixed table of codes. Instead, the user and system follow rules to expand the letter codes.

Message prediction is a dynamic strategy, whereby the system makes use of earlier portions of the message to interpret and predict the next icon, letter, word, or larger segment of text (Beukelman & Mirenda, 1992, pp. 42-44). On a spelling-based system, for example, after the user selects the letters “s”, “t”, and “u”, the system may predict the word in progress to be “student”, “stupid”, or “studying”. If the intended word is in this list, the user can select it immediately rather than continuing to spell it. A system could also make use of syntax rules (VanDyke et al., 1992) to better predict the intended word.

2.5 Example AAC Techniques

To illustrate possible techniques for AAC, two systems are described here in detail. The first, Compansion, is different from other systems in that it uses natural language processing techniques to complete a telegraphic sentence once the user has entered it. The second system, the Liborator^{TM} (Prentke Romich Company), is an example of a commercially available system, and contains a variety of tools with which the user can edit and navigate through stored text. One participant in the evaluation study, discussed in Chapter 5, normally communicates using the Liberator.

2.5.1 Compansion

The strategies discussed in Section 2.4 attempt to accelerate message production by reducing the number of selections required to retrieve a word or sequence of words. To produce a novel sentence, the user must specify each word or phrase, and the resulting sentence may or may not be well-formed. A well-formed sentence may be retrieved as a single unit, but only in the fixed form in which it was encoded.

Compansion (McCoy et al., 1994a; Demasco & McCoy, 1992) is a strategy for producing a well-formed, novel sentence by selecting only uninflected content words. The system expands this input to produce a complete sentence. Compansion, in its current implementation, consists of three stages: a word-order parser, a semantic parser, and a translator/generator.

The sequence of input words enters the word-order parser. Here, each word is matched with an entry in the system lexicon, and its part of speech and semantic attributes are identified. The word-order parser then uses a grammar either to reject the sequence or to identify the main verb, attach modifiers to the words they modify, and identify embedded sentences. As the name implies, the word-order parser assumes words appear in the input in the same order they appear in the completed sentence. In cases where more than one part of speech is possible, multiple interpretations of the input are processed.

The lexicon contains nouns represented in a hierarchical meaning structure. "Hammer", for example, is classified specifically as a tool, more generally as a physical object, and so on. Verbs are also represented hierarchically, and are associated with case frames. A verb case frame contains cases for noun phrases that can occur with the verb in a sentence, and preferences for the kinds of noun phrases that can fill each case.

The word-order parser passes the information derived from the input to the semantic parser, which uses the lexicon entry for the main verb in the input to build a case frame for the entire sentence. Noun phrases identified by the word-order parser are then used to fill cases in this case frame, moderated by the verb's preferences for case fillers.

When all required cases for a verb have been filled, the case frame is passed to the final component of the compansion system, the translator/generator. Here, the case frame is translated into words to produce the sentence, with function words and morphological endings added as required.

As an example, consider the input: "John break window hammer yesterday." Both "break" and "hammer" can be verbs, but only "break" allows for a complete parse of the sentence. "John" is identified as a human object, "window" as a fragile inanimate object, and "hammer" as a tool. The preferences for the case fillers of "break" suggest that the agent, John, used an instrument, the hammer, to break the theme, the window. "Yesterday" modifies the verb and specifies the past tense. Definite articles, "with" before the instrument, and the past tense marker are filled in automatically to produce the completed sentence: "John broke the window with the hammer yesterday." Any alternative interpretations of the input would be produced if the user did not accept the one given.

A modification to the current system that is underway involves replacing the current lexicon with a database that can access information from various sources (Zickus, 1995). A study into interactions between AAC users and conversational partners, and the implications for Compansion, is also in progress (McCoy et al., 1994b; Vanderheyden et al., 1994).

2.5.2 The Liberator™ and Minspeak™

The Liberator consists of a "keyboard", a small liquid crystal display (LCD), and a speaker for outputting synthesized speech (as well as a printer, serial communication port, and other features not discussed here). The keyboard is typically configured to use the maximum of 128 keys, but may also function as a set of 8 or 32 keys. A light-emitting diode (LED) is situated next to each key, and a key guard (a plastic layer with a round hole for each key) lies on top, to help the user's finger find and stay on the intended key.

In the 128-key configuration, approximately 64 of the keys are associated with pictographic icons, as well as characters, digits, and a number of keys regularly found on a typewriter or computer keyboard. A short sequence of keys can be associated with a word or phrase, fewer keys than would be required to spell the word or phrase. Baker and Barry (1990) noted that, in English orthography, 26 letters are combined into words that are often much more than three characters in length. If these words were retrieved by 3-key sequences, between 60% (Baker & Barry, 1990) and 85% (Prentke Romich Company, 1991) fewer key presses would be required to produce text than by spelling.

Minspeak (Baker, 1982) is a strategy for encoding text, similar to letter category encoding, using icons instead of letters. A sequence of two to three multi-meaning icons is associated with each word or phrase. For example, the sentence "When do we eat?" can be encoded as a sequence of the icons "question mark", "clock", and "apple" (Prentke Romich Company, 1991, Appendix - Tutorial Demo Vocabulary, p. 1), representing the concepts question, time, and food. Rule-based programs are available to provide the user with a systematic strategy for encoding words and phrases. In the Words Strategy™ (Prentke Romich Company, 1992) program, for example, "turtles" is encoded as the sequence of icons "zebra", "treasure chest", and "noun+s". The icons represent the semantic category and specific word meaning, and the syntactic category of the word, respectively.

Minspeak is a strategy for lexical retrieval only, a "transducer" for associating key sequences directly with stored text, and performs no linguistic processing. The user may, intentionally or by mistake, produce an ungrammatical or nonsense sentence. A future version is planned that will process sentence syntax and infer semantic roles, and will be able to produce a complete sentence from partial input (Baker & Nyberg, 1989).

The Liberator, using Minspeak, provides access to pre-stored words, phrases, and sentences, or any other segments of text, by selecting a sequence of keys. Organizing a set of words into a "Theme" saves the user one key selection, as long as only words within the current Theme are being retrieved. The user may also choose to spell text letter by letter, and edit text on the LCD display.

A key sequence can also retrieve a sentence template, or a sentence that contains a blank (a "Minsert") to be filled in at retrieval time. One or more words are selected to fill the Minsert, and complete the phrase or sentence. This provides a certain amount of flexibility for using pre-stored text. However, there are no linguistic processing mechanisms to ensure agreement between the template and the Minsert contents. For example, separate sentence templates might be required to handle the case when either a singular or plural noun phrase could fill a Minsert.

Stored text can be organized and reorganized by the user in a structure called a "Notebook". A key can be programmed to allow the user to have the Liberator speak the contents of the Notebook a sentence at a time, pausing for the user to press a key before speaking the next sentence. Thus, a user could store a speech or a conversation and proceed through it, controlling the rate of sentence production.

Chapter 3

APPLYING CONVERSATIONAL STRUCTURE TO AAC

In the previous chapter, a number of general approaches currently being applied to AAC were described briefly, and two systems were discussed in more detail. Existing AAC systems are accessible using a wide variety of input devices, and contain large vocabularies of words and phrases which users can combine together and speak.

A number of systems attempt to facilitate and accelerate augmented communication further by doing one or more of the following:

(1) organizing the vocabulary to reduce the time and effort required to find and access it;

(2) providing the opportunity for users to store and re-use their own utterances, in addition to the regular stock vocabulary;

(3) applying natural language processing techniques to the user's input.

However, a fourth consideration, one that many speaking individuals take for granted, has not been sufficiently addressed in AAC:

(4) organizing text to support participation within the context of conversation.

A conversation-based organization of text does not prohibit a system from pursuing points (1), (2), and (3) as well.

Conversation is not just a random exchange of utterances by two or more people. Rather, it is often a structured and cooperative construct in which participants follow well-established, though rarely stated, rules and expectations. Yet participants in unaugmented conversation may pay little attention to controlling how the conversation develops. By contrast, augmented communicators must continuously be thinking about how to produce their utterances using their AAC system. Alm urges that this should not be, that "a conversation aid should ideally be able to simulate the non-conscious control mechanisms which are available to unimpaired speakers." (Alm, 1988, p. 29)

3.1 Stages of Conversation

A conversation can be described in terms of the following sequence of stages (Alm, 1988, p. 107), understanding that the same conversation need not contain them all:

(1) greetings;

(2) smalltalk;

(3) main business;

(4) wrap-up remarks;

(5) farewells.

Greetings and smalltalk can be used to open the conversation, wrap-up remarks and farewells to close it. These four stages serve a social and pragmatic purpose, exchanging little or no real information. The main topic is discussed in the third stage of the conversation.

3.1.1 CHAT

During those stages in which the content of an utterance was not important, the prototype CHAT system (Alm, 1988; Alm et al., 1987; Alm et al., 1992) allowed the user to construct and generate an utterance, to choose from several prepared utterances, or simply to request the system to choose and produce an utterance on its own. To avoid repeating the same utterance, the system made a constrained random choice of pre-stored utterances for the current stage of the conversation. At any time, the user could advance the system to the next, or any other, stage.

The user could select a mood (polite, informal, angry, and so on) for the conversation and the name of the other participant. Pre-stored utterances of the selected mood were then made available, and the name of the other participant was inserted where appropriate. For example, an informal greeting to "Bill" might have taken the form: "Hi, Bill. How's it been going with you?" (Alm et al., 1987, p. 129)

CHAT also provided the user with easy access to utterances for filling pauses during speech (e.g. "so", "well"), and for giving feedback when another person was speaking to indicate that the user was listening attentively (e.g. "uh-huh").

Example conversations using CHAT (transcribed in Alm, 1988, and in Alm et al., 1987) were brief, suggestive of what might occur between acquaintances passing on the street and exchanging a few words.

3.1.2 TOPIC

To assist the augmented communicator during the less predictable main body of the conversation, a database management system and interface called TOPIC (Alm et al., 1989; called CHAT4 in Alm, 1988, p. 159) was developed. Segments of text could be stored by the user as records in the database. The purpose of TOPIC was to, whenever possible, predict what text the user would choose next, and to help the user find relevant segments of text stored in the system (Alm et al., 1989, p. 148).

Database records could contain text several sentences in length, and were identified by their speech acts, subject keywords, and frequency of use. The system suggested possible next utterances by considering subject keywords and speech acts of the previous item selected in the conversation. It then searched for records with matching subject keywords in the database, giving preference to those records that were accessed frequently.

Conversations produced by TOPIC tended to be longer and similar to monologues. Examples included jokes, stories, and lectures (Alm, 1988, p. 171-176).

3.2 Schemata

While CHAT and TOPIC can provide valuable aids to conversation, an AAC system may well be able to provide more support to the augmented communicator. CHAT provided interactional utterances, but with little content. TOPIC provided text with content, but gave little support for interaction.

In this section, an approach is developed that supports both meaningful and interactive conversation. As well, this approach provides a more formal definition of the structural components in a conversation and how they are organized.

3.2.1 Schema Theory

Schank and Abelson (1977) suggested that a person developed mental scripts as a result of repeated experiences. Each script represented the typical sequence of events that occurred in a particular situation. New experiences were interpreted on the basis of existing scripts. Events which did not already match an action in a script but which were considered important were either added to the script for subsequent use, or developed into a new script.

Schank (1982) extended and modified the idea of scripts into a hierarchy of schema structures. Memory organization packets (MOPs) replaced scripts in representing typical situations. Meta-MOPs represented higher-level goals, and contained MOPs corresponding to situations that, together, could satisfy those goals. For example, the meta-MOP "trip" (Schank, 1982, p. 99) might involve planning, preparation, travel, arrival, and so on, each of which was represented as a MOP.

Each MOP was associated with a sequence of scenes, and each scene represented the sequence of actions leading to a specific goal. For example, the MOP for a trip by airplane could include scenes for getting tickets, driving to the airport, boarding the plane, and so on. Each scene was associated with any number of alternative scripts, where a script contained one instance of the actions that fulfilled the scene's goal.

MOPs and scenes could each be organized in hierarchies, with MOPs and scenes at higher levels representing more general situations. For example, "travel by train" and "travel by plane" MOPs could both contain a "buy ticket" scene that contained actions specific to buying train and plane tickets, respectively. The "buy ticket" scene in a more general "travel" MOP would then capture aspects common to both.

To demonstrate the use of schemata in understanding stories and answering questions, Schank developed the computer program CYRUS (Schank, 1982, p. 207). CYRUS contained databases of information about two former Secretaries of State, and could integrate new information into these databases. On the basis of the information it gathered, CYRUS could correctly answer such questions as "Have you been to Europe recently?" and "Why did you go there?"

Schematization (Matlin, 1989, p. 118) is a term for the abstraction of experiences and the development of schemata (the plural of schema) in memory. Experimental evidence suggests that people may remember and interpret events according to schemata (see Matlin, 1989, pp. 230-236, for a review). How an experience is abstracted into a schema is a highly individualized process, relying on a person's own history of experiences.

In addition to organizing past experiences in memory, schemata can drive expectations during new experiences (Schank, 1982, p. 37). For example, when visiting a newly opened fancy restaurant for the first time, one may expect to be seated by a maitre d' and to pay a great deal for dinner because of "fancy restaurant" experiences in the past.

3.2.2 Schemata for Conversation

Whereas Schank (1982) had concentrated on schemata involving typical sequences of events for story understanding, Kellermann et al. (1989) observed conversations between undergraduate students meeting for the first time, and described their conversations in terms of scenes in a MOP. Scenes were grouped into three phases: initiation, maintenance, and termination. In the initiation phase, for example, participants might exchange greetings, introduce themselves, and then discuss their current surroundings.

A number of interesting patterns were found in this study. Scenes were weakly ordered within each phase, but strongly ordered between phases, so that a person rarely entered a scene in an earlier phase from a scene in a later phase. Some scenes shared what the investigators called “subroutines” (used in an informal, not computational, sense), or common sequences of generalized acts. In one example, several scenes contained the sequence “get facts”, “discuss facts”, “evaluate facts”, and so on.

Thus, conversations have been represented in terms of schemata. Scenes in these conversations demonstrated some degree of temporal ordering, and could be described as being in a hierarchy with phases and subroutines.

An example of a natural language interface that represented conversation in terms of schemata was JUDIS (Turner & Cullingford, 1989, p. 75). JUDIS was the interface for an interactive system that played the part of a caterer’s assistant and helped the user plan a meal. JUDIS represented its goals and the goals of the user as MOPs. Each MOP contained a list of characters (the caterer and customer), scenes (either mandatory or optional), and the sequence of events. Higher-level MOPs handled higher-level goals, such as the goal of getting information, while lower-level MOPs handled lower-level goals, such as answering yes-no questions.

JUDIS differed significantly from the interface developed later in this thesis because JUDIS tried to “understand” and model the user’s goals, and, consequently, was limited to operating in a narrow domain. The latter approach, by making use of MOPs that have been constructed by the user, can be applied to any domain with which the user is familiar.

3.3 Conversational Schemata in AAC

In this thesis, the notion of schemata and MOPs is explored as a method for organizing the user’s pre-stored text in an AAC system. Several researchers have suggested that scripts might be applied to AAC (Newell, 1990, p. 50; Elder & Goossens', 1994), though nobody, to the author’s knowledge, has taken this idea and developed it into a system. Yet the notion shows promise in giving the AAC user access to large amounts of prestored text in a way that is transparent, since it presumably mirrors the way that information is stored in memory.

The system of schemata developed here is motivated by, and loosely based on, Schank’s hierarchy of MOPs and scenes. The user’s text is stored in a large hierarchy. At the highest level in this hierarchy is the MOP.

3.3.1 MOP

The MOP identifies the conversational topic, or conversational context. In this system, a user defines MOPs to represent topics or contexts that are relevant to him. Each MOP contains a list of scenes (Section 3.3.2) and a list of slots (Section 3.3.3).

