Simple NLP Techniques for Expanding Telegraphic Sentences

Kathleen F. McCoy
CIS Department
University of Delaware
Newark, DE 19716
mccoy@cis.udel.edu

Abstract

Some people have disabilities which make it difficult for them to speak in an understandable fashion. The field of Augmentative and Alternative Communication (AAC) is concerned with developing methods to augment the communicative ability of such people. Over the past 9 years, the Applied Science and Engineering Laboratories (ASEL) at the University of Delaware and the duPont Hospital for Children, has been involved with applying natural language processing (NLP) technologies to the field of AAC. One of the major projects at ASEL (The COMPANSION project) has been concerned with the application of primarily lexical semantics and sentence generation technology to expand telegraphic input into full sentences. While this project has shown some very promising results, its direct application to a communication device is somewhat questionable (primarily because of the computational power necessary to make the technique fast). This paper describes some of the problems with bringing Compansion to a standard communication device and introduces some work being done in conjunction with the Prentke Romich Company (PRC) (a well known communication device manufacturer) on developing a pared-down version of Compansion for people with cognitive impairments.

1 Introduction

Some people have disabilities which make it difficult for them to speak in an understandable fashion. The field of Augmentative and Alternative Communication (AAC) is concerned with developing methods to augment the communicative ability of such people. In addition to problems that make "speaking" difficult, AAC users often have difficulties in coordinating extremities (so typing on a standard keyboard may be impossible and access to large keys is often very slow). Cognitive difficulties may also be present. The field of AAC is concerned with developing methods that provide access to communicative material under reasonable time and cognitive constraints.

Over the past 9 years, the Applied Science and Engineering Laboratories (ASEL) at the University of Delaware and the duPont Hospital for Children, has been involved with applying natural language processing (NLP) technologies to the field of AAC. One of the major projects at ASEL (The COMPANSION project) has been concerned with the application of primarily lexical semantics and sentence generation technology to expand telegraphic input into full sentences (McCoy et al., 1989), (Demasco and McCoy, 1992), (McCoy et al., 1994).

The project can best be thought of as a rate enhancement technique used in the context of a writing tool. Assuming the user is selecting full words at a time (so time of word selection is basically constant and is independent of the number of letters in the word), the technique shows the most gain when used by a linguistically sophisticated user who desires well-formed English constructions. The system speeds rate by allowing the user to select basic content and having the system provide expansions into well-formed sentences. The user may then select among the generated expansions with 1 additional keystroke (for example).

Consider the following input:

```
Mary think 3 watch give John Andrew.
```

expanded as:

```
Mary thinks that the 3 watches were given
        to John by Andrew.
```

Notice that assuming a root word can be selected in a single keystroke and endings added with additional keystrokes, the initial input would take 7 keystrokes, while the expanded version would have required 16.

The Compansion prototype contains three processing modules and requires a great deal of lexical knowledge:

Word Order Parser – encodes a loose grammar of telegraphic sentences and determines and attaches modifiers (e.g., `3` is an adjective which is modifying `watch`), determines part of speech information (e.g., `think` is a verb, `watch` is a noun), and passes sentence sized chunks to the next phase of processing (e.g., first `3 watch give John Andrew` would be passed on to the next phase, and then `Mary thinks` with the result of the previous processing).

Semantic Parser – uses semantic information associated with words to create a semantic representation (Fillmore, 1968), (Fillmore, 1977), (Allen, 1995), (Palmer, 1984), (Hirst, 1987) for each sentence. E.g., verb frame information is associated with each verb. This information indicates which cases the verb is likely to have and the semantic type of words that are likely to fill those cases. Individual nouns have associated semantic types.

Sentence Generator – creates an actual English sentence from the semantic representation (Elhadad, 1991). In this phase the system attempts to keep the word order that was originally input.

While the Compansion prototype is viewed as a promising and successful application of NLP to AAC, it raises some questions when viewed as a practical AAC system.

