AAC-USER THERAPIST INTERACTIONS: PRELIMINARY LINGUISTIC OBSERVATIONS AND IMPLICATIONS FOR COMPANSION

Kathleen F. McCoy, Wendy M. McKnitt, Denise M. Peischl,
Christopher A. Pennington, Peter B. Vanderheyden, and Patrick W. Demasco

Applied Science and Engineering Laboratories
University of Delaware/A.I. duPont Institute

ABSTRACT

Intelligent AAC Systems attempt to provide a communication system that can interpret input from the user in much the same way a familiar listener would. The COMPANSION system is a research demonstration prototype which "interprets" compressed input given by a user of a word based system into a full grammatical sentence. In developing a usable system from the prototype the needs of the user must be specified in well-defined ways.

This paper reports some preliminary observations from an experiment in which word board users interact with their therapist to tell a story from a picture book. The analysis compares the therapist's output with what could be achieved by a system like COMPANSION and discusses the necessary functionality for a second generation prototype as well as some of the potential difficulties that will be faced.

BACKGROUND

In recent years, a number of AAC researchers have attempted to develop techniques and systems that translate symbol or word input into well formed sentences (1, 2, 3). Common to the various approaches is the ability to inflect words (e.g., verb conjugation) and to add function words (e.g., determiners).

In the COMPANSION technique, a primary emphasis was the inclusion of a sophisticated semantic knowledge base and numerically-based heuristics for reasoning about relative word roles (2, 4, 5). For example, the system might take a set of input such as: "<apple> <pear> <eat> <john>" and generate "An apple and a pear were eaten by John". Note that in order to generate such a sentence the machine had to recognize that 'apple' and 'pear' were the things being eaten (recognizing a conjoined theme), and that John was doing the eating. In addition, appropriate determiners (e.g., "a") were added (but not to proper nouns such as John), and the appropriate passive construction was used (requiring the past tense form of "be" and a past participle ending on the main verb) in order to maintain the input order used by the user.

The COMPANSION approach has been implemented as a Lisp-based demonstration running on a Sun Workstation. In developing a second generation prototype that will form the basis for a practical system, it is necessary to validate the inferencing methods previously used and to understand any additional needs for a future product. To accomplish this, we need a methodology for deciding the specific functionality needed. Ideally, we would like our system to act like a familiar human partner does. Thus, this paper attempts to uncover interaction patterns that occur between an AAC user and a listener with an emphasis on the types of linguistic transformations performed in translating word sequences to sentences.

METHOD

Pilot data was collected by transcribing videos originally recorded by van Balkom. Adolescent students with cerebral palsy described pictures in a children's book to their primary speech therapists, using their own manual symbol charts. Four such adolescent-therapist dyads were videotaped and analyzed.

Each student was instructed to describe the pictures as if telling a story to younger children. The therapist was instructed to repeat each word as it was selected by the student, paraphrase the sentence when it was completed, and then ask the student for confirmation that the paraphrased interpretation was correct. A single camera was used to videotape both the student and the therapist. Students took between 11 minutes and one hour to retell their stories.

RESULTS

Some Interactions Consistent with the COMPANSION Approach:

Standard Compansion: Some interactions with the therapist followed the "standard" operation of the compansion system.[1]

S: <girl> <make> <in> <pan> <egg> <breakfast>
T: Girl *will* make *the* egg*s* in *the* pan *for* breakfast.

Here the therapist has added *tense,* and *determiners.* In addition the *plural* form of "egg" was chosen. Though not indicated by the student, the plural form may have been chosen using default knowledge (that people generally eat multiple eggs for breakfast) or it

1. In this and subsequent examples "S" stands for the student input and "T" the therapist. Words/letters added by the therapist are in *italics*. Words of particular interest are in **bold**.

may have been the result of extra-linguistic information (e.g., the picture being described at the time). Notice that the preposition *for* was also included in the expanded message. This addition required reasoning about the semantics of the input sequence. For example, breakfast was the "reason" for making the eggs and should be introduced with a *for* preposition.

Word Order Changes: An assumption of the COMPANSION system has been that the words will be given to the system in the same order that they should be output in a sentence. However, some of our analysis reveals that the therapist sometimes did not follow the word order initially given by the student. The above example falls into this category: the eggs and the pan have been switched in the therapist's output. Consider the following example:

S: <boy> <table> <dusting> <grandmom> <floor> <sweep>
T: Boy *is* **dusting** ***the* table** *and the* grandmom is **sweep*ing*** ***the* floor**.

Notice that in this instance the student is not following a standard subject-verb-object ordering of the words. The therapist changes the order to follow standard English word order (it is not obvious how to form an English sentence while keeping the word order given by the student).

Agent Inference: Another assumption of the COMPANSION system is that a user might omit an agent when referring to him/herself. An agent might also be omitted if it was obvious from context. This behavior was also found in our analysis. Because the story was about a boy and a girl, students sometimes did not specify an agent, yet it was inferred by the therapist:

S: <wash> <clothes>
T: ***They*** *are* wash*ing* clothes.

Verb Inference: Another assumption of the COMPANSION system is that the main verb may be left out in some situations (particularly when the main verb is either *have* or *be)*. We have argued previously that a system must have the ability to reason about which verb is most appropriate in the given situation. Our default rule (i.e., if there is an animate agent and an inanimate object, then the verb *have* should be inferred) is consistent with examples found in the transcripts. Consider the following where both the agent ("they") and the verb ("have") have been inferred.

S: <toys>
T: ***They have*** toys.

Conjunctions: Students sometimes left out conjunctions in the pilot study:

S: <boy> <girl> <made> <bed> <up> <in> <morning>
T: *The* boy ***and*** *the* girl made up *the* bed in *the* morning.