Based on Schank's definition of schemata as representing actions, as well as on observations by Kellermann et al. (1989) of actual conversations, scenes are expected to be traversed in the order in which they appear in the MOP. Once the conversation leaves a scene, the expectation is that the scene will not be re-entered. However, some conversational situations may follow the schema more closely than others. Matlin (1989, p. 225) distinguished between "strong scripts" that are ordered and "weak scripts" that are not. During weakly ordered conversational situations, the order in which scenes are traversed may be different from the order in which they appear in the MOP.

MOPs can be organized in a hierarchy, so that lower-level MOPs inherit structures from more general higher-level MOPs. Thus, a scene that occurs in a number of different restaurant MOPs needs to be defined only once, in the common parent MOP.

The actual text of the conversation is contained in the scenes and slots, and these are accessed by first selecting a MOP.

3.3.2 Scene

Formally, a scene contains actions that, together, fulfil a specific goal in a MOP. Looked at in another way, scenes are subtopics of the MOP, and an individual can define them in whatever manner is logical to her. In the conversational setting, associated with each scene is a set of sentences (or, more generally, utterances). These are the sentences that the user may wish to select when discussing the actions, or subtopic, captured in the scene.

Sentences grouped together in a scene are treated, to some extent, as a unit. Child MOPs inherit complete scenes from their parent. Sentences within a scene are expected to be traversed in sequence. These expectations are motivated on the same basis as those for the ordering of scenes in a MOP.

3.3.3 Slot

A particular situation may have associated with it words or phrases that do not form complete utterances. For example, a list of food items may be associated with a particular restaurant. A new schema structure, a slot, is introduced here to group together words or phrases (slot fillers) that are related to each other in the context of the MOP.

3.3.4 Sentence Templates

There may be sentences in a scene that are identical except for one word or phrase, such as the food items in the "order" scene of a restaurant MOP. Rather than listing in a separate sentence each food item that an individual might order, a sentence template could be used. A sentence template contains some words and a reference to a previously defined slot. So, for example, the "order" scene of a restaurant MOP might contain a template: "I'll have a <food slot>." When this template was selected, the system would give the user access to the food slot, to select the specific food items.

This technique would require an additional selection over listing each item in a separate sentence, but reduces the number of sentences that need to be included in the scene. Sentence templates are especially advantageous if the slot contains many items.

To this point, sentence templates and slots here are similar to sentence templates and Minserts in the Liberator.

Another advantage to using a sentence template in this system could be realized if the template occurs in a scene that is inherited from a parent, or higher ancestor, MOP. The same template could be associated with slot fillers contained in each MOP. Thus, the sentence template would appear to be uniform across all related MOPs, while retrieving fillers appropriate to the current MOP.

The use of sentence templates may require the system to perform some natural language processing. For example, if a singular slot filler in a sentence template is filled by a plural filler, morphological endings on other words in the sentences may need to be modified to match the filler number.

Chapter 4

SCHEMATALK: DESIGN AND IMPLEMENTATION

SchemaTalk was designed to complement an individual's existing augmentative communication device, adding a schematic organization of pre-stored and reusable text to the device's own capabilities. In order to investigate issues related to design and implementation, a prototype system was developed and tested at the Applied Science and Engineering Laboratories (ASEL).

The most common way to use SchemaTalk is to enter sentences organized as schemata into a specially designed text file. This may be done using the user's regular AAC system, or using a standard computer keyboard and mouse. Once the schemata are stored, SchemaTalk gives the user access to this pre-stored text in an organized manner, making it easy to step through the schema and select this text as the conversation moves along. The SchemaTalk user may:

(1) select and speak pre-stored text, as is;

(2) select pre-stored text, edit it to more accurately reflect current needs, and speak this edited text;

(3) use the standard AAC system to compose and edit novel text in SchemaTalk, and then speak this text.

For the remainder of the chapter, the term "keyboard" will be used to refer either to a keyboard or to an AAC system emulating a keyboard. Similarly, "mouse" will refer to either a computer mouse or a mouse emulator.

SchemaTalk currently runs on a Sun system, in a UNIX environment. The program code consists of several files of Tcl/Tk scripts, and extensions to the Tcl/Tk library, written in C++. These C++ functions are used to build the internal representation of the user-specified schema hierarchy. The Tcl/Tk script files define the SchemaTalk window and user interface, calling Tcl/Tk functions to display and control various graphical widgets, including dialog boxes, buttons, and scrolling boxes.

4.1 Initial Setup

At start-up, the schema file (see Section 4.5) is loaded and the SchemaTalk window (Figure 4.1) is displayed. This window consists of, from top to bottom:

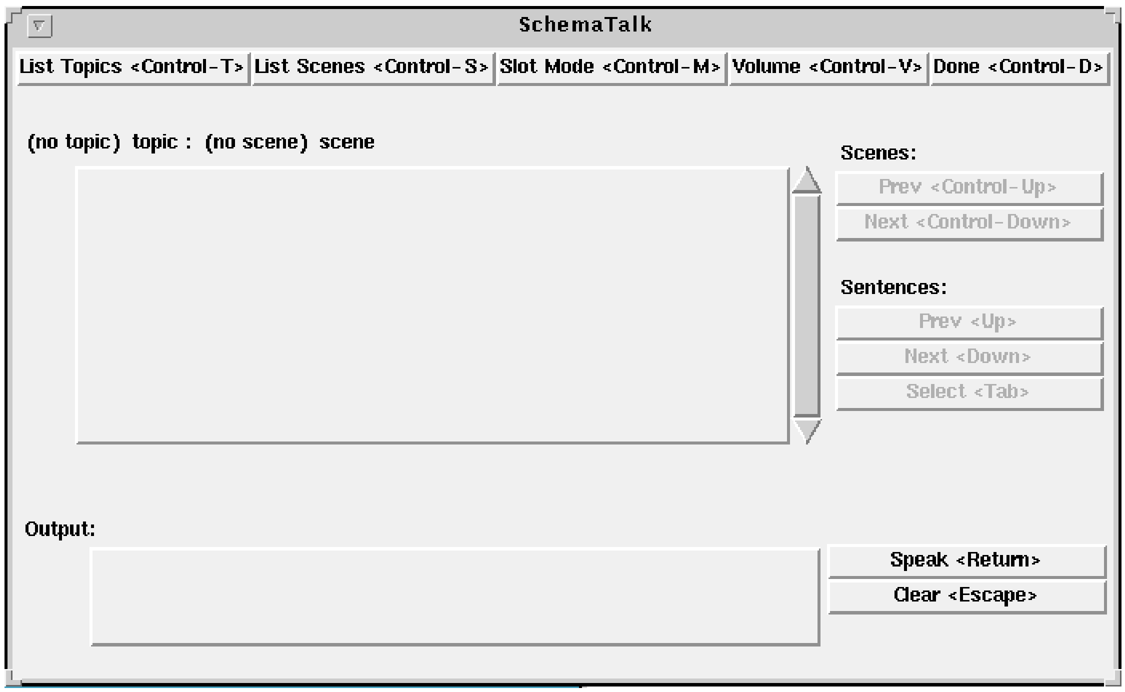

Figure 4.1: SchemaTalk window at start-up

(1) a title bar, containing the title "SchemaTalk", and, on the left, a small triangle in a square that, if pressed using a mouse, will iconify the window (provided by the OpenLook window manager);

(2) a menu bar, consisting of a row of menu buttons for listing topics (MOPs), listing scenes, setting the mode, setting speech volume, and exiting;

(3) the schema display area, consisting of a scrolled list box that will contain the schema text, and buttons to its right that are used for navigation through the schema; and,

(4) the speech output area, consisting of the output text box with the heading "Output:", and the buttons "Speak" and "Clear" to its right.

Buttons can be pressed using a mouse, or by entering a key from a keyboard. The key that will “press” the button is specified in the button label. For example, the “Speak” and “Clear” buttons can be “pressed” by entering the “Return” and “Escape” keys, respectively. For the remainder of the chapter, the term “pressing” a button will refer to either pressing a button with a mouse or mouse emulator, or by entering the appropriate key.

All dialog boxes contain two buttons: “OK”, and “Cancel” (e.g., Figure 4.2). Although they do not say so explicitly, these buttons are pressed either with a mouse or by pressing the keys “Return” and “Escape”, respectively. If the dialog box contains a list, the list can be scrolled using a mouse, or the “Up” and “Down” arrow keys on a keyboard. The current selection (highlighted with a black background and white text) is returned and the dialog box disappears when the “OK” button is pressed, or nothing is returned when the “Cancel” button is pressed.

Also, all dialog boxes are modal. This means that, while a dialog box is present, the SchemaTalk window will not accept any input.

4.1.1 Menu Bar

The five menu bar buttons can be pressed using a mouse, or by pressing the “Control” key followed by a key specific to the button. If the “Control” key is pressed and released, all of the menu buttons are highlighted (black background, white text), and pressing the button-specific key presses the button. If the “Control” key was pressed by mistake, pressing it again will release the “Control” highlight.

4.1.1.1 “List Topics” Button

Pressing the “List Topics” button (or “Control” and “T”) causes the “List Topics” dialog box to be displayed (Figure 4.2), listing all MOPs contained in the schema file. Approximately the first 20 characters of each entry in a dialog box is visible. Indentation indicates the hierarchical structure, with the names of child MOPs indented two spaces to the right of their parents.

A MOP is selected by scrolling to the topic name (using “up” and “down” arrow keys, or controlling the scroll bar with a mouse) and pressing the “OK” button (or “Return” key). The schema list box is then updated to contain sentences from this new MOP, which becomes the current MOP.

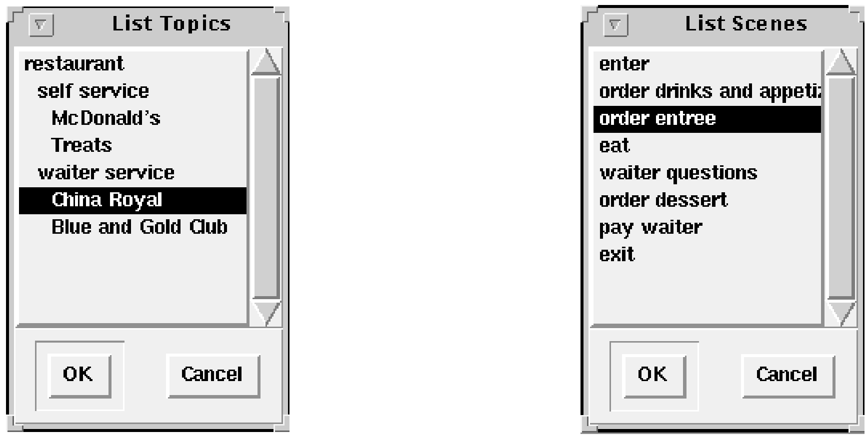

Figure 4.2: "List Topics" (left), and "List Scenes" (right) dialog boxes

4.1.1.2 "List Scenes" (or "List Slots") Button

The SchemaTalk window is initially in scene mode (Section 4.2), and the schema list box contains sentences from the current MOP. Pressing the "List Scenes" button (or "Control" and "S") causes the "List Scenes" dialog box to be displayed (Figure 4.2), listing all scenes in the current MOP. The scene selected from this dialog box becomes the current scene, and the schema list box is scrolled so that sentences from this scene are visible.

In slot mode (Section 4.3), the same process takes place, involving slots rather than scenes. A new current slot is selected from the dialog box, and the schema list box is updated to contain all, and only, the fillers from this new slot.

4.1.1.3 "Slot Mode" (or "Scene Mode") Button

Pressing this button causes the mode to toggle from scene mode to slot mode, or slot mode to scene mode (described in Section 4.2 and Section 4.3). The button label indicates the mode that will be active after the button has been pressed: when SchemaTalk is in scene mode, this button is labelled "Slot mode", and when SchemaTalk is in slot mode, this button is labelled "Scene mode".

4.1.1.4 "Volume" Button

The user can adjust the sound level of the speech output by pressing the "Volume" button (or the keys "Control" and "V" together), and scrolling to the desired sound level on the dialog box that appears. Possible sound levels range from zero to the maximum limit, by 10% increments.

4.1.1.5 "Done" Button

When the user is ready to quit from SchemaTalk, pressing the "Done" button (or "Control" and "D") causes a dialog box to be displayed, which asks the user to confirm or cancel this request to quit. Pressing the "OK" button exits from the interface and the SchemaTalk window disappears.

4.1.2 Schema Display Area

The schema display area is initially empty, and the buttons for navigating through the schema are disabled, as indicated by their grayed lettering. When a MOP has been selected, sentences (in scene mode) or fillers (in slot mode) are displayed in the list box. Schema buttons are enabled only if they are appropriate. For example, the "Next" scene button is enabled (the text is black, not gray) only if there is a scene after the current one in the current MOP.

The schema list box displays up to approximately 60 characters per line. (The list box width is based on the average width of a character. Using a proportional font, it is possible that a particular sentence will have more than or less than 60 characters actually visible.)

4.1.3 Speech Output Area

Any text that the user enters will appear in the speech output area's text box. This text box will grow if more text is entered than can fit into its current size, and shrink back to its original size as text is deleted.

To speak the text, the user can press the "Speak" button (or the "Return" key). Text is automatically cleared from the speech output area after it has been spoken, but will be spoken again each time the "Speak" button is pressed before any new text is entered.

Text in the speech output area can also be cleared by pressing the "Clear" button (or the "Escape" key). Pressing the "Delete" or "Backspace" key will delete the last character, while pressing "Control" and "Delete" or "Control" and "Backspace" together will delete the last word, space, or punctuation symbol.

4.2 Scene Mode

In scene mode, the schema list box contains sentences from the current MOP. If SchemaTalk can not find the schema file, or if there are no MOPs in the file, this list box is empty. To illustrate the operation of SchemaTalk, a possible MOP for visiting the Blue & Gold Club restaurant (Figure 4.3) shall be used as an example. The first sentence has been

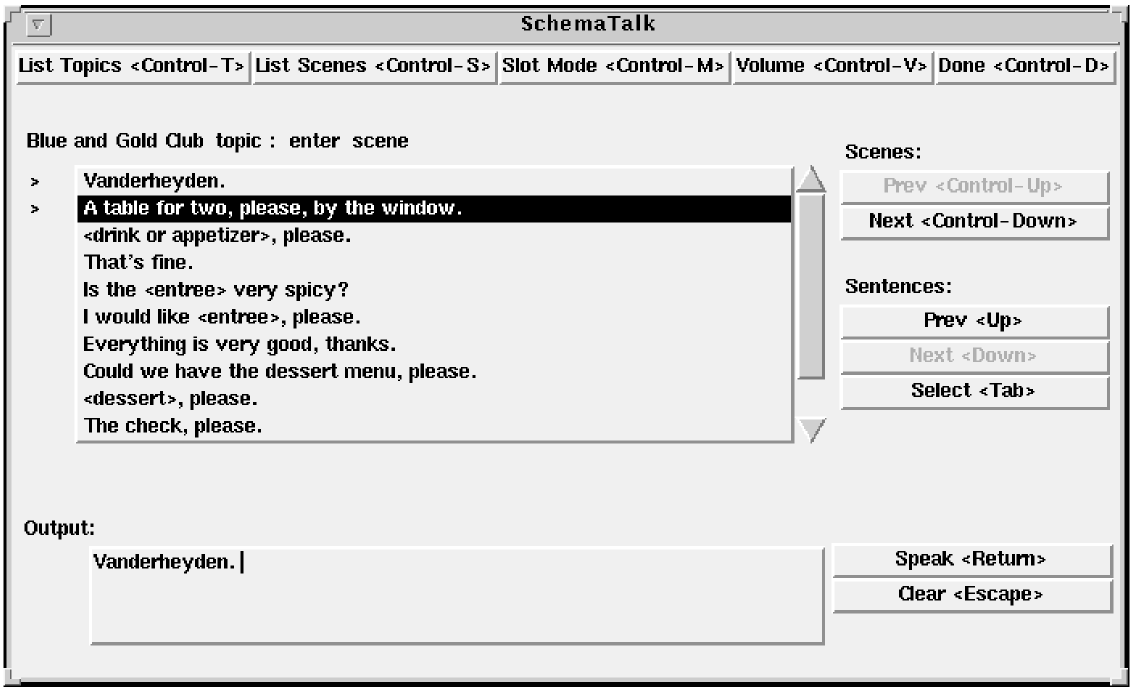

Figure 4.3: The “enter” scene in a possible restaurant schema

selected and appears in the speech output text box. After a sentence has been selected, the next sentence is highlighted, ready to be selected, to further facilitate involvement in conversations that closely follow the schema.