- Unlimited Vocabulary – Compansion relies on having a large amount of semantic information associated with each word for the processing within the semantic parser. We have been investigating gathering as much of this information as possible through online lexical resources (Zickus, 1995), (Zickus et al., 1995), but much of this information must still be hand encoded. This is particularly true of verbs which are the cornerstone of the semantic reasoning. While some information on noun semantic categorization can be gleaned from online lexical resources such as WordNet (Miller, 1990; Miller, 1995), it is well beyond the state of the art to glean the kind of verb semantics necessary from online resources.
- Sophisticated Grammatical Input – Sophisticated writers are apt to want to use complicated grammatical constructions which may lie outside the processing ability supplied by the Compansion technique. In some instances the fault may be that the system has not yet been programmed to deal with certain constructions (e.g., certain kinds of verb compliments). Such deficiencies can be remedied. Much more serious is the possibility of some grammatical constructions not being understandable in telegraphic form even to a human reader. For example, relative clauses in telegraphic input may be impossible for a human to interpret correctly (at least, unless a great deal of world knowledge information is applied). Thus, there is a limit to the sophistication of grammatical constructions that are possible to disambiguate in telegraphic input.
- Sophisticated Grammatical Output – The sentence generator used by Compansion relies on a grammar which necessarily encodes some limited set of grammatical constructions.

These problems, coupled with the relatively low processing power and space on devices used for AAC, led us to question whether or not NLP is possible in viable AAC devices. Notice that the rate enhancement power of Compansion is heightened when sophisticated linguistic constructions are used. On the other hand, it is exactly the situations where such constructions are most used that the other problem areas of the system are most prevalent.

The question is: Are there uses for techniques such as Compansion that avoid some of the above mentioned problems? Is it feasible and beneficial to provide some kind of "pared-down" version of Compansion on an AAC device?

In conjunction with the Prentke Romich Company (PRC) (a well known communication device manufacturer) we have been working on developing a pared-down version of Compansion for people with cognitive impairments (McCoy et al., 1997). In the next section we motivate the focus on this population. We indicate that not only might the technique prove very useful for this population, but by focusing on this population some of the problems with Compansion can be eliminated. We describe our processing (a simplification of the processing in Compansion) and note some challenges that still remain.

2 The Need: Target Population Description

One population of AAC users that might greatly benefit from expanded telegraphic input are those who are young in age but who suffer from some cognitive impairments which affect their expressive language production. According to (Kumin, 1994) and (Roth and Casset-James, 1989), language production from a population that has such cognitive impairments includes: (1) short telegraphic utterances; (2) sentences consisting of concrete vocabulary (particularly nouns); (3) morphological and syntactical difficulties such as inappropriate use of verb tenses, plurals, and pronouns; (4) word additions, omissions, or substitutions; and (5) incorrect word order. While people exhibiting these kinds of production problems may be understandable, they will often be perceived negatively in both social and educational situations. Therapy is often geared toward developing strategies that overcome or compensate for these production problems. Children who use AAC and have these kinds of difficulties face additional problems over speaking children with the same impairments both because they have additional obstacles in accessing language elements (i.e., language elements must be accessed through a device) and because language and literacy acquisition are not well understood in children who use AAC. Because of this, it is not clear what kind of interventions will be effective with these children.

Our aim is to provide an AAC device which will be useful to this population both by allowing their output to be more standard, and as a potential language intervention (therapy) tool.

3 Challenges with This Population

Before any intelligent device can be developed for this population, the problem of lexical access must be solved. That is, we must find a method that enables the user to select the lexical items that they wish to communicate. The speech output communication aids that PRC designs for commercial use incorporate an encoding technique called semantic compaction, commercially known as Minspeak^R (a contraction of the phrase "minimum effort speech") (Baker, 1982), (Baker, 1984). Minspeak^R is ultimately an abbreviation expansion system, but it is designed to eliminate much of the cognitive load associated with abbreviation expansion. In using abbreviation expansion, users are required to memorize a set of abbreviations that, if typed, will be expanded into full words. For example, the user might type `''t ''` for `''the ''` and `''sch ''` for `''school ''`. Of course, there is a tremendous amount of cognitive load associated with memorizing abbreviations (and with developing memorable abbreviations for a large number of words). Suffice it to say that regular abbreviation expansion is not a viable option for the population being considered here.