The conjunction could involve the agent role (as above) or other semantic roles:

S: <mom> <help> <shirt> <shoes>
T: Mom*'s* help*ing with the* shirt ***and*** *the* shoes.

Omitted conjunctions were also observed at the sentence level:

S: <girl> <make> <bed>
<boy> <help> <girl> <bed>
T: *The* girl make*s up the* bed ***and*** *the* boy help*s the* girl *make up the* bed.

Possessives: The inference of when a conjunction is necessary is complicated by the need to correctly indicate possessive information. The following example contains an inferred possessive.

S: <boy> <girl> <girl> <clothes> <my> <mother>
T: ***They're*** giving **their** clothes *to* ***their*** mother.

This example is interesting in that it points out several of the difficulties inherent in inferring when a possessive is needed. Note above there was both a conjunction ("boy" and "girl" combined to "they") and two possessives. A possible possessive rule might require that if you want a possessive followed by a noun, just put the two items next to each other (e.g., <girl> <clothes> for "the girl's clothes"). Note here <girl> <clothes> was translated as "***their*** clothes" as if <girl> was now "standing for" the combined agent. However, this strategy was not followed for the second possessive (the strategy would have resulted in <girl> <mother> being used). Rather the student chose the first person possessive pronoun, "my", to indicate the recipient in the message.

It is not clear in the data how much of the therapist's interpretation were influenced by the picture book itself. Nonetheless, it raises important questions about how to determine when a possessive form is desired.

Some Interactions Beyond the Scope of the Current COMPANSION Approach:

Dropped Word (included in interpretation): In some instances the therapist did not include words given by the student in the interpretation even though they often contributed to the intended meaning. Consider:

S: <two> <table> <in> <table> <dining> <room>
T: *There were things on the* table in *the* dining room.

AAC User/Therapist Interaction

Notice that <table> occurs twice in the student's input, but only once in the interpretation. In some sense, the student's input is "linguistically" sound. He is saying two things about a table (a) there are two things on it, and (b) the table is in the dining room. If these two assertions were stated as two separate sentences, then "table" would occur twice. However, as a single sentence there is a way to combine the thoughts without repeating "table". Compare this example with the possessive case above for an illustration of the difficulty in distinguishing this case from that of a possessive.

Replacing a Word (not included in interpretation): In some instances the therapist ignored words selected by the student, even though there was no obvious indication from the student to ignore the word.

S: <girl> **<help>** <clothes> <up>
T: Girl clothes up.
She's ***hanging*** *the* clothes up.

Note that in the above example **<help>** does not occur in the output. The example also shows a case where a new verb has been inferred (probably from the extra-linguistic context).

More Complicated Verb Inference (Adding or Replacing a Word): In some instances the therapist inferred a verb which was not actually included in the input:

S: <boy> <girl> <up> <table> <for> <lunch>
T: ok. *They're setting* up *the* table for lunch.

Dropped Words (not contributing to meaning): In some instances the therapist dropped words from the interpretation:

S: <girl> <look> <at> <to> <boy>
T: *The* girl*'s* look*ing* at *the* boy.

DISCUSSION

Throughout the study, we observed that much of the communication between the student and the therapist was **not** done through the word board. Several of the students were quite apt at getting their meaning across by a combination of vocalizations, gesturing, and pointing (both at the picture book and around the room). Such multi-modal communication is beyond the ability of most AAC systems available today, but would be a fruitful area of future research. However, in developing systems, it is important to understand which of these interactions are critical to the message construction process (e.g., yes/no gestures to confirm or correct interpretations), and develop means to support them in the system's interface.

In the development of "intelligent" AAC systems, it is useful and appropriate to look at user-partner interactions as a source of input into the analysis and design process.The information obtained from such a methodology can provide guidance and justification for the development of system knowledge bases and inferencing mechanisms. We believe that grounding system design in actual human interactions will insure that systems developed will be relevant to the needs of individuals with disabilities.

REFERENCES

[1] Hunnicutt S. Bliss symbol-to-speech conversion: 'Blisstak'. Journal of the American Voice I/O Society 1986;3.

[2] McCoy KF, Demasco P, Gong Y, Pennington C, Rowe C. Toward a communication device which generated sentences. In: Proceedings of the 12th Annual RESNA Conference. New Orleans, LA: RESNA: 1989.

[3] Reich Peter & Shein F. VOICI: A voice output intelligent communication system. In: Presented at the Fourth Biennial ISAAC Conference. 1990.

[4] Jones M, Demasco P, McCoy K, Pennington C. Knowledge representation considerations for a domain independent semantic parser. In: Proceedings of the 14th Annual RESNA Conference. Kansas City, MO: RESNA: 1991.

[5] Demasco PW, McCoy KF. Generating text from compressed input: An intelligent interface for people with severe motor impairments. Communications of the ACM May 1992;35(5):68–78.

ACKNOWLEDGEMENTS

This work has been supported by a Rehabilitation Engineering Center Grant from the National Institute on Disability and Rehabilitation Research (#H133E30010). Additional support has been provided by the Nemours Foundation.

The original data collection was performed in collaboration with Hans van Balkom and Harry Kamphuis, Institute of Rehabilitation Research (IRV), Hoensbroek, The Netherlands. IRV is an institute for research, development and knowledge transfer in the field of rehabilitation and handicaps. The authors would also like to thank the student collaborators and the staff at HMS School for Children with Cerebral Palsy, Philadelphia, for their interest and participation.

Kathleen F. McCoy
Applied Science and Engineering Laboratories
1600 Rockland Road, P.O. Box 269
Wilmington, Delaware 19899 USA
Internet: mccoy@asel.udel.edu