Sentences in the current scene are marked with a “>” in the left margin. A total of 10 sentences are visible at any time. Sentences can be 400 characters long, but only approximately the first 60 characters of each are displayed.

4.2.1 Navigation

The user can move to the next or previous sentence within the current scene by pressing the “Next” and “Prev” sentence buttons, or the “Down” and “Up” arrow keys.

These buttons are disabled when the current sentence is the last or first sentence in the scene, respectively.

Pressing the “Next” or “Prev” scene buttons (or the “Control” and “Down” or “Up” arrow keys) will cause the next or previous scene, respectively, to be made current, and the first sentence of that scene to be highlighted. Alternatively, the user can press the “List Scenes” button and directly select the scene to be made current. The list box scrolls to make the current scene visible.

4.2.2 Sentence Selection

The currently highlighted sentence is selected by pressing the “Select” sentence button (or the “Tab” key). The selected sentence is appended to the text in the speech output text box. If there is insufficient space in the text box, the box grows, and shrinks when the text is deleted.

When a sentence is selected, if it is not the last sentence in the scene, the next sentence becomes current, and the list box is updated to display and highlight the new current sentence.

Text in the speech output area can include both text entered directly from the keyboard, and sentence selected from the schema.

4.3 Slot Mode

In slot mode, the schema list box contains fillers for the current slot. For example, in a restaurant schema, a slot might contain food items for the user to order at a particular restaurant (Figure 4.4).

4.3.1 Navigation

Similar to navigation among sentences and scenes, the user can move to the next or previous filler within the current slot by pressing the “Next” and “Prev” filler buttons, or the “Down” and “Up” arrow keys. These buttons are disabled when the current filler is the last or first filler in the scene, respectively.

Pressing the “Next” or “Prev” slot buttons (or the “Control” and “Down” or “Up” arrow keys) will cause the next or previous slot, respectively, to be made current, and the first filler of that slot to be highlighted. Alternatively, the user can press the “List Slots” button and directly select the slot to be made current.

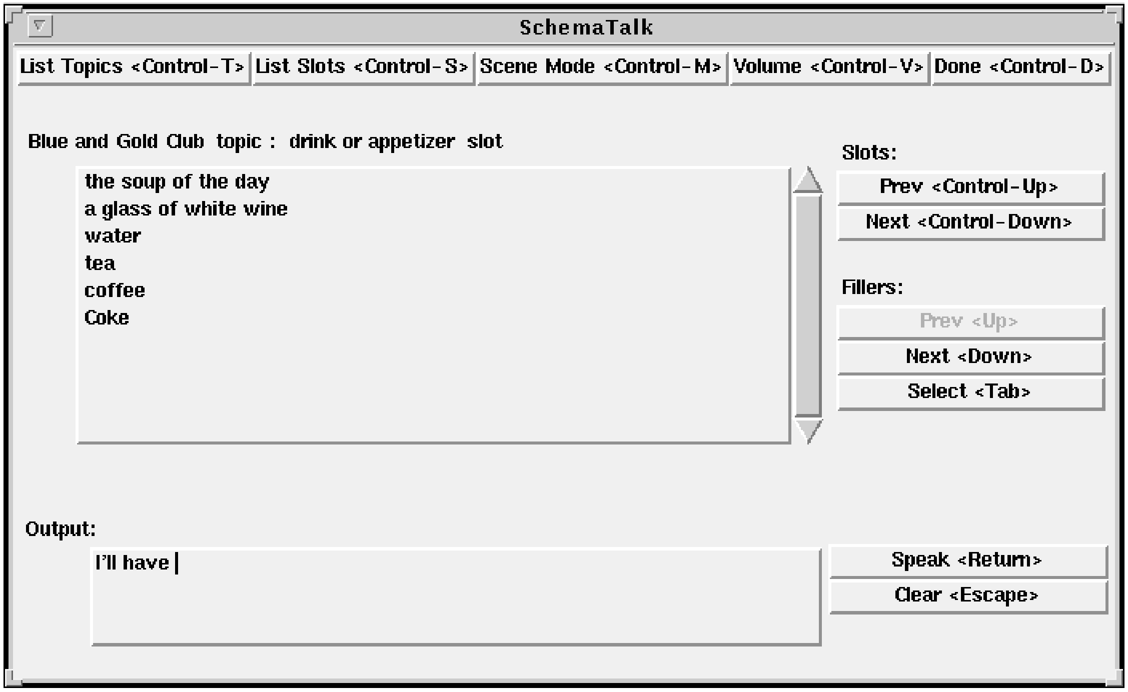

Figure 4.4: The “drink or appetizer” slot in a possible restaurant schema

4.3.2 Filler Selection

The currently highlighted filler is selected by pressing the “Select” filler button (or the “Tab” key). The selected filler is appended to the text in the speech output text box.

4.4 Sentence Templates

A slot can be linked to a sentence using a sentence template. Continuing with the Blue & Gold Club restaurant MOP in scene mode (Figure 4.3), consider the fifth sentence: “Is the <entree> very spicy?” The angle brackets around the word “entree” indicate a slot associated with that position in the sentence, in this case the subject position. When this sentence is selected, the SchemaTalk window changes to slot mode, and the “entree” slot becomes the current slot (Figure 4.5).

The cursor is initially located in the sentence-slot box, which takes the place of the slot name in the template (in this example, “<entree>”) and is empty to begin with. The user can navigate through the fillers in the slot in the normal fashion, using the “Next” and “Prev” buttons, or the arrow keys. Pressing the “Select” button (or “Tab” key) selects the highlighted current filler and appends it to the text at the location of the cursor. Pressing the “Right” arrow key moves the cursor out of the sentence-slot box and to the end of the

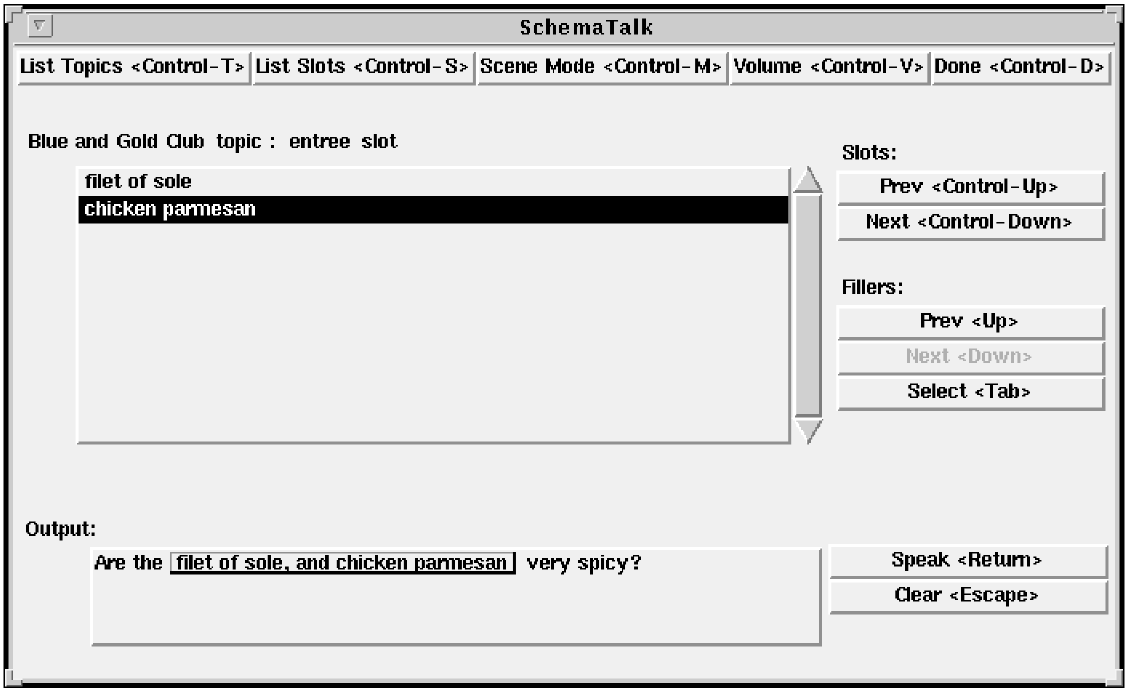

Figure 4.5: A sentence template in a possible restaurant schema
The "entree" slot was made current, after the sentence "Is the <entree> very spicy?" was selected. Then "filet of sole" and "chicken parmesan" were selected from the entree slot.

sentence, while pressing the "Left" arrow key moves the cursor to the end of the text within the sentence-slot box.

The sentence template can be configured to automatically change its form to agree with the text in the sentence-slot box. In this preliminary implementation of SchemaTalk, there is no capability for parsing sentences. However, the user can include more than one sentence form in the schema, and set syntactic and semantic features to match fillers with the correct sentence form. Features are described in Section 4.5.4.

In the example in Figure 4.5, the original sentence form was "Is the <entree> very hot?" When the first filler, a singular "filet of sole", was selected, the speech output text was "Is the filet of sole very hot?" When a second singular filler, "chicken parmesan", was selected, the sentence-slot text formed a plural noun phrase. The plural sentence form was substituted for the singular form, and the text became "Are the filet of sole and chicken parmesan very spicy?"

It should be noted that the "and" between the two fillers was added by SchemaTalk. If yet a third filler were selected, perhaps "rice pilaf", SchemaTalk would replace the “and” with a comma, and place “and” before the last filler, to obtain “Are the filet of sole, the chicken parmesan, and the rice pilaf very spicy?”

4.5 Schema File

The user defines the schemata for SchemaTalk in a text file named “st.schemata”. To illustrate the various schema structures, a schema for visiting restaurants, including the Blue & Gold Club, is developed below. The complete restaurant schema is included in Appendix A.

4.5.1 Grammar

The format of the schema file must conform to the grammar in Figure 4.6, written in Extended BNF (Sethi, 1989, p. 18). Parentheses, “()”, indicate that their contents should be treated as a unit. Square brackets, “[]”, indicate that only one item of their contents should be chosen. The contents of braces, “{ }”, are optional, repeated zero or more times. The terminal **CR** refers to the carriage return.

SchemaFile::=
    { *Comment* | *SchemaDefn* | **WhiteSpace** }

Comment::=
    ‘<comment>’ *Text* **CR**

SchemaDefn::=
    { *MopDefn* | *SceneDefn* | *SlotDefn* }

MopDefn::=
    ‘<mop>’ **WhiteSpace** *MopPathAndName* **CR**
    { *SceneDefn* | *SlotDefn* }
    ‘</mop>’ **CR**

SceneDefn::=
    ‘<scene>’ **WhiteSpace** *SceneName* **CR**
    { *TextSentenceDefn* | *ImportedSentenceDefn* | *SentenceTemplateDefn* }
    ‘</scene>’ **CR**

Figure 4.6: EBNF grammar of schema structures

SlotDefn::=
‘<slot>’ **WhiteSpace** *SlotName* **CR**
{ *TextFillerDefn* | *ImportedFillerDefn* }
‘</slot>’ **CR**

TextSentenceDefn::=
‘<sentence>’ **WhiteSpace** *Text* **CR**

ImportedSentenceDefn::=
‘<sentence>’ **WhiteSpace** ‘{’ *ScenePathAndName* ‘}’ **CR**

SentenceTemplateDefn::=
‘<sentence’ { *PosFeature* } ‘>’ **WhiteSpace** *Text* ‘<’ *SlotName* ‘>’ *Text* **CR**
{ *SformDefn* }

SformDefn::=
‘<sform’ { *NumFeature* } ‘>’ **WhiteSpace** *Text* ‘<’ *SlotName* ‘>’ *Text* **CR**

TextFillerDefn::=
‘<filler’ { *NumFeature* } ‘>’ **WhiteSpace** *Text* **CR**

ImportedFillerDefn::=
‘<filler>’ **WhiteSpace** ‘{’ *SlotPathAndName* ‘}’ **CR**

MopPathAndName::=
MopName | *MopPath MopName*

ScenePathAndName::=
SceneName | *MopPath SceneName*

SlotPathAndName::=
SlotName | *MopPath SlotName*

MopPath::=
‘:’ | *MopName* ‘:’ | *MopPath MopName* ‘:’

MopName::=
Text

Figure 4.6: EBNF grammar of schema structures (cont.)

SceneName::=
    Text

SlotName::=
    Text

PosFeature::=
    ‘ pos=s’ | ‘ pos=o’ | ‘ pos=x’

NumFeature ::=
    ‘ num=s’ | ‘ num=p’ | ‘ num=x’

Text ::=
    { **Letter** | **Digit** | **Punctuation** | **WhiteSpace** }

Figure 4.6: EBNF grammar of schema structures (cont.)

A sentence, sentence form, filler, or identifier name (MOP_NAME, SCENE_NAME, and SLOT_NAME) is limited to 400 characters in length.

The terminals **Letter**, **Digit**, **Punctuation**, and **WhiteSpace** are defined using regular expressions in Figure 4.7. Square brackets, braces, and the vertical line are definition

Letter::=
    [a-zA-Z]

Digit::=
    [0-9]

Punctuation::=
    [!@#$%^&*-=+;’”‘~,.?]

WhiteSpace::=
    { **Space** | **Tab** }

Figure 4.7: Regular expression definitions of schema terminals

symbols and not legal characters. **Space** and **Tab** refer to the space key and tab key, respectively.

4.5.2 The MOP Tree

The schemata read from the schema file are organized in a tree of MOPs. At the top of this tree is the global or root MOP. Scenes and slots are defined within a MOP and are said to be member objects of that MOP. In this case, the MOP which contains these scenes and slots is referred to as their parent MOP. Scenes and slots not defined within a particular MOP are member objects of the root MOP. Similarly, MOPs defined within a MOP become member objects, subMOPs, of the parent. Thus, a MOP contains three (possibly empty) lists: subMOPs, scenes, and slots.

Member objects can be defined in two ways: implicitly, and explicitly. An object is implicitly defined as a member if it is defined within its parent MOP. A waiter-service restaurant MOP, for example, can be defined implicitly as a subMOP of a restaurant MOP (Figure 4.8). To define an object to be a member explicitly, the object is defined outside of

```
<mop> restaurant
        ...
        <mop> waiter service
                ...
                <mop> Blue & Gold Club
                        ...
                </mop>
                ...
        </mop>
        ...
</mop>
```

Figure 4.8: Implicit definition of a schema member object

the body of the parent MOP, and the name of the object is preceded by the path to the parent MOP (Figure 4.9). The path lists the names of all ancestor MOPs of the member, beginning with the most general MOP (but not including the root) and ending with the parent MOP, separated by colons. A second form for the path is a solitary colon, which indicates the parent of the current MOP.

```
<mop> restaurant
        ...
</mop>

<mop> restaurant : waiter service
        ...
</mop>

<mop> restaurant : waiter service : Blue & Gold Club
        ...
</mop>
```

Figure 4.9: Explicit definition of a schema member object

Member scenes and slots can be defined for a MOP in the same way as subMOPs are defined, explicitly or implicitly. However, a MOP that contains no scenes will, by default, inherit all of the scenes of its parent, and a MOP that contains no slots will inherit all of the slots of its parent. On the other hand, no inheritance of scenes (or slots) occurs if any scene (or slot) is defined in the subMOP. Thus, a subMOP resembles its parent MOP, unless defined otherwise.