Minspeak^R deals with the cognitive load associated with standard abbreviation systems by forming abbreviations using multi-meaning icons rather than letters. Because the icons are rich in meaning and associations, a small number of icons (keys) can be used to represent a large vocabulary where each item can be selected using a memorable sequence of 2-3 icons. The success of the general Minspeak^R paradigm of vocabulary access led PRC to start designing tailored prestored vocabulary programs known as Minspeak Application Programs (MAPs^{TM}) for specific populations of users. These programs are concerned with providing both an appropriate vocabulary and a set of icons appropriately placed on the keyboard so as to allow communication in an "automatic" fashion.

One of these MAPs^{TM}(the Communic-Ease MAP^{TM}) was developed for users chronologically 10 or more years of age with a language age of 5-6 years. This MAP provides access to a basic vocabulary and has proven to be an effective interface for users in our target population. It provides access to approximately 580 single words divided into 38 general categories. Most of these words are coded as 2-icon sequences. The first icon in the sequence (the category icon) establishes the word category. For example, the <SKULL> icon indicates a body part word, the <MASKS> icon indicates a feeling word, and the <APPLE> icon indicates a food word. The second icon denotes the specific word. For example, <MASK> followed by <SUN> produces the word "`happy`"; <APPLE> followed by <APPLE> produces the word "`eat`". The learning and use of icon sequences is facilitated by the incorporation of "icon prediction". In icon prediction the user is "prompted" for valid icon sequences using lights on keys. For example, once the first icon is hit (e.g., <MASK>) lights will appear on icons that lead to a word (e.g., all icons that complete a valid feeling word will be lit).

In addition to the words which are accessed via the icon sequences, Communic-Ease^{TM}contains some morphology and allows the addition of endings to regular tense verbs and regular noun plurals. However, note that to accomplish this, additional keystrokes are required. Also, it is possible to spell words that are not included in the core vocabulary.

In practice, however, users with either slow access methods or poor language ability tend to produce telegraphic messages consisting of sequences of core vocabulary items without embellishing morphology.

Our project builds a processing method on top of the Communic-Ease MAP^{TM}which will expand the telegraphic/mis-ordered input on the part of the user into well formed sentences. Notice that developing a system for this particular population will overcome some of the difficulties faced with the general Compansion system built as a writing tool for people with sophisticated linguistic ability. First, this population will rely on a *limited vocabulary.* As was noted above, the users of this particular vocabulary access system generally use only the vocabulary items programmed into the Communic-Ease MAP^{TM}. While the method allows them to spell any word, in actuality spelling is rather limited and the spelled vocabulary items may be easier to anticipate. Second, the output structures of the sentences will not require sophisticated syntactic constructions. This population requires *limited output structures* comprised primarily of fairly simple sentence structures. Finally, the system will not face the same sorts of input problems described in conjunction with the Compansion system. Again, this primarily follows from the language sophistication of the chosen population.

On the other hand, this population of users does bring with it other difficulties. For example, it is likely that users may produce *unusual sentence input.* While we do not expect to see the same sorts of complications with the input described above with respect to Compansion, it is likely that the input will display unusual characteristics. For example, with linguistically sophisticated users we expect the input word order to mirror the desired surface form. This assumption does not hold for users with cognitive impairments. Thus we must carefully study this population to determine exactly what kind of input to expect.

Other problems faced by this system have to do with the ability of the user to handle the decisions that are required of them. For instance, it may be the case that selecting from a set of expanded sentences may prove very difficult for this population who may become confused or may be unable to retain their desired sentence when given a list of sentences to choose from. Thus it becomes extremely important that the system present *appropriate* expansions and that these expansions be ordered using a heuristic that accurately predicts the most likely expansion.

Because of the cognitive impairments, it is also likely that the user will have a great deal of difficulty if the system acts in unexpected ways. Thus, the system must be extremely *robust* and capable of handling any input given to it.

Finally, the system's interface must be carefully designed so as to make it easy for users to learn and use. In addition, the system must be fast and runnable on relatively inexpensive and portable PC platforms so as to make it cost effective.