4.5.3 Organization of Scenes and Slots

Each scene contains a (possibly empty) list of sentences. These sentences may be listed in the scene definition, or they can be imported from a scene of an ancestor MOP. A scene can contain both listed and imported sentences. Similarly, slots can contain both listed and imported fillers.

In the slot example below (Figure 4.10), a drink slot is defined for the Blue & Gold Club restaurant MOP. Braces in the filler text indicate that fillers will be imported, and not listed.

The first “filler” in Figure 4.8 contains a path and slot name, and can be read as “import the fillers of the restaurant MOP drink slot into the current slot.” All of the fillers of the drink slot in the restaurant MOP are then copied into the drink slot of the Blue & Gold MOP, beginning at the location of this first “filler”.

The second “filler” contains a colon and slot name, which can be read as “import the fillers of the parent MOP drink slot into the current slot.” All of the fillers from the drink slot of the waiter service MOP are then copied into the drink slot of the Blue & Gold MOP.

```
<mop> restaurant : waiter service : Blue & Gold Club
        ...
        <slot>                  drink
                <filler>        {restaurant : drink}
                <filler>        {: drink}
                <filler>        Canada Dry ginger ale
        </slot>
        ...
</mop>
```

Figure 4.10: Definition of a slot, containing both listed and imported fillers

The third filler contains text not enclosed in braces, and defines a simple filler.

4.5.4 Sentence Templates and Features

A sentence template contains text and a slot associated with a specific location within that text. For example, “Is the <entree> very spicy?” contains the text “Is the “ and “ very spicy”, with the “entree” slot associated with the subject position of the verb “is”. Fillers and sentence forms can be defined with a number feature (Figure 4.11) to enable automatic generation of sentences with number agreement.

```
<scene> order entree
        ...
        <sentence pos=s>                Is the <entree> very spicy?
                <sform num=s>           Is the <entree> very spicy?
                <sform num=p>           Are the <entree> very spicy?
        ...
</scene>

<slot>                                  entree
        <filler num=s>                  filet of sole
        <filler num=p>                  steamed mussels
        ...
</slot>
```

Figure 4.11: A sentence, sentence forms, and fillers, all with features

Two aspects must be specified:

(1) what the instantiated sentence template should look like for the allowable values of the number feature;

(2) what value for the number feature is associated with the filler (or fillers) in the sentence slot.

For example, Figure 4.11 shows the definition of a sentence template and its associated slot filler definition. The top half of the figure defines the template. the first line shows how the template should be displayed in the SchemaTalk schema display area. Next, the two lines beginning with “<sform” show how the template should be instantiated (what sentence should appear in the speech output area) for fillers of number “s” (singular) and “p” (plural), respectively. The lower half of the figure shows how a number feature value can be associated with different fillers of the slot referred to in the sentence template.

SchemaTalk does the following processing to determine the number feature to be associated with an instantiation. When a sentence template is selected by the user, the instantiation is associated with an initially empty list of sentence slot fillers. SchemaTalk then switches to slot mode, displaying the set of fillers defined for the slot referred to in the sentence template, and the user is given the opportunity to select fillers to fill this slot. The initial instantiation of the selected template is also associated with a number feature, initially empty. When a filler is selected, it is added to the list of sentence slot fillers, and the number feature for the instantiated template is updated to be the same as the number feature associated with the selected filler. The sentence form with the number feature matching the number feature of the instantiated template appears in the speech output text box. If the user selects another filler, this new filler is conjoined to the existing filler(s) in the slot, the feature is updated to “p”, and the sentence form associated with the “num=p” feature in the template definition will appear in the speech output text box.

The sentence template can also have a position feature defined (e.g., “pos=s” in Figure 4.11), though the position feature is not made use of in this implementation. The position feature specifies the semantic and/or syntactic position of the slot in the template. This would be necessary, for example, to ensure that the verb agrees in number with a sentence slot if it appears in the subject position, but not in the object position. In the current implementation, however, the position feature is not used. If the sentence slot is in the object position, no sentence forms are given, and no number agreement is performed.

Chapter 5

AN EVALUATION OF SCHEMATALK: ESPN INTERVIEWS

A series of mock job interviews was conducted in order to make a preliminary evaluation of the SchemaTalk interface, and to gain insight into the development of conversational schemata and the design of a communication interface. In each interview, one participant played the role of job applicant, and was instructed to speak using only the specified communication system. A second participant played the role of an employee of ESPN, a television sports channel, and interviewed the applicant for a job as a sportswriter at ESPN.

In the first interview, the applicant participated with no text pre-stored in schemata. On the basis of this initial interview, schemata were developed to capture the content and structure that the applicant felt was potentially useful for subsequent interviews. Before each of the following interviews, in which applicants used the SchemaTalk interface, schemata were updated and encoded in the system. Of particular interest in examining the result of the interviews are how often applicants chose text from their schemata rather than entering it directly, and whether using the interface enabled them to participate more effectively.

All interviews took place at the Applied Science and Engineering Laboratories (ASEL), with the approval of the University of Delaware Human Subjects Review Board (see Appendix G).

5.1 Participants

Six participants took part in the mock interviews: two played the roles of applicant (A1 and A2), and the remaining four played the role of interviewer (I1 through I4).

The first applicant, A1, was a 39-year-old man, who was always eager to discuss sporting events, sports personalities, and statistics. A1's speech and movement were severely impaired by cerebral palsy. He regularly communicated using a Liberator, as well as gestures and a small amount of vocalization (using a head gesture or vocalization to express yes or no, for example). His communication rate was slow, due to his physical impairments and his frequent self-corrections. A1 produced fairly complex sentences, but had difficulty with spelling and grammar.

The second applicant, A2, was a 29-year-old man, a graduate student on staff at ASEL. A2 was unimpaired in either speech or movement, and not accustomed to communicating using an augmentative device.

Neither applicant had experience in an interview situation, though at least A2 presumably had developed expectations on the basis of his current and past employment. A1 was not and had never been employed, and may not have previously developed such expectations. The dissimilarity of these two SchemaTalk users will discourage any direct comparisons of their results, but will also allow the results to be viewed on more general grounds, and not limited to a specific kind of user.

The interviewers were three men and one woman (I3), all speaking adults with no physical impairments. One interviewer, I1, was nominally acquainted with A1, while the others were not at all. All four interviewers, however, were familiar with A2, as they were all staff at ASEL.

5.2 Equipment

The SchemaTalk interface ran on a Sun SPARCstation IPC, connected to an Apple Macintosh Quadra 900 on the ASEL computer network. SchemaTalk ran in a UNIX environment on the SPARCstation, and was displayed remotely on the Macintosh.

The interface accepts input from a regular keyboard and mouse, or any keyboard emulator. Thus, A1 was able to use his Liberator to access SchemaTalk, and A2 used a mouse to type on an on-screen virtual keyboard program, ScreenDoors (Madenta Communications). Using ScreenDoors, a ceiling of approximately 1 character per second was put on A2's typing rate, a rate more representative of augmented communication.

Voice output was achieved by sending text to an early beta test version of the text to speech synthesizer under development at ASEL, then the speech signal was output over a speaker.

A video camera stood behind the applicant and interviewer, and recorded each interview.

5.3 Procedure

A job description and an announcement of a job vacancy at ESPN were distributed to all participants prior to their interviews. The job was for a sportswriter for the Sports

Center television show on the ESPN sports channel, requiring the writing of news articles about professional sporting events in the tri-state area (Delaware, Maryland, and Pennsylvania).

The applicants were then asked to compose resumes, bearing in mind the job description. Applicants had complete freedom to write anything they wanted in their resume. A1 composed the details of his resume, which was structured and typed up by the author. A2 composed his resume on his own.

Prior to their interviews, interviewers were supplied with the job description, the applicant's resume and covering letter, and a page of typical interview questions (Appendix C, derived from Allen, 1994), to provide consistent guidelines to all.

All participants were made aware of the fact that they would be recorded on video tape, and signed consent forms stating that they understood and agreed to participate.

At the time of the interview, the job applicant and the interviewer were officially introduced to each other, were told that they would be notified when one hour had passed, and the interview began. The interview ended when the interviewer and applicant decided that it was finished, with no limits on time. (All interviews did in fact last between 25 and 65 minutes.)

The four interviews then followed for each applicant, scheduled over four months. The first interview provided a baseline measure, with applicants communicating with no schematized text available (referred to as "direct" selection for the remainder of this chapter). Before each subsequent interview, the applicants were allowed to develop and update their schemata. During these subsequent interviews, applicants were instructed to communicate using SchemaTalk, and to produce utterances either by constructing sentences directly, or by retrieving text from schemata.

The sequence of interviews differed slightly for the two applicants. Prior to the first interview, A1 was given a short practice interview, in order to give him an idea of what to expect. In the first interview, A1 communicated by using his Liberator normally, with no connection to SchemaTalk. In the second interview, he was instructed to use SchemaTalk, accessing it with his Liberator, and the sentences from his practice and first interviews were available to him as a single list in SchemaTalk. For his third and fourth interview, the regular procedure was followed, where he developed and updated his schemata prior to the interview, then accessed SchemaTalk through his Liberator to either construct new utterances or retrieve text from schemata. Transcripts of prior interviews were made available to assist A1 in updating the schemata.

In contrast, A2 was not given a practice interview, because he did not appear to need it. During his first interview, A2 accessed SchemaTalk using the mouse-driven on-screen keyboard, and having no pre-stored text. During the second and third interviews, transcripts were made available to him, and he developed and updated his own schemata.

A1 and A2 were both instructed to communicate using the devices as outlined above, but were not specifically told not to use other modes of communication, such as head nodding or non-verbal vocalizations, and in fact they did use these other modes as well.

5.4 Transcription of Conversational Turns

The videotape of each interview was viewed and transcribed in full by the author. Interviewers' statements were transcribed as accurately as possible from the videotapes. Most statements made by applicants during interviews were available either by printing out the text buffer of A1's Liberator, or from the text file in which SchemaTalk stored its output for both A1 and A2.

The transcribed text was divided into conversational turns (Duncan, 1972; Buzolich & Wiemann, 1988; Kraat, 1987, pp. 78-87). Conversation is in most cases a cooperative interaction between two or more people, during which only one person is usually talking at a time. The person who is speaking is described as holding the conversational turn. The turn holder is temporarily in control of the conversation, either speaking or remaining silent, but in either case inhibiting anybody else from speaking. For another person to become turn holder, the current holder must either relinquish the turn, or the other person must interrupt and take it. Turns are a more natural unit in conversation than sentences. In spoken conversation, incomplete or run-on sentences are the rule and not the exception, and speakers often do not clearly identify the end of one sentence and the beginning of the next.

Turns were defined in the transcripts as ending for the current turn holder and beginning for the new turn holder when the current turn holder fell silent and gave no indication of wanting to continue. In most cases, the interviewer's turn ended with a question, and the applicant then held the turn long enough to respond. By this definition, both conversants always produce the same number of turns (plus or minus one). Several times during the interviews, the interviewer spoke during delays that occurred while text was being spoken by SchemaTalk. In these cases, the interviewer's utterance was treated as a mistake by both participants and thus was not counted as a turn.

In general, turns do not have to contain words. For example, a yes-no question can be answered with a nod or shake of the head. However, a listener may nod to indicate paying attention to the speaker's words, and not answering a question, or even agreeing with the speaker. This back-channel communication is not considered a conversational turn (Buzolich & Wiemann, 1988, p. 4). During the interviews in this study, interview participants occasionally moved off to the side and out of camera range, and any gestures that they might have produced could not be seen. Non-verbal ("other") turns have been included in transcripts when possible, but, for all of these reasons, are not analyzed in detail.

Turns were grouped by the manner in which the words contained within were spoken. They were classified as "schema" if they contained text retrieved from schemata and no other text, and "direct" if they contained only newly input text. They were termed "mixed' if they contained text from both sources, and "other" if they contained no text but consisted of gestures, non-verbal vocalizations, or no apparent response but occurred when the interviewer appeared to pause and wait for one ("no response" turns are discussed in Kraat, 1987, p. 82).

The number of words contained in each turn was counted, as well. A word was defined as any group of characters bounded by spaces, or by punctuation other than a hyphen. This may appear an overly formalized definition for such a simple item as a word, but it allows systematic handling of non-words, including intentional use of abbreviations. Without this definition, a difficulty arises when it is unclear whether a non-word is an abbreviation, or a misspelling, and if it is a misspelling, whether or not the applicant recognized it as such.

Words produced unintentionally, for example by inadvertently hitting the "Speak" key on the device, were not counted, nor were turns composed exclusively of these unintentional words. Unintended utterances also occurred when the sentence previously spoken was accidentally repeated as a result of hitting a wrong key. In many cases, the speaker waved his arms, vocalized, and in other ways made it clear that the text was not intended.

The transcripts for each applicant's last interview are included (see Appendix E), as well as the results derived from all of the interviews (Appendix F). Interview results are summarized for A1 in Table F.6, and for A2 in Table F.10.

5.5 Measures

In conversations with individuals who speak with their natural voices, augmented communicators can be at a disadvantage. As a result of the long time required to form a message on most AAC devices, it is usually the natural-voiced speaker who will utter the first greeting or sentence, and take control of the conversation thereafter. Occasionally taking a more passive role in conversation may be a strategy that is effective and satisfactory for the augmented communicator. However, taking a passive role most or all of the time can impede language development and social interaction.

There may be a strong temptation for both conversants to structure the interchange as a sequence of yes/no questions from the natural-voiced speaker and answers from the augmented communicator. The answers may be verbal, or may simply be gestures for yes and no. This puts the onus on the natural-voiced speaker to guess what the augmented communicator would like to say, and to phrase the information in the form of a question. The augmented communicator finds little need to develop language skills and form sentences in order to communicate needs, likes and dislikes, or more abstract information, when only simple answers are required. Similarly, family members and others familiar with the augmented communicator may become quite effective at guessing the augmented communicator's message after only a few words. The augmented communicator may have little motivation to learn to produce complete and correct sentences in this environment, and will not learn the communication skills necessary to interact successfully with the world outside.

The important question for SchemaTalk, then, is whether it enables the augmented communicator to participate in conversation more effectively. It is not certain, however, what factors determine effective participation, or how these factors vary when different individuals are involved and in different situations. A yes-no question and answer interchange may be highly effective in terms of providing specific facts, but ineffective in terms of developing a social relationship or learning the context around those facts. Effectiveness is at least as much a function of the semantic content of each utterance, as its form, or the structure of the conversation.

The measures described below are not intended to prove or disprove effective participation, but do give some indication of the kind of interaction the took place.

The term "verbal turns" is used to refer to all turns that contained words, or, equivalently, the union of schema, mixed, and direct turn types.

5.5.1 Turn Count

The number of turns (Figure 5.1) varied considerably across interviews for A1. Interviews lasted nearly 57 minutes on average (ranging between 53 and 61 minutes), and the number of verbal turns ranged from 8 to 23 (mean: 13.25). During each interview, a large number of turns were non-verbal. The majority (78%) of verbal turns in interview #2 involved retrieval of pre-stored text from schemata (in schema and mixed turn types), but only 37% in interview #3, and 54% in interview #4.

For A2, the situation was different. Interviews lasted between 25 and 41 minutes (mean: 33.33 minutes), and included between 19 and 26 verbal turns (mean: 22.33). Although very few non-verbal turns were recorded, this may be due to the fact that participants often moved outside of the camera area.

From the transcripts, it appears that the interviewers generally took control of the interactions, which seems appropriate for their role. The total number of turns, then, may be more indicative of characteristics of interviewers than applicants. However, the propor-

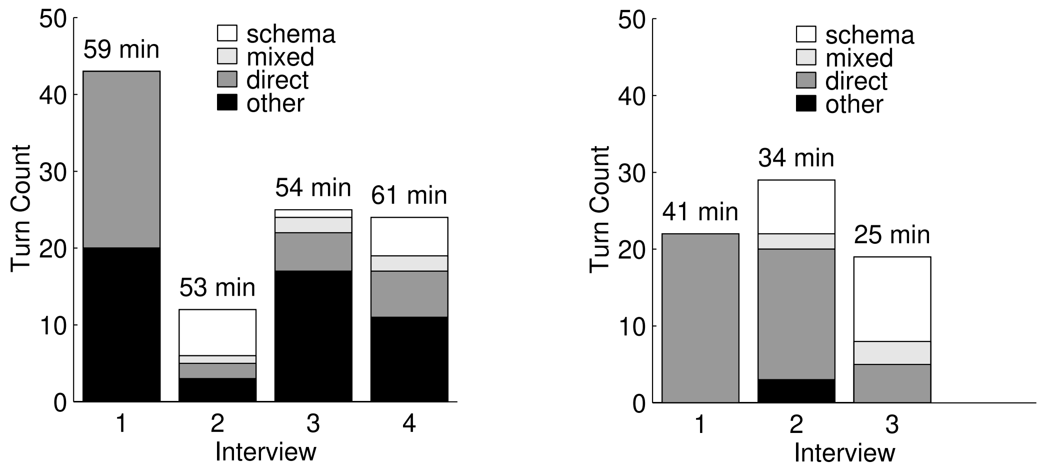

Figure 5.1: Turn count
The number of conversational turns during each interview, for A1 (left) and A2 (right); by turn type, with the total duration of the interview written above.

tion of turn involving pre-stored text from schemata provides an indication of the degree to which applicants' schemata accurately predicted the type of questions interviewers would ask, as well as greetings and other utterances of the applicants' own choice. A1 used pre-stored text in almost all of this turns in interview #2, in almost no turns in interview #3, and in roughly half of his verbal turns in interview #4. Evidently, A1 had not expected many of the questions asked by the interviewer in interview #3. A2, however, used pre-stored text during an increasing number of turns in each interview.

5.5.2 Word Count

As well as the number of turns, however, it is important to consider the number of words produced in each turn. Figure 5.2 shows the average number of words in each verbal turn, by turn type.

Considering A1's information first, interview #1 contained the most turns, but the shortest turns. Turns grew longer with each interview, and turns involving text retrieved from SchemaTalk were longer than turns containing only text entered from his Liberator. Mixed turns were the longest. It is possible that A1 deliberately pre-stored longer text in order to maximize the savings in time, or it may be that situational demands simply encourage his production of shorter novel messages during a conversation.

A2's data were again very different. While mixed and direct turns increased in length in interview #3, schema turns decreased. Mixed turns again contained the most

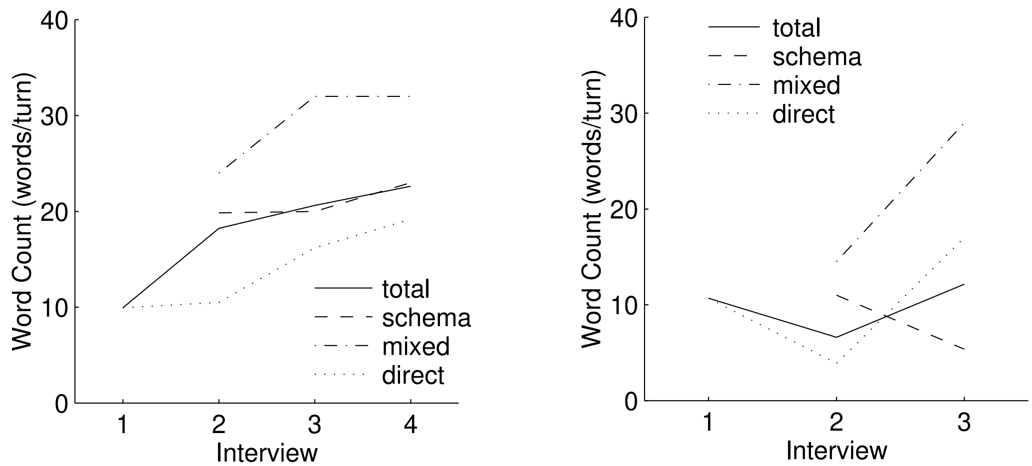

Figure 5.2: Word count (words/turn)
The average number of words per turn, for A1 (left) and A2 (right); by turn type.

words. From the transcripts of interview #3, it was evident that A2, unlike A1, had decided to store both short and long messages in his schemata, leading to the drop in schema turn word counts. In the schema file for interview #3 (Section D.2.2), for example, A2 has a "greetings" scene that contains single words and short phrases for responses to questions.

5.5.3 Turn Duration

As mentioned earlier, a turn ended for the current turn holder and began for the new turn holder when the current turn holder fell silent and gave no indication of wanting to continue. To investigate SchemaTalk's goal of reducing the time required to produce a response using an AAC device, the average duration of turns (Figure 5.3) was calculated. For both A1 and A2, schema turns were produced in a shorter amount of time than either direct turns or mixed turns. Mixed turns, which contained the most words of the different turn types, took the longest time to produce.

It is unclear how much of any reduction in turn duration was due to practice answering interview questions or practice using the interface, and how much was due to improvements to the schemata.

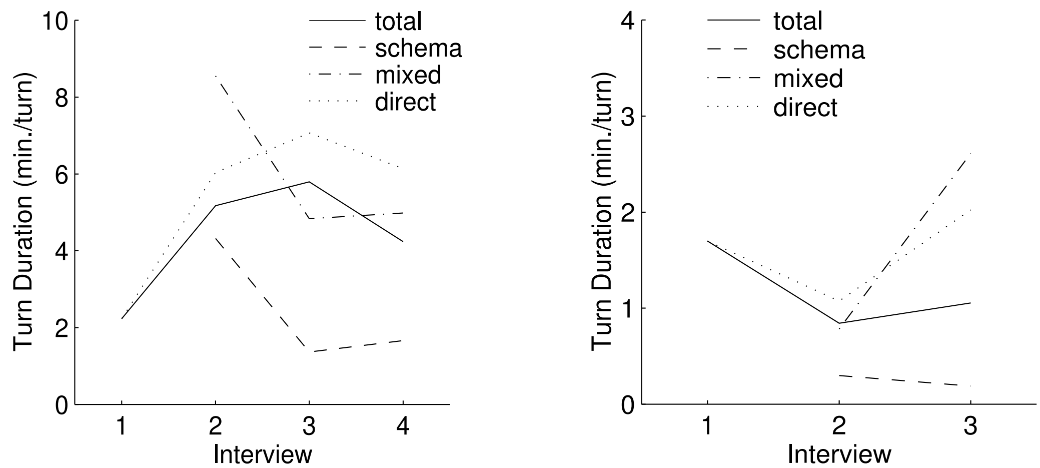

Figure 5.3: Turn duration (minutes/turn)
The average duration of conversational turns of each turn type, for A1 (left) and A2 (right); separated by turn type.

5.5.4 Speech Rate

Finally, the speech rate (Figure 5.4), or the number of words produced per minute,

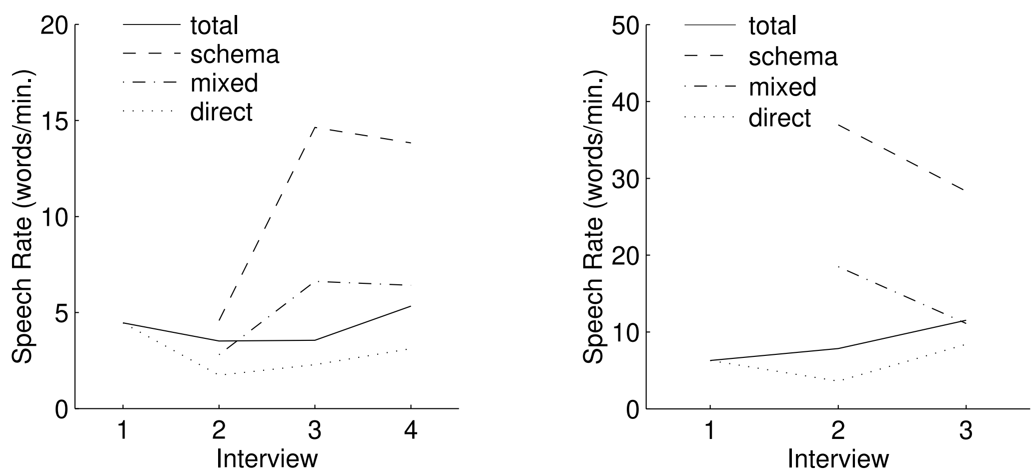

Figure 5.4: Speech rate (words/minute)
The average rate of word production during turns of each verbal turn type, for A1 (left) and A2 (right); separated by turn type.

was computed for each turn type. Foulds (Foulds, 1987) observed that the speech rate of augmented communicators can be far lower than that of speech during conversation by individuals speaking with their natural voices. Augmented communicators spoke at a rate of 2 to 15 words per minute, compared to rates of 150 to 250 words per minute for natural-voiced speakers.

A1's overall speech rate was approximately 4 words per minute (ranging from 3.56 to 5.33, with a mean of 4.22). A2, who does not normally use an AAC device, and whose typing rate was artificially limited, produced words at nearly 9 words per minute (ranging from 6.29 to 11.53, with a mean of 8.55). The speech rate was highest during schema turns and lowest for direct turns.

5.6 Observations

The SchemaTalk interface was designed with the goal of being easy to use from a keyboard-like or mouse-like input device. The Liberator has the ability to function as a keyboard and mouse, but to someone unfamiliar with the necessary steps, it can be quite complicated. A1 needed first to decide what he wanted to say. For a direct turn, he then had to think of how to produce this text using the Liberator, a task with which he was familiar but which took some effort and time, as evidenced by his speech rate of under 5 words per minute in interview #1. To produce a schema or mixed turn, A1 needed to remember the different configuration of keys used to operate the Liberator as a computer keyboard, and then how to operate SchemaTalk to retrieve and produce the text. That A1 found this difficult, especially in interviews #1 through #3, could be seen by problems that he encountered, including producing unintentional turns, attempting to select functions in SchemaTalk that were not appropriate, and showing obvious signs of frustration.

Comments were sought from participants after each interview. Interviewers mentioned that the slow speech rates of both applicants precluded some regular interview procedures. An interviewer normally attempts to ask questions of both breadth and depth, covering different subjects, but also asking a number of questions in each topic. Here, interviewers were less likely to investigate any one subject deeply by following up on applicants' responses with more detailed questions. This was even the case on the occasions when the interviewer did not understand the applicant's response, or when it seemed incorrect or inappropriate.

The responses of A1 were judged inappropriate by interviewers more often than the responses of A2. It is difficult to be certain to what degree this was due to A1's lack of familiarity with interviews, impairments in his language abilities, or possibly his use of personal strategies for reducing message production time, as when he used abbreviations. When text was retrieved from schemata and used inappropriately, it is uncertain whether A1 was doing this deliberately, choosing the rapid production of a near-appropriate message over the lengthy production of an appropriate message, as one interviewer had suggested. However, when asked whether he had used some pre-stored sentences to save time even if they were not perfectly appropriate, A1 responded that he had not.

As well, interviewers mentioned having difficulty understanding the speech output of SchemaTalk. On occasion, interviewers read the message from the computer screen in order to understand what was being said. The same was true for messages spoken by the Liberator in A1's practice and first interview, but less frequently so.

Chapter 6

DISCUSSION AND CONCLUSIONS

6.1 Discussion

In the evaluation study, both applicants appeared to have no problem organizing their sentences into scenes. Scenes that they constructed focussed either on a subtopic of the conversation, or a stage of the conversation (e.g., A2's "greetings").

Neither applicant chose to include slots or sentence templates in his schemata. Although this could be interpreted as implying that slots or templates are somehow less intuitive, it could also simply be the case that the questions interviewers asked did not encourage the use of slots in applicants' responses.

Although the evaluation study was limited and no definite conclusions can be drawn from the results, it was encouraging to observe the much higher word rate that was achieved on turns that used SchemaTalk.

6.2 Some Possible Improvements

The opportunity for reflection naturally brings with it suggestions for how SchemaTalk and the evaluation study could be improved.

6.2.1 Extensions to SchemaTalk

To begin with, the need for two changes to the current design of SchemaTalk became apparent during the evaluation study. Whereas an original design goal was to provide the user with a mode in which the current MOP could be traversed with the need to select from as few as one or two keys or buttons, in its current state, SchemaTalk requires at least six. For a user like A1, who is controlling SchemaTalk remotely from an AAC system, a smaller number of keys would reduce the time spent searching for the correct one.

The second change concerns the creation and updating of the schema file. In the current implementation of SchemaTalk, the user must enter all of the text in the schema file prior to it being available for use. Although this gives the user complete control over what is included in the schemata, it also requires a fair bit of time and attention to the schema file grammar specifications (Section 4.5.1). A sibling application, tentatively called “SchemaEdit”, could be developed to assist the user in creating and updating this schema file. During sessions of SchemaTalk, the user’s utterances could be collected by SchemaEdit and prepared to be easily added to the current, or some other, scene or slot. When SchemaTalk is not running, SchemaEdit could prompt the user for any desired changes to the schema file, automatically formatting them appropriately.

A third extension to SchemaTalk is not motivated by the evaluation study, in which users each defined a single MOP, but rather by considerations of use in the longer term. As a user adds MOPs to the list of topics, scrolling through this potentially very long list may take considerable time. An alternative representation for choosing the current MOP, and for organizing a large number of MOPs would be helpful. One example is a semantic network that would graphically depict relations between MOPs. Another possibility would be associating a MOP with icons that represent the situation or the location contained in the MOP.

Finally, the natural language processing capabilities used during sentence template instantiation could be improved by integrating a parser and lexical database into SchemaTalk, the first to produce and interpret slot filler combinations, and the second to contain features of slot fillers. This would remove the need for the user to enter all of this information into the schema file.

6.2.2 Evaluation Study

The study to evaluate the current implementation of SchemaTalk presented several unforeseen difficulties.

The majority of participants, both interviewers and applicants, had only limited experience with job interviews. This made it possible to consider how applicants developed their own internal schemata for the interviews. However, this also made it impossible to determine the extent to which changes in the applicants’ performance was due to the text in their schema files, practice in the interview situation, or practice in using SchemaTalk. A different, possibly more structured, conversational context may have been more suited for demonstrating the effectiveness of the approach being investigated with SchemaTalk. As well, for A1, the complexity of controlling SchemaTalk with a Liberator likely contributed to his slow message production, especially in the early interviews.

In a study similar to this one, Todman et al. (1994) considered not only the rate of communication, but also its content and form. Speech acts, topic maintenance, and turn-taking were included among the measures. It would be interesting to compute similar measures of communication quality for the evaluation study described in this thesis. (Unfortunately, differences in the experimental methods prevent a comparison of rate measures in the two studies.)

6.3 Conclusions

The approach to AAC described in this thesis specifically focuses on supporting conversational interaction. An individual can store conversational text, navigate through it, and retrieve it in a form based on the individual's own schematized representation of the conversational situation.

The preliminary implementation, SchemaTalk, produced encouraging results, despite various confounding factors. This suggests that a schema-based technique can facilitate the participation of augmented communicators in conversation.

Furthermore, schemata can be integrated with other AAC strategies and techniques. For example, MOPs and scenes can provide semantic constraints for systems that use natural language processing, by modelling the context of the conversation. A database-type system could benefit from the structure supplied by schemata, while SchemaTalk could benefit from a database approach in situations that do not closely follow a schema.

Finally, a schema approach to the representation of conversational text could support language training. Schemata developed by a facilitator would be used as a guide to selecting and producing appropriate utterances (Elder & Goossens', 1994).

Further work must be done in refining the basic ideas described here, but the study suggests that there is great potential for this technique.

BIBLIOGRAPHY

Allen, J. G. (1994). *Successful job search strategies for the disabled*. New York, NY: John Wiley & Sons.