4 Simple Techniques

In this project we have decided to collapse the three levels of processing found in Compansion into one level. The system is implemented in C++ for economy of memory and for speed and compatibility considerations.

The major processing in the system takes place in an augmented transition network type of grammar. The network itself encodes a grammar of the telegraphic input expected from this population. The tests in the grammar may be made on the basis of syntactic or semantic features stored in the lexicon on each word. Some of the actions in the grammar are responsible for manipulating a particular register which encodes the "generated string" or expansion associated with each state in the network. Thus these actions are responsible for adding determiners etc. Sets of registers are also maintained for recording semantic aspects of the partial sentence (e.g., information such as what word is the agent). This information is primarily used to reconstruct an appropriate expansion if later input indicates inappropriate decisions were made in earlier states of the parse.

The augmented transition network formalism was chosen for this work mainly because it allowed parallel traversal of all possible parses and therefore the ability to predict next input words. This allowed us to extend the icon prediction mechanism described in the Communic-Ease MAP^{TM}to the word level. For instance, in a situation where the user has typed an adjective, only icons that begin valid next words (e.g., adjective, noun) will be highlighted thus facilitating learning. One can imagine this particular aspect of the system being expanded for therapy sessions. For example, it might be used to teach a user to use standard agent-verb-object sentences by highlighting only words that fit into that pattern.

5 Methodological Issues

Our system functionality has been determined by a collection of transcripts from Communic-Ease^{TM}users. We have collected both raw keystroke data (so that we can establish the range of input we expect from the population) and keystroke data from videotaped sessions where interpretations of the keystroke data are provided by a communication partner. This data allows us to ensure the output from the system is in fact appropriate. In addition it has been used to validate expected sentence structures, validate the expectation that the core vocabulary will comprise most of the input, allow us to better anticipate the spelled vocabulary, and validate input expectations.

Our methodology in using the data has been to partition it into several sets. First, some portion of each of the two kinds of data has been set aside for testing purposes. Thus it is not seen for purposes of system development. Because the videotaped sessions contain both input and its expansion, these are being used primarily as a means for tuning the expansion rules used in the grammar and the appropriateness heuristics that order the expansions produced by the system. We will attempt for the system to mimic the partner on the videotaped sessions.

The raw keystroke data is being used in two ways. Most obviously it is being used to tune the grammar to the range of input. Secondly, some of the raw data is being associated with multiple interpretations deemed reasonable by one of the team members. These interpretations will be used to further tune the grammar expansions.

Several evaluations of the completed prototype system are planned and made possible by the set-aside collected data. First, the robustness of the grammar can be tested by determining the number of completed input utterances found in the collected data that can be handled by the grammar. Second, the appropriateness of the grammar can be tested by determining how often the grammar's output matches the interpretation provided by the communication partner (in the video sessions containing interpretations by the partner) or by a human faced with the same sequence of words (for the raw data to which interpretations have been added).

In addition to the theoretical grammar testing described above, we also plan an informal evaluation of the usability of the system. We plan to iteratively refine the interface by doing usability studies of our prototype with current users of the Communic-Ease MAP^{TM}. We anticipate beginning this testing during the summer and fall of 1997.

6 Conclusions

We have motivated and described a system that is under development via a joint venture between the Applied Science and Engineering Laboratories of the University of Delaware and the duPont Hospital for Children and the Prentke Romich Company. Important features of the effort include a multidisciplinary team with technical expertise in various areas including NLP and clinical expertise with the target population. This effort focuses on a particular user population which enables us to constrain the system processing sufficiently to make the NLP application feasible. Our effort involves designing the system around the specific needs and abilities of the particular population.

Characteristics of the language used by the particular population being studied has permitted us to apply some simple NLP techniques which are proving to be sufficiently robust for this task. We anticipate the addition of some statistical reasoning particularly as a heuristic for ordering possible expansions that the system deems appropriate.

The system is in its prototype stage and currently consists of a PRC Liberator (a standard piece of hardware which provides access to vocabulary items through the Communic-Ease^{TM}MAP) attached to a Pentium-based desktop PC running the user interface with Windows NT and a software-based text-to-speech synthesizer. When the final details of the user interface and the other software components are worked out, the completed prototype will be implemented on a tablet-based PC and will be field tested with current users of the Communic-Ease^{TM}MAP.