Alm, N., Arnott, J. L., & Newell, A. F. (1989). Database design for storing and accessing personal conversational material. In J. J. Presperin (Ed.), *Proceedings of the 12th Annual Conference of the Rehabilitation Engineers Society of North America (RESNA), "Technology for the Next Decade"* (pp. 147–148). New Orleans, LA: RESNA Press.

Alm, N., Arnott, J. L., & Newell, A. F. (1992). Prediction and conversational momentum in an augmentative communication system. *Communications of the ACM*, *35*(5), 46–57.

Alm, N., Newell, A. F., & Arnott, J. L. (1987). A communication aid that models conversational processes. In *Proceedings of RESNA 10th Annual Conference* (pp. 127–129). San Jose, California.

Alm, N. A. (1988). *Towards a conversation aid for severely physically disabled non-speaking people*. Unpublished doctoral dissertation, University of Dundee.

American Speech-Language-Hearing Association (1991). Report: Augmentative and alternative communication. *ASHA*, *33*, 9–12. (Suppl. 5).

Baker, B. (1982). Minspeak. *Byte*, (pp. 186ff).

Baker, B. R. & Barry, R. F. (1990). A mathematical model of Minspeak(TM). In *Proceedings of the Second Annual European Minspeak(TM) Conference* Nottingham, England.

Baker, B. R. & Nyberg, E. (1989). Semantic compaction: A basic technology for artificial intelligence in AAC. In *Proceedings of the 4th Annual Minspeak Conference* St. Louis, Missouri.

Beukelman, D. R. & Mirenda, P. (1992). *Augmentative and alternative communication: Management of severe communication disorders in children and adults*. Paul H. Brookes: Baltimore.

Beukelman, D. R. & Yorkston, K. M. (1989). Alternative and augmentative communication application for persons with severe acquired communication disorders: An introduction. *Augmentative and Alternative Communication*, *5*(1), 42–48.

Buzolich, M. J. & Wiemann, J. M. (1988). Turn taking in atypical conversations: The case of the speaker/augmented-communicator dyad. *Journal of Speech and Hearing Research*, *31*, 3–18.

Demasco, P. W. & McCoy, K. F. (1992). Generating text from compressed input: An intelligent interface for people with severe motor impairments. *Communications of the ACM*, *35*(5), 68–78.

Demasco, P. W. & Mineo, B. (1995). AAC technology: Next year and next decade. *Presentation at the Pennsylvania Speech Hearing Language Association*. March 1995.

Duncan, S. (1972). Some signals and rules for taking speech turns in conversations. *Journal of Personality and Social Psychology*, *23*, 283–292.

Elder, P. S. & Goossens', C. (1994). *Engineering training environments for interactive augmentative communication: Strategies for adolescents and adults who are moderately/severely developmentally delayed*. Clinician Series. Birmingham, AL: Southeast Augmentative Communication Conference Publications.

Foulds, R. (1987). Guest editorial. *Augmentative and Alternative Communication*, *3*, 169.

Kellermann, K., Broetzmann, S., Lim, T.-S., & Kenji, K. (1989). The conversation MOP: Scenes in the stream of discourse. *Discourse Processes*, *12*, 27–61.

Kraat, A. W. (1987). *Communication interaction between aided and natural speakers: A state of the art report*. Madison, WI: Trace R&D Center, University of Wisconsin, 2nd edition.

Light, J. (1988). Interaction involving individuals using augmentative and alternative communication systems: State of the art and future directions. *Augmentative and Alternative Communication*, *4*(2), 66–82.

Light, J. (1989). Toward a definition of communicative competence for individuals using augmentative and alternative communication systems. *Augmentative and Alternative Communication*, *5*(2), 137–144.

Light, J. & Lindsay, P. (1992). Message-encoding techniques for augmentative communication systems: The recall performances of adults with severe speech impairments. *Journal of Speech and Hearing Research*, *35*, 853–864.

Matlin, M. W. (1989). *Cognition*. Holt, Rinehart and Winston, Inc., second edition.

McCoy, K. F., Demasco, P. W., Jones, M. A., Pennington, C. A., Vanderheyden, P. B., & Zickus, W. M. (1994a). A communication tool for people with disabilities: Lexical semantics for filling in the pieces. In *Proceedings of ASSETS '94* (pp. 107–114). Los Angeles, CA: ACM.

McCoy, K. F., McKnitt, W. M., Peischl, D. M., Pennington, C. A., Vanderheyden, P. B., & Demasco, P. W. (1994b). AAC-user therapist interactions: Preliminary linguistic observations and implications for compansion. In M. Binion (Ed.), *Proceedings of the RESNA '94 Annual Conference* (pp. 129–131). Nashville, Tennessee: RESNA RESNA Press. mccoy@asel.udel.edu.

Mirenda, P. & Mathy-Laikko, P. (1989). Alternative and augmentative communication applications for persons with severe congenital communication disorders: An introduction. *Augmentative and Alternative Communication*, *5*(1), 3–13.

Newell, A. F. (1990). Using models of human interaction for rate enhancement. In B. A. Mineo (Ed.), *Proceedings of the Visions Conference, "Augmentative and Alternative Communication in the Next Decade"* (pp. 49–51). Wilmington, DE.

Prentke Romich Company (1991). *Liberator with Minspeak software and DECtalk speech*. Prentke Romich Company, 1022 Heyl Road, Wooster, Ohio 44691 USA.

Prentke Romich Company (1992). *Words Strategy(TM) for Liberator: Workshop handout*. Prentke Romich Company, 1022 Heyl Road, Wooster, Ohio 44691 USA.

Sandler, M. H. (1994). Eyeray: A modeling environment for the exploration of the parameters of eye gaze tracking. Unpublished master's thesis, University of Delaware.

Schank, R. (1982). *Dynamic memory: A theory of reminding and learning in computers and people*. Cambridge University Press.

Schank, R. C. & Abelson, R. P. (1977). *Scripts, plans, goals and understanding: An inquiry into human knowledge structures*. Hillsdale, NJ: Erlbaum.

Sethi, R. (1989). *Programming languages: Concepts and constructs*. Addison Wesley.

Stum, G. M. & Demasco, P. (1992). Flexible abbreviation expansion. In J. J. Presperin (Ed.), *Proceedings of the RESNA International '92 Conference* (pp. 371–373). Washington, D.C.: RESNA Press.

Todman, J., Alm, N., & Elder, L. (1994). Computer-aided conversation: A prototype system for non-speaking people with physical disabilities. *Applied Psycholinguistics*, *15*(1), 45–73.

Turner, E. H. & Cullingford, R. E. (1989). Using conversational MOPs in natural language interfaces. *Discourse Processes*, *12*, 63–90.

Vanderheyden, P. B., Pennington, C. A., Peischl, D. M., McKnitt, W. M., McCoy, K. F., Demasco, P. W., van Balkom, H., & Kamphuis, H. (1994). Developing AAC systems that model intelligent partner interactions: Methodological considerations. In M. Binion (Ed.), *Proceedings of the RESNA '94 Annual Conference* (pp. 126–128). Nashville, Tennessee: RESNA RESNA Press. vanderhe@asel.udel.edu.

VanDyke, J., McCoy, K., & Demasco, P. (1992). Using syntactic knowledge for word prediction. Presented at ISAAC-92. Abstract appears in Augmentative and Alternative Communication, 8.

Zickus, W. M. (1995). A software engineering approach to developing an object-oriented lexical access database and semantic reasoning module. Unpublished master's thesis, University of Delaware.

Appendix A

SAMPLE SCHEMA FILE

```
<comment>     This file contains a sample hierarchy of restaurant schemata.


<slot>                                      natural numbers
        <filler num=s>                      one
        <filler num=p>                      two
        <filler num=p>                      three
        <filler num=p>                      four
        <filler num=p>                      five
        <filler num=p>                      six
        <filler num=p>                      seven
        <filler num=p>                      eight
        <filler num=p>                      nine
        <filler num=p>                      ten
        <filler num=p>                      twenty
        <filler num=p>                      fifty
</slot>

<slot>                                      numbers
        <filler num=p>                      zero
        <filler>                            {natural numbers}
</slot>

<mop>                                       restaurant
        <slot>                              drink
                <filler num=s>              water
                <filler num=s>              tea
                <filler num=s>              coffee
                <filler num=s>              Coke
        </slot>
        <slot>                              food
                <filler>                    {drink}
```

```
        </slot>

        <scene>                         enter
                <sentence>              Hello.
        </scene>

        <scene>                         order
                <sentence pos=o>        I'd like <food>.
                <sform num=x>           I'd like <food>.
                <sentence>              Thanks.
        </scene>

        <scene>                         pay cashier
                <sentence pos=n>        Here's <natural numbers> dollars.
                <sform num=s>           Here's <natural numbers> dollar.
                <sform num=p>           Here are <natural numbers> dollars.
        </scene>

        <scene>                         exit
                <sentence>              Thank you.
                <sentence>              Good bye.
        </scene>
</mop>

<mop>                                   restaurant : self service
</mop>

<mop>                                   restaurant : self service : McDonald's
        <slot>                          food
                <filler num=s>          Big Mac
                <filler num=s>          shake
                <filler>                {: food}
        </slot>
</mop>

<mop>                                   restaurant : self service : Treats
        <slot>                          food
                <filler num=s>          pepperoni pizza
                <filler>                {: food}
        </slot>
</mop>

<mop>                                   restaurant : waiter service
```

```
<slot>                              drink or appetizer
        <filler>                    {: drink}
</slot>

<slot>                              entree
</slot>

<slot>                              dessert
</slot>

<scene>                             enter
        <sentence>                  Vanderheyden.
        <sentence>                  A table for two, please, by the window.
</scene>

<scene>                             order drinks and appetizers
        <sentence pos=x>            <drink or appetizer>, please.
                <sform num=x>       <drink or appetizer>, please.
        <sentence>                  That’s fine.
</scene>

<scene>                             order entree
        <sentence pos=s>            Is the <entree> very spicy?
                <sform num=s>       Is the <entree> very spicy?
                <sform num=p>       Are the <entree> very spicy?
        <sentence pos=o>            I would like <entree>, please.
                <sform num=x>       I would like <entree>, please.
</scene>

<scene>                             eat
</scene>

<scene>                             waiter questions
        <sentence>                  Everything is very good, thanks.
        <sentence>                  Could we have the dessert menu, please.
</scene>

<scene>                             order dessert
        <sentence pos=x>            <dessert>, please.
                <sform num=x>       <dessert>, please.
</scene>

<scene>                             eat
```

```
        </scene>

        <scene>                         pay waiter
                <sentence>              The check, please.
        </scene>

        <scene>                         exit
        </scene>
</mop>

<mop>                                   restaurant : waiter service : China Royal
        <slot>                          drink or appetizer
                <filler num=s>          egg roll
                <filler num=s>          hot and sour soup
                <filler>                {: drink or appetizer}
        </slot>

        <slot>                          entree
                <filler num=s>          moo goo gai pan chicken
                <filler num=s>          sweet and sour pork
                <filler>                {: entree}
        </slot>

        <slot>                          dessert
                <filler>                {: dessert}
        </slot>
</mop>

<mop>                                   restaurant: waiter service: Blue & Gold Club
        <slot>                          drink or appetizer
                <filler num=s>          the soup of the day
                <filler num=s>          a glass of white wine
                <filler>                {: drink or appetizer}
        </slot>

        <slot>                          entree
                <filler num=s>          filet of sole
                <filler num=s>          chicken parmesan
                <filler>                {: entree}
        </slot>

        <slot>                          dessert
                <filler num=s>          Snickers bar pie
```

<filler> {: dessert}
</slot>
</mop>

Appendix B

ESPN JOB DESCRIPTION

* * *

Job Opening -- Sportswriter for ESPN SportsCenter

* * *

Description

The SportsCenter program provides television viewers with a summary of current sports events and results. SportsCenter is recorded four times daily, and covers professional football, hockey, baseball, basketball, and tennis.

We are looking for one qualified individual to report current sports events in Pennsylvania and the surrounding areas. Support in researching events will be provided by the ESPN research team, but the individual will be responsible for arranging the relevant information into a concise and exciting format.

Qualifications

A minimum of two years experience reporting sports events. Television experience is preferred, but radio or newspaper work is also acceptable. The individual must be self-motivated, knowledgeable and interested in local sports, and must demonstrate an ability to report these events in an engaging style.

Salary

$28,000 per year; negotiable, based on experience

Available

September 1, 1995

Contact

Denise Peischl, General Manager
ESPN

Appendix C

"POSSIBLE INTERVIEW QUESTIONS" FOR INTERVIEWERS

C.1 Career Path

(1) Why are you leaving your present job?

(2) What do you expect to be doing in five years?

(3) What are your long-term career goals?

(4) Why did you change jobs so often?

(5) Why do you want this job?

C.2 Background and Character

(1) Tell me about yourself.

(2) What is your main strength?

(3) What is your main weakness?

(4) What specific strengths would you bring to this position?

(5) What do you do when you have trouble solving a problem?

C.3 Work Style

(1) Do you prefer working as a member of a team, or would you rather work alone?

(2) Do you like to work with people?

(3) How do you react to criticism from superiors, if you believe it is unwarranted?

(4) Do you work well under pressure?

(5) Are you able to work alone without direct supervision?

(6) What three areas of your job do you like the least?

C.4 Availability

(1) Are you able to work overtime if necessary?

(2) Do you have a good work attendance record?

(3) May we contact your present employer?

(4) Do you have any questions?

C.5 Disability-Related

(1) Do you think you'll fit in with the rest of the staff?

(2) Will your disability interfere with your attendance?

(3) Will you require adaptive equipment to perform your work?

(4) (a) Do you think you'll require assistance from your coworkers to perform job functions?

(b) My staff is not trained to work with disabled people.