7 Acknowledgments

This work has been supported by a Small Business Research Program Phase I Grant from the Department of Health and Human Services Public Health Service, and a Rehabilitation Engineering Research Center Grant from the National Institute on Disability and Rehabilitation Research of the U.S. Department of Education (#H133E30010). Additional support has been provided by the Nemours Foundation.

The author would like to thank Arlene Badman, Patrick Demasco, David Hershberger, Clifford Kushler, and Christopher Pennington for their collaboration on the design and implementation of this project. In addition we thank John Gray for his discussions and implementation of many of the C++ aspects of the system, and Marjeta Cedilnik for her work on the grammar (and transformation rules).

References

Allen, James. 1995. *Natural Language Understanding, Second Edition.* Benjamin/Cummings, CA.

Baker, B. 1982. Minspeak. *Byte*, page 186ff, September.

Baker, B. 1984. Semantic compaction for sub-sentence vocabulary units compared to other encoding and prediction systems. In *Proceedings of the 10th Conference on Rehabilitation Technology*, pages 118–120, San Jose CA. RESNA.

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

Elhadad, Michael. 1991. FUF: The universal unifier user manual version 5.0. Technical report, Columbia University, Computer Science Department.

Fillmore, C. J. 1968. The case for case. In E. Bach and R. Harms, editors, *Universals in Linguistic Theory*, pages 1–90, New York. Holt, Rinehart, and Winston.

Fillmore, C. J. 1977. The case for case reopened. In P. Cole and J. M. Sadock, editors, *Syntax and Semantics VIII: Grammatical Relations*, pages 59–81, New York. Academic Press.

Hirst, Graeme. 1987. *Semantic interpretation and the resolution of ambiguity.* Cambridge University Press, Cambridge.

Kumin, L. 1994. *Communication skills in children with Down Syndrome: A guide for parents.* Woodbine House, Rockville, MD.

McCoy, K., P. Demasco, Y. Gong, C. Pennington, and C. Rowe. 1989. Toward a communication device which generates sentences. In *Proceedings of the 12th Annual Conference*, New Orleans, Louisiana, June. RESNA.

McCoy, Kathleen F., Patrick W. Demasco, Mark A. Jones, Christopher A. Pennington, Peter B. Vanderheyden, and Wendy M. Zickus. 1994. A communication tool for people with disabilities: Lexical semantics for filling in the pieces. In *Proceedings of the First Annual ACM Conference on Assistive Technologies (ASSETS94)*, pages 107–114, Marina del Ray, CA:.

McCoy, Kathleen F., Patrick W. Demasco, Christopher A. Pennington, and Arlene Luberoff Badman. 1997. Some interface issues in developing intelligent communication aids for people with disabilities. In *Proceedings of the 1997 International Conference on Intelligent User Interfaces, IUI97*, Orlando, Florida.

Miller, G. A. 1990. WordNet: An on-line lexical database. *International Journal of Lexicography*, 3(4).

Miller, G. A. 1995. A lexical database for English. *Communications of the ACM*, pages 39–41, November.

Palmer, Martha S. 1984. *Driving Semantics for a Limited Domain.* Ph.D. thesis, University of Edinburgh. Chapter 2: Previous Computational Approaches to Semantic Analysis.

Roth, F. P. and E. Casset-James. 1989. The language assessment process: Clinical implications for individuals with severe speech impairments. *Augmentative and Alternative Communication*, 5:165–172.

Zickus, Wendy M. 1995. A software engineering approach to developing an object-oriented lexical access database and semantic reasoning module. Technical report 95-13, Department of Computer and Information Sciences, University of Delaware, Newark, DE.

Zickus, Wendy M., Kathleen F. McCoy, Patrick W. Demasco, and Christopher A. Pennington. 1995. A lexical database for intelligent AAC systems. In Anthony Langton, editor, *Proceedings of the RESNA '95 Annual Conference*, pages 124–126, Arlington, VA. RESNA Press.