Appendix D

APPLICANT SCHEMA FILES

D.1 Schema Files of A1

D.1.1 Schema File of A1 for Interview #2

my name is __(name)__
I did like hearing people ideas abouts sports
when I did worked at 610 w i p. I did take calls from people calling in
television
when I worked for k y w t v I did my writing myself
they had people to help with writing but I wanted to say my own words myself.
any but not golf
no I would like more calls about tennis
pete sampras
The Davis Cup was in Florida that year in December
I love to go to any thing
I am like a kid in toys are us when some I am at a sport
michael chang & andre agassi for United States
I know vic b
I was happy they won the n l
like eighty and eighty-six
people getting mad at m w
with out him wouldn't no play off
he didn't NEED one more
game six of n l. in ninety three
when ashe die
he said what he thought
my love of sports and knowing who to call for in put
game six of 93 n l
to long for air time
that is I why I want this job
you have the time to do sports the right way
after game six of 93 world series I had one on one c s

writing therefore for e s p n may be helping chess along to a sport
yes I had been bad man
for Washington D.C. I did it
I will have to live in Connecticut
my name is jeff endler
yes the add in the u,s,a today
I has lived in that part of United States 35 of thirty nine years of my life
other four years my family was in Connecticut
I try to read five or six papers a day and last year I got myself a dish and I love to go sports myself
I am like a kid in toys are us when I am at big sport
from 90 to 93 I did worked at k y w television giving sport
I always wrote my words myself
they had person they would write I didn't used them
I felt it part of my job to keep my lines to sports people but it wasn't in my job outline
ninety
the n l only three hits all 3 was homerun 3-2
I would like to do some chess
week
my parents looked fast
knowledge and lines to sports people
telephone and going to sports
begin talking
talk to long
Canada
went I did television before I did sit with the person when they did that it was my ideas they did the work
if I move will you pay for a dish and my telephone call to keep my lines to sport people

D.1.2 Schema File of A1 for Interview #3

```
<mop>                           sports interview

        <scene>                 introducing myself
                <sentence>      My name is __(name)__.
                <sentence>      I am 39.
        </scene>

        <scene>                 about old jobs
                <sentence>      I had two jobs.
                <sentence>      My first one was at KYW-TV giving sports on the news.
                <sentence>      I always wrote my words myself.
```

<sentence> They had people to help with writing, but I wanted to say my own words myself.
<sentence> I thought people could tell if I didn't.
<sentence> After three years of just having a few minutes a day, I wanted to try talk radio.
<sentence> From 93 until now, I am at 610-WIP radio.
<sentence> I did take calls from people calling in.
<sentence> I did like hearing people ideas about sports.
</scene>

<scene> what I am good at and why
<sentence> I am good at knowing where to go to get an answer to a question.
<sentence> Many sports people know and like me.
<sentence> They are willing to help me on and off the air.
<sentence> I am like a kid in Toys Are Us when I am at a big sport event.
</scene>

<scene> what sports I like and why
<sentence> I really like all sports except golf.
<sentence> I like hockey for it is fast moving.
<sentence> I like football for its team work.
<sentence> I love basketball for its movements with or without the ball.
<sentence> I like baseball for its thinking.
<sentence> I think tennis is pretty. It and basketball are my most favorite.
</scene>

<scene> why I want this job
<sentence> When I did worked on television before, I had just a few minutes on the air.
<sentence> I know I will not be on the air much but I will have to write longer stories.
</scene>

<scene> about moving and the dish
<sentence> If I move, will you pay for a dish and my telephone calls to keep my lines to sport people?
</scene>

</mop>

D.1.3 Schema File of A1 for Interview #4

<mop> sports interview

<scene> introducing myself
<sentence> My name is __(name)__.
<sentence> I am 39.
<sentence> I am happy you are taking your time seeing me about this job.
</scene>

<scene> about old jobs
<sentence> I had two jobs.
<sentence> My first one was at K Y W - T V giving sports on the news.
<sentence> I always wrote my words myself.
<sentence> They had people to help with writing, but I wanted to say my own words myself.
<sentence> I thought people could tell if I didn't.
<sentence> After three years of just having a few minutes a day, I wanted to try talk radio.
<sentence> From 93 until now, I am at 610 - W I P radio.
<sentence> I did take calls from people calling in.
<sentence> I did like hearing people ideas about sports.
<sentence> After two years of radio I had got tired of some people crazys ideas.
<sentence> I am ready for something new.
</scene>

<scene> what I am good at and why
<sentence> I am good at knowing where to go to get an answer to a question.
<sentence> Many sports people know and like me.
<sentence> They are willing to help me on and off the air.
<sentence> I am like a kid in Toys Are Us when I am at a big sport event.
</scene>

<scene> what sports that I like and why
<sentence> I really like all sports except golf.
<sentence> I like hockey for it is fast moving.
<sentence> I like football for its team work.
<sentence> I love basketball for its movements with or without the ball.
<sentence> I like baseball for its thinking.

<sentence> I think tennis is pretty. It and basketball are my most favorite.
</scene>

<scene> why I want this job
<sentence> When I did worked on television before, I had just a few minutes on the air.
<sentence> I know I will not be on the air much but I will have to write longer stories.
<sentence> I like the idea you do sport s joined all day long.
</scene>

<scene> about moving and the dish
<sentence> If I move, will you pay for a dish and my telephone calls to keep my lines to sport people?
</scene>

</mop>

D.2 Schema Files of A2

D.2.1 Schema File of A2 for Interview #2

<mop> INTERVIEW

<scene> GREETINGS / CLOSINGS
<sentence> Hi! Nice to meet you.
<sentence> Thank you for this interview.
<sentence> Have a good day.
<sentence> I am looking forward to talking with you again.
</scene>

<scene> EDUCATION
<sentence> I graduated in the top 5 percent of my class at Shippensburg.
<sentence> I helped supervise three other news writers.
</scene>

<scene> EXPERIENCE
<sentence> My main strength's include concise and interesting writing, especially editorials, and good organizational skills.
<sentence> My main weakness is a lack of experience working with a

large staff.
<sentence> Probably my best work was a week-long series on how endorsements have changed college football.
</scene>

<scene> EXPECTATIONS
<sentence> My eventual goal is to become a national sports commentator.
<sentence> I am looking for opportunities to gain national exposure and experience.
<sentence> To leave my current position I would need to receive a 10 to 15 percent increase in my salary.
</scene>

<scene> QUESTIONS
<sentence> What are the possibilities for advancement within the organization?
<sentence> In what community would I be located?
<sentence> What would my initial responsibilities include?
</scene>

</mop>

D.2.2 Schema File of A2 for Interview #3

<mop> INTERVIEW

<scene> GREETINGS / CLOSINGS
<sentence> Yes.
<sentence> No.
<sentence> That would be fine.
<sentence> That sounds promising.
<sentence> No, not really.
<sentence> Hi! Nice to meet you.
<sentence> Thank you for this interview.
<sentence> Have a good day.
<sentence> I am looking forward to talking with you again.
</scene>

<scene> EDUCATION
<sentence> I graduated in the top 5 percent of my class at Shippensburg.
<sentence> I helped supervise three other news writers.

<sentence> Working on the Gazette was a requirement of my major.
</scene>

<scene> EXPERIENCE
<sentence> My main strength's include concise and interesting writing, especially editorials, and good organizational skills.
<sentence> My main weakness is a lack of experience working with a large staff.
<sentence> Probably my best work was a week-long series on how endorsements have changed college football.
<sentence> I enjoyed doing interviews and got along well with the players.
<sentence> I primarily covered the Harrisburg Senators.
</scene>

<scene> EXPECTATIONS
<sentence> My eventual goal is to become a national sports commentator.
<sentence> I am looking for opportunities to gain national exposure and experience.
<sentence> Travel is not a problem as I am still single and have no other significant obligations.
<sentence> To leave my current position I would need to receive a 10 to 15 percent increase in my salary.
</scene>

<scene> QUESTIONS
<sentence> What are the possibilities for advancement within the organization?
<sentence> In what community would I be located?
<sentence> Will any travel be involved?
<sentence> What would my initial responsibilities include?
</scene>

</mop>

Appendix E

TRANSCRIPTS OF APPLICANTS' FINAL INTERVIEWS

Table E.1: Legend for Transcripts

schema	pre-stored text selected from SchemaTalk, including: (SS) sentence selected and spoken (output) by SchemaTalk
direct	text input directly by user, including: (LL) input and output on Liberator (LS) input on Liberator, output on SchemaTalk
other	communication by gesture, vocalization, or no response
un	text produced unintentionally; not included in measures

[...]	deleted text, not included in measures
(...), <...>	comments

__(name)__	consistent with subject consent forms, all names of participants have been removed from transcripts.

Table E.2: Transcript of Interview #4 with A1

	Text	Time	Source	Words	Turn
I4:	Um, hi __(name)__. (A1 smiles) I'm __(name)__, and today I'm going to be playing the director of the ESPN Sports Center. (A1 presses some keys on Liberator) Um, my background is in broadcast television as opposed to sports broadcasting, um, so I'm more interested in running a studio than necessarily a sports show. Now, so why don't you start by telling me a little about yourself. (pause) Where are you from? Originally.	0:01:11			
A1:	(A1 moves Liberator closer) (A1 vocalizes "Ya" while typing on Liberator)	0:01:39			1
	My name is __(name)__.		SS	5	
I4:	(speaking during pause between A1's sentences, because 'tts' pauses after periods) Hi __(name)__. (A1 motions with hand)				

Table E.2: Transcript of Interview #4 with A1 (cont.)

	Text	Time	Source	Words	Turn
A1:	(‘tts’ continues turn)				1a[a]
	I am 39.		SS	3	
	most of my life New Jersey and Delaware		LS	8	
	I love sports [some]		LS	3	
	I am happy you are taking your time seeing me about this job.		SS	13	
I4:	It’s my pleasure.	0:05:45			
	Um.				
	(pause)				
	So what’s your favorite sport?				

Table E.2: Transcript of Interview #4 with A1 (cont.)

	Text	Time	Source	Words	Turn
A1:	(this text is unintentional. While the sentences are being spoken, A1 motions, looks at camera; hits keys on Liberator, gestures, vocalizes “Nnn”, looks at I4. I4 smiles)	0:05:52			2
	My name is __(name)__.		un		
	I am 39.		un		
	most of my life New Jersey and Delaware		un		
	I love sports		un		
	I am happy you are taking your time seeing me about this job.		un		
	(A1 starts typing again)				
	I really like all sports except golf.		SS	7	
	(A1 smiles)				
I4:	(speaking during pause between sentences; stops when next sentence starts)				
	(laughs)				
	You don’t like golf?				

Table E.2: Transcript of Interview #4 with A1 (cont.)

	Text	Time	Source	Words	Turn
A1:	[!] I love basketball for its movements with or without the ball. I like baseball for its thinking. I think tennis is pretty. It and basketball are my most favorite. ? (unintentional; repeated by ‘tts’) It and basketball are my most favorite.		LS SS SS SS LS un	0 11 6 12 0	2a[b]
I4:	Basketball and tennis.	0:09:06			
A1:	(nods)	0:09:09	other		3
I4:	(nods) Um. So you, um, I noticed you, you (pause) have done most of your work in the Philadelphia area, and we’re looking for somebody, um, who also covers the rest of Pennsylvania, which, of course, includes Pittsburghs... (A1 smiles, nods) Pittsburgh. What is, what’s your feeling about the cross-state rivalries between some of the Pittsburgh teams and some of the Philadelphia teams?	0:09:10			

Table E.2: Transcript of Interview #4 with A1 (cont.)

	Text	Time	Source	Words	Turn
A1:	(vocalization) “Ha” I think in on hockey and [of] baseball it is great [to] but it [doesn‘t like] is not like [were] boston vs phula [up] in [on] basketball not yet (smiles)	0:09:36	other LS	1 22	4
I4:	What’s that about Boston basketball?	0:18:29			
A1:	(smiles, gestures with hand)	0:18:33	other		5
I4:	(A1 searches for key on Liberator to repeat last sentence) You think the Sixers and the Celtics are a better rivalry?	0:18:38			
A1:	(smiles) (repeated) I think in on hockey baseball it is not boston vs phula basketball not yet (smiles)	0:18:46	 LL	 15	6
I4:	So the Celtics are better than Pittsburgh.	0:18:53			
A1:	(shakes head slightly) (focuses on Liberator)	0:19:02	other		7
I4:	Okay, let’s get to your, uh, work, work experience. Um, tell me a little bit about the call-in talk show that you hosted.	0:19:08			

Table E.2: Transcript of Interview #4 with A1 (cont.)

	Text	Time	Source	Words	Turn
A1:	(repeat) I think in on hockey and baseball it is great but it is not like boston vs phula in basketball not yet	0:19:22	un		8
	(shakes head, smiles)				
	From 93 until now, I am at 610 - W I P radio.		SS	12	
	I did take calls from people calling in.		SS	8	
	I did like hearing people ideas about sports.		SS	8	
	(repeated by tts) I did like hearing people ideas about sports.		un		
I4:	(nods)	0:21:36			
	Did you then respond to their ideas?				
A1:	(nods)	0:21:42	other		9
I4:	Okay.	0:21:44			
	Um, well unforturnately this job is for a writer here, so most of what gets written, you're going to be writing for what we refer to as the talking heads, ...				
	(A1 nods)				
	...the men who actually read during Sports Center.				
	Um, how do you feel about being a behind the scenes player in an organization like that?				

Table E.2: Transcript of Interview #4 with A1 (cont.)

	Text	Time	Source	Words	Turn
A1:	(nods)	0:22:04			10
	I am ready for something new.		SS	6	
	(pause; looks at I4)				
I4:	(no response)	0:22:31			
A1:	(as A1 prepares statement, I4 talks aside to C; not transcribed)	0:22:32			11
	I know I will not be on the air much but I will have to write longer stories.		SS	18	
	I like the idea you do sport s joined all day long.		SS	12	
I4:	(nods)	0:24:34			
	You said you liked hearing about people's ideas about sports.				
	That sounds like you'd be interested in commentary, in commentary direction on...				
	(A1 nods)				
	...um, sort of reviewing coaches decisions and Monday, ...				
	(I4 and A1 both smile)				
	...Monday armchair quarterbacking and all that, uh, that deal.				
	Would you prefer to be in the commentary end as opposed to the reporting end, simply reporting what's happening?				
A1:	(nods)	0:25:01	other		12

Table E.2: Transcript of Interview #4 with A1 (cont.)

	Text	Time	Source	Words	Turn
I4:	(nods) You think that would be, be a better...? You like that?	0:25:02			
A1:	(nods)	0:25:06	other		13
I4:	Okay, good. Um, let's see. (and two words mumbled) (pause; reading through pages) You think chess will ever be an Olympic sport? (smiles)	0:25:07			
A1:	[=] no because they have they only [track] thing ever[ ][,] two the years it is same [the idea] idea	0:25:29	LS	15	14
I4:	(laughs) You say they, they come up with that idea every year. (looks at A1)	0:33:12			
A1:	(no response)	0:33:19	other		15
I4:	So it'll never be in.	0:33:20			
A1:	(focuses attention on Liberator) (I4 looks at computer screen) (repeated, intended) no because they have they only thing ever two the years it is same idea	0:33:22	 SS	 15	16

Table E.2: Transcript of Interview #4 with A1 (cont.)

	Text	Time	Source	Words	Turn
I4:	(mumbles something, like "Too bad.") Um. Now, at ESPN we're a bunch of sportscasters and so we don't know much about disabilities, so you're going to have to educate us. Um, what particular accomodations do you require to perform your job?	0:33:44			
A1:	(smiles) (shakes head slightly)	0:34:05	other		17
I4:	Do you understand the question?	0:34:10			
A1:	(smiles; nods) I [would] think you [think] might have to do some little things but [and] [I was] I don't see it real big problem (smiles)	0:34:19	 LS	 18	18
I4:	(nods) Good, good. Um. Have you, have you found in the past that the staffs accept you,... (A1 smiles) ...and , and what can I do as a manager to help you if they don't.	0:41:01			

Table E.2: Transcript of Interview #4 with A1 (cont.)

	Text	Time	Source	Words	Turn
A1:	(pause) (before SchemaTalk spoke output, A1 gestured to monitor, and I4 turned to read) [q] if I do get the job I would like [at twenty-ninth] [in] [in] at first to not do anything because I would like see people are like if [of] I NEED some help I want see you	0:41:16	LS LS	23 9	19
I4:	(nods) Okay (A1 smiles) That makes, that makes sense. Let people accept you first. Now, now one of the things that a job interview entails is, is not only my determining what I think of you, but your determining what you think of us. Are there any questions you'd like to ask me about the, the job, um, about what it entails, or, uh, anything else you're, curious about?	0:50:12			

Table E.2: Transcript of Interview #4 with A1 (cont.)

	Text	Time	Source	Words	Turn
A1:	what sports it doesn‘t sound like you will not move me [c] but... If I move, will you pay for a dish and my telephone calls to keep my lines to sport people? < ‘tts’ died for a second on “If I...” > [z]	0:50:41	LS SS LS	12 20 0	20
I4:	(nods) Oh, certainly. (A1 smiles) Any, any calls to keep up with sports are considered business related around here, so telephone is certainly considered part of your, your normal office, office supplies. Interested in what sports you might be... (A1 smiles, nods) ... asked to, to report on?	0:56:33			
A1:	(nods)	0:56:55	other		21

Table E.2: Transcript of Interview #4 with A1 (cont.)

	Text	Time	Source	Words	Turn
I4:	Well we haven't, haven't completely decided but, uh, any, w-it's, uh, it's seasonal, so,... (A1 focuses on Liberator) ...if you want a job full-year, sometimes you're going to have to report on football, sometimes you're going to have to report on baseball,... (A1 smiles) ...since they occupy different halves of the year.	0:56:56			
A1:	(not spoken, I4 turns and reads it from the monitor) Ido love your new adds with art d about football some someup questiogood (smiles)	0:57:11	LS	13	22
I4:	(smiles) You dig the advertisements.	1:01:40			
A1:	(smiling, nods)	1:01:44	other		23
I4:	(nods) Good. Um, any last things you'd like to add, or... (A1 shaking head) ...that, uh...	1:01:45			
A1:	(shakes head)	1:02:01	other		24

Table E.2: Transcript of Interview #4 with A1 (cont.)

	Text	Time	Source	Words	Turn
I4:	Okay, well, I'll discuss your resume with some of the other people around here, and we'll be in touch. Nice meeting you __(name)__... (A1eff smiles, nods) ...and good luck with your career. (I4 gets up)	1:02:02			

a. Continuation of previous turn
b. Continuation of previous turn

Table E.3: Transcript of Interview #3 with A2

	Text	Time	Source	Words	Turn
I3:	< A2’s face was off-screen part of the time > Hi, __(name)__. I’m __(name)__, and I’m the director of personnel for ESPN, and I thought I’d take just a minute and tell you a little bit about what we’re doing, and then I’ll ask you some questions about why you’re interested in us. (A2 nods) As you know, ESPN is an all-sports network, and we are looking for someone to do the writing of material for our Sports Center broadcast four times a day. What you’ll be doing is putting the creative twist on research that other people have done so that it is more attractive to the audience at large. And let me ask you why you were interested in this particular job at ESPN.	1:20:39			
A2:	I am looking for opportunities to gain national exposure and experience.	1:21:19	schema	11	1
	My eventual goal is to become a national sports commentator.		schema	10	
	i have enjoyed writing opinion pieces i[k]n the past and hope to bring this [fl]experience to bear on my future work		direct	21	

Table E.3: Transcript of Interview #3 with A2 (cont.)

	Text	Time	Source	Words	Turn
I3:	Um, I see that you have been a... sports columnist, or are a staff reporter and columnist with the Harrisburg Patriot. Is there some... some event at your current position that's prompted you to look for this position now, or was it just a matter of serendipitous timing that you, um, are applying with ESPN at this point?	1:25:47			
A2:	the timing and opportunity to further my career goals	1:26:09	direct	9	2
I3:	(nods) Great. Tell me what it is about writing that particularly excites you? What, what are the, the elements of being a sports writer that are particularly attractive to you?	1:27:50			
A2:	The ability to express my views in an effective manner as well as the opportunity to convey information about a subject i enjoy	1:28:03	direct	23	3

Table E.3: Transcript of Interview #3 with A2 (cont.)

	Text	Time	Source	Words	Turn
I3:	Well, clearly you write well based on your resume and your experiences. When you are writing for an audience such as our ESPN audience, what kind of an approach will you take with your writing? Do you think the public's interested in, in the facts in a humorous way, do you think they want jokes, do you think they want, um, a more serious approach? What kind of, of a, of a form would your writing take in trying to appeal to our audience?	1:31:02			
A2:	Mostly objective with the use of subtle humor to get my point across in a clear but not heavy-handed manner	1:31:30	direct	20	4
I3:	Great, that's just what our directors like. Would you mind, uh, doing two things for us. One is providing us with some samples of your writing from the past. And the other is, when you and I finish talking... (A2 starts typing) ...would you go down the hall to our research office and they'll have some papers for you, and write us a sample, uh, broadcast, this afternoon while you're here.	1:33:44			

Table E.3: Transcript of Interview #3 with A2 (cont.)

	Text	Time	Source	Words	Turn
A2:	That would be fine.	1:34:07	schema	4	5
I3:	Good. Um, we're a wonderful staff here, you would like us enormously, but there's always things that kind of irritate a person about a working environment, or the kinds of activities they have to do on a day in and day out basis. What are the things that make you crazy about a job, in a negative fashion, the things that irritate you, get your goat, or make you wish you weren't there?	1:34:11			
A2:	My main weakness is a lack of experience working with a large staff. i guess that my hesitation would be an overabundance of meetings and other administrative duties	1:34:34	schema direct	13 15	6
I3:	Well, naturally no one likes those, but you realize that, if you join our staff, fortunately or unfortunately,... (A2 begins typing) you'll be involved with some of those things to a greater or lesser degree,... (A2 nods) depending on the day of the week and the project at hand.	1:36:53			
A2:	Yes.	1:37:06	schema	1	7

Table E.3: Transcript of Interview #3 with A2 (cont.)

	Text	Time	Source	Words	Turn
I3:	Um, tell me without, I don't want you to tell me negatives about where you worked before, but I would like to know what was some of the things, the experiences and attributes of the previous job that you liked especially well. I guess it's your current job.	1:37:09			
A2:	The closeness with the community. also, I enjoyed doing interviews and got along well with the players.	1:37:26	direct schema	6 11	8
I3:	You would be doing some of that here, that's for sure. Um, speaking of the players means, reminds me that there is also activity directly outside the office, not all of it would be writing internally. Is travelling going to be a problem for you?	1:38:29			
A2:	Travel is not a problem as I am still single and have no other significant obligations.	1:38:44	schema	16	9
I3:	Do you like to travel?	1:39:03			
A2:	Yes.	1:39:05	schema	1	10
I3:	What is your time frame for being able to leave your current position?	1:39:15			
A2:	I would prefer giving 6 [e]weeks notice but i am only required to give 2	1:39:20	direct	15	11

Table E.3: Transcript of Interview #3 with A2 (cont.)

	Text	Time	Source	Words	Turn
I3:	That's great. Six weeks would not be a problem for us, we appreciate your desire to clean up your obligations with your current employer. Do you have any questions for me about how we operate, or any other related activities?	1:40:45			
A2:	I would prefer giving 6 weeks notice but i am only required to give 2 (smiles) In what community would I be located?	1:41:00	un schema	 7	12
I3:	For starters, we'd probably have you in Los Angeles. How does that sound to you?	1:41:29			
A2:	that might be o k but i would probably prefer living in the region i would be covering	1:41:35	direct	18	13
I3:	(A2 continues to type) Ultimately you would be in that region. Initially, the Los Angeles position would be some general training and overall understanding of national woodworks.	1:43:23			
A2:	What would my initial responsibilities include?	1:43:34	schema	6	14

Table E.3: Transcript of Interview #3 with A2 (cont.)

	Text	Time	Source	Words	Turn
I3:	Initially, after the training period is over, you would be responsible for producing the on-air commentary for our sports-casters for these four daily broadcasts of Sports Center.	1:43:51			
A2:	(nods) What are the possibilities for advancement within the organization?	1:44:07	 schema	 9	15
I3:	Outstanding! Um, generally people in this position work in the actual production of broadcast materials for... (A2 begins typing) ...oh, probably three to four years, and then there any number of opportunities both in doing research or in the editing areas, or in actual broadcasting, depending on your skills and what you find you par-ticularly enjoy.	1:44:18			
A2:	That sounds promising.	1:44:45	schema	3	16
I3:	Excellent. Um, I'd like for you to do some writing, if you haven't any other questions, or do you have some more questions?	1:44:57			
A2:	No, not really. (repeated by tts) Not really.	1:45:04	schema un	3	17

Table E.3: Transcript of Interview #3 with A2 (cont.)

	Text	Time	Source	Words	Turn
I3:	Great. Well then let's get on to seeing a little bit of what you'd actually be doing if you were working here.	1:45:12			
A2:	Thank you for this interview.	1:45:19	schema	5	18
I3:	It was my pleasure. I look forward to talking to you again.	1:45:26			
A2:	(nods) Have a good day.	1:45:28	 schema	 4	19
I3:	Thank you.	1:45:33			

Appendix F

RESULTS OF INTERVIEWS

Table F.1: Legend for Results

schema	pre-stored text selected from SchemaTalk
direct	text input directly by user
mixed	turn containing both schema and direct text
other	communication by gesture, vocalization, or no response

F.1 Results of Interviews with A1

Table F.2: Results of Interview #1 with A1

Turn	Type				Turn Duration (sec.)			Words		
	schem.	mix.	dir.	other	schem.	mix.	dir.	schem.	mix.	dir.
1			x				2			5
2			x				195			25
3				x			11			
4			x				73			27
5			x				67			4
6			x				71			8
7			x				24			2
8				x			2			
9			x				90			8
10				x			2			
11			x				268			22
12				x			7			
13			x				105			8
14				x			3			
15				x			2			
16			x				77			4
17				x			4			
18				x			1			
19			x				106			8
20			x				87			4
Sum	0	0	23	20	0	0	3078	0	0	229
Total	43				59 (min.)			229		

Table F.2: Results of Interview #1 with A1 (cont.)

Turn	Type				Turn Duration (sec.)			Words		
	schem.	mix.	dir.	other	schem.	mix.	dir.	schem.	mix.	dir.
21			x				233			13
22				x			2			
23				x			5			
24			x				47			5
25			x				230			8
26			x				104			3
27			x				13			3
28				x			5			
29			x				61			5
30			x				197			12
31				x			6			
32			x				101			6
33				x			3			
34				x			2			
35			x				334			23
36			x				276			14
37				x			2			
38				x			15			
39				x			2			
40				x			3			
41			x				234			12
Sum	0	0	23	20	0	0	3078	0	0	229
Total	43				59 (min.)			229		

Table F.2: Results of Interview #1 with A1 (cont.)

Turn	Type				Turn Duration (sec.)			Words		
	schem.	mix.	dir.	other	schem.	mix.	dir.	schem.	mix.	dir.
42				x			4			
43				x			2			
Sum	0	0	23	20	0	0	3078	0	0	229
Total	43				59 (min.)			229		

Table F.3: Results of Interview #2 with A1

Turn	Type				Turn Duration (sec.)			Words		
	schem.	mix.	dir.	other	schem.	mix.	dir.	schem.	mix.	dir.
1			x				492			12
2				x						
3			x				232			9
4	x				405			15		
5	x				156			6		
6	x				75			4		
7	x				109			30		
8	x				486			37		
9				x						
10		x				513			24	
11	x				325			27		
12				x						
Sum	6	1	2	3	1556	513	724	119	24	21
Total	12				53 (min.)			164		

Table F.4: Results of Interview #3 with A1

Turn	Type				Turn Duration (sec.)			Words		
	schem.	mix.	dir.	other	schem.	mix.	dir.	schem.	mix.	dir.
1				x						
2				x						
3			x				334			12
4				x						
5				x						
6		x				269			36	
7			x				519			17
8				x						
9				x						
10				x						
11			x				596			26
12				x						
13				x						
14		x				311			28	
15				x						
16			x				532			19
17				x						
18				x						
19			x				139			7
20				x						
21				x						
Sum	1	2	5	17	82	580	2120	20	64	81
Total	25				54 (min.)[a]			165		

Table F.4: Results of Interview #3 with A1 (cont.)

Turn	Type				Turn Duration (sec.)			Words		
	schem.	mix.	dir.	other	schem.	mix.	dir.	schem.	mix.	dir.
22				x						
23				x						
24	x				82			20		
25				x						
Sum	1	2	5	17	82	580	2120	20	64	81
Total	25				54 (min.)[a]			165		

a. Not including time for problem resolution.

Table F.5: Results of Interview #4 with A1

Turn	Type				Turn Duration (sec.)			Words		
	schem.	mix.	dir.	other	schem.	mix.	dir.	schem.	mix.	dir.
1		x				246			32	
2	x				194			36		
3				x						
4			x				533			22
5				x						
6			x				7			15
7				x						
8	x				134			28		
9				x						
10	x				27			6		
11	x				122			30		
12				x						
13				x						
14			x				463			15
15				x						
16	x				22			15		
17				x						
18			x				402			18
19			x				536			32
20		x				352			32	
21				x						
Sum	5	2	6	11	499	598	2210	115	64	115
Total	24				61 (min.)			294		

Table F.5: Results of Interview #4 with A1 (cont.)

Turn	Type				Turn Duration (sec.)			Words		
	schem.	mix.	dir.	other	schem.	mix.	dir.	schem.	mix.	dir.
22			x				269			13
23				x						
24				x						
Sum	5	2	6	11	499	598	2210	115	64	115
Total	24				61 (min.)			294		

Table F.6: Summary of Results of Interviews with A1

	Turn Type	Interview #1	Interview #2	Interview#3	Interview #4
Number of Turns	schema	0	6	1	5
	mixed	0	1	2	2
	direct	23	2	5	6
	other	20	3	17	11
	total	43	12	25	24
Turn Duration (sec.)	schema	0	1556	82	499
	mixed	0	513	580	598
	direct	3078	724	2120	2210
	total	59 min.	53 min.	54 min.	61 min.
Number of Words	schema	0	119	20	115
	mixed	0	24	64	64
	direct	229	21	81	115
	total	229	164	165	294

F.2 Results of Interviews with A2

Table F.7: Results of Interview #1 with A2

Turn	Source				Turn Duration (sec.)			Words		
	schem.	mix.	dir.	other	schem.	mix.	dir.	schem.	mix.	dir.
1			x				12			1
2			x				8			1
3			x				41			4
4			x				55			6
5			x				88			8
6			x				143			18
7			x				130			11
8			x				120			13
9			x				171			18
10			x				5			1
11			x				218			22
12			x				245			25
13			x				123			12
14			x				160			20
15			x				199			19
16			x				166			18
17			x				89			10
18			x				99			12
19			x				124			10
20			x				33			3
Sum	0	0	22	0	0	0	2242	0	0	235
Total	22				41 (min.)			235		

Table F.7: Results of Interview #1 with A2 (cont.)

Turn	Source				Turn Duration (sec.)			Words		
	schem.	mix.	dir.	other	schem.	mix.	dir.	schem.	mix.	dir.
21			x				6			1
22			x				7			2
Sum	0	0	22	0	0	0	2242	0	0	235
Total	22				41 (min.)			235		

Table F.8: Results of Interview #2 with A2

Turn	Source				Turn Duration (sec.)			Words		
	schem.	mix.	dir.	other	schem.	mix.	dir.	schem.	mix.	dir.
1			x				27			1
2			x				103			6
3			x				95			6
4				x						
5			x				25			2
6				x						
7			x				192			8
8				x						
9			x				58			3
10	x				31			15		
11			x				66			5
12			x				91			6
13			x				45			4
14			x				100			7
15	x				20			14		
16			x				23			2
17	x				14			10		
18		x				58			14	
19			x				90			3
20			x				25			1
21			x				63			5
Sum	7	2	17	3	125	94	1097	77	29	66
Total	29				34 (min.)			172		

Table F.8: Results of Interview #2 with A2 (cont.)

Turn	Source				Turn Duration (sec.)			Words		
	schem.	mix.	dir.	other	schem.	mix.	dir.	schem.	mix.	dir.
22			x				9			1
23	x				31			19		
24	x				15			9		
25			x				25			3
26	x				10			6		
27			x				60			4
28		x				36			15	
29	x				4			4		
Sum	7	2	17	3	125	94	1097	77	29	66
Total	29				34 (min.)			172		

Table F.9: Results of Interview #3 with A2

Turn	Source				Turn Duration (sec.)			Words		
	schem.	mix.	dir.	other	schem.	mix.	dir.	schem.	mix.	dir.
1		x				268			42	
2			x				101			9
3			x				179			23
4			x				134			20
5	x				4			4		
6		x				139			28	
7	x				3			1		
8		x				63			17	
9	x				19			16		
10	x				10			1		
11			x				85			15
12	x				29			7		
13			x				108			18
14	x				17			6		
15	x				11			9		
16	x				12			3		
17	x				8			3		
18	x				7			5		
19	x				5			4		
Sum	11	3	5	0	125	470	607	59	87	85
Total	19				25 (min.)			231		

Table F.10: Summary of Results of Interviews with A2

	Turn Type	Interview #1	Interview #2	Interview#3
Number of Turns	schema	0	7	11
	mixed	0	2	3
	direct	22	17	5
	other	0	3	0
	total	22	29	29
Turn Duration (sec.)	schema	0	125	125
	mixed	0	94	470
	direct	2242	1097	607
	total	41 min.	34 min.	34 min.
Number of Words	schema	0	77	59
	mixed	0	29	87
	direct	235	66	85
	total	235	172	231

Appendix G

APPROVAL LETTER FROM HUMAN SUBJECTS REVIEW BOARD

Insert HSR Board review here

- photoreduce to 72%