An Evaluation of METAL: the LRC Machine Translation System

Jonathan Slocum
Microelectronics and Computer Technology Corp.

Winfield S. Bennett
Lesley Whiffin
Edda Norcross
Siemens Communication Systems, Inc.

Abstract

The Linguistics Research Center (LRC) at the University of Texas at Austin is currently developing METAL, a fully-automatic high quality machine translation system for market introduction in 1985. This paper will describe the current status of METAL, emphasizing the results of the most recent post-editors’ evaluation, and will briefly indicate some future directions for the system. A 6-page German original text and a raw (unedited, but automatically reformatted) METAL translation of that text into English are included as appendices.

Introduction

The Linguistics Research Center (LRC) at the University of Texas at Austin is currently developing METAL, a fully-automatic high quality machine translation system for market introduction in 1985. This paper will describe the current status of METAL, including the results of the most recent evaluation, and will briefly indicate some future directions for the system. Exhibits A and B (attached) are, respectively, a German original text and a raw (unedited, but automatically reformatted) METAL translation of that text into English.

History and Status

Machine translation research at the University of Texas began in 1956; the LRC was founded in 1961. For much of the history of this project, funding was provided by the U.S. Air Force’s Rome Air Development Center and other U.S. government agencies. In 1979, Siemens AG began funding the development phase of the METAL machine translation system, at which point implementation of the current system was initiated. A prototype has recently been delivered to the sponsor for market testing.

The current system is a unidirectional German-English system, although work to add other target languages, as well as creating an English-German MT system, is now underway. The present staff for the METAL project consists of seven full-time and five half-time personnel.

Application Environment

Software has been developed to handle the formatting problems associated with technical manuals. This software, written in SNOBOL, automatically marks and prepares texts for the METAL translation system (Slocum and Bennett, 1982; Slocum et al., 1984). The only human intervention prior to translation is checking and correcting the results of the automatic formatting routines. Postediting is expected for the output texts. The system does not expect (or provide for) human intervention during the actual translation phase.

Pre-processing and post-editing are presently done on a DEC-2060; the actual translation, on a Symbolics Lisp machine. The “production system” design envisions a Lisp Machine as the translation unit connected to 4-6 translator workstations, from which the prepared texts will be sent to the translation unit and on which the output texts will be postedited.

METAL uses a transfer approach for translation. The entire process consists of four phases: analysis, integration, transfer, and generation (synthesis). The integration phase works with whole parse tree structures, following analysis and preceding transfer. Until recently, transfer and generation were essentially a single phase, but work is currently underway to separate this single phase into two, with a much more powerful generation phase.

Linguistic Component

The current METAL lexicon consists of over 20,000 German and English monolingual entries, consisting of morphological, syntactic, and semantic features and values, and an appropriately large number of transfer entries. The features and values in monolingual lexical entries supply necessary information for the analysis and/or synthesis of these items during the machine translation process. Most entries are reasonably simple, but entries for verb stems are significantly more complex. Inflected adjectives, nouns, and verbs are parsed by word-level grammar rules, with the stems and endings assigned to appropriate lexical categories.
Each transfer lexical entry is a structure equating the source language canonical form with an appropriate target language canonical form. Certain significant transformations (i.e., lexical category, subject area, and preference) is coded in the entry to guide the system in selecting the appropriate translation. Furthermore, tests and operations (including transformations) may be included within transfer entries.

The grammar for METAL consists of over 600 augmented phrase structure rules, each of which is used in both analysis and transfer/generation. METAL's grammar rules are used in the parsing of all levels of structure from the word level to the sentence level, including phrases and clauses. A METAL grammar rule consists of five analysis sections, plus an additional section for each target language: a top line describing the phrase structure (with an optional enumeration of each constituent); a series of restrictions, which test the appropriateness of individual constituents on the right-hand side of the rule; TESTS, which enforce agreement among the right-hand constituents; a CONSTR section, which constructs the analysis of the phrase; an INTEG section, which is executed (once a complete analysis of the sentence is achieved) in order to, e.g., resolve anaphoric references; and one or more target-language-dependent Transfer sections, which control lexical and structural translation into the target language.

Homograph resolution and disambiguation are handled uniformly (i.e., without special passages), in various ways: by orthographic tests, such as the test to ensure that a word that looks like a German noun is not all lower case; by positional constraints, which disallow co-occurrence of ambiguous strings in the same clause location; and, most especially, by the case frame mechanism.

The case (valency) frame mechanism is vital in METAL's analysis of German source language sentences. This mechanism is invoked in clause-level rules and uses features on the verb stem to define the functions of the various central arguments to the predicate. In addition, the case frame mechanism is used to test for such things as subject-verb agreement.

The METAL grammar makes extensive use of transformations to modify structure or perform certain tests. Transformations may be used in the TEST, CONSTR, INTEG, and Transfer phases of the rules; transformations may also be used in transfer lexical entries. A transformation may be written as part of a rule or called by name.

Computational Component

The lexicon for METAL is maintained via a DBMS written in LISP. Input of lexical entries is facilitated by an INTERCODER, a menu-driven system which asks the user for information in English and encodes the answers into the internal form used by the system. An integral part of the INTERCODER is the "lexical default" program which accepts minimal information about the particular entry (root form and lexical category) and encodes most of the remaining necessary features and values. Entries may also be created using any text editor, without the aid of the INTERCODER or lexical defaulter.

Interfacing with the lexical database is done by means of a number of menu-driven functions which permit the user to access, edit, copy, and/or delete entries individually, in groups (using specific features), or entirely. In order to assure a high degree of lexicon integrity the METAL system includes validation programs which identify errors in format and/or syntax. The validation process is automatically used to check lexical items which have been edited, to ensure that no errors have been introduced during editing.

The grammar is also in a database and may be accessed and/or edited in much the same way as the lexicon. System software and named transformations are stored in individual source files.

METAL's parser is a "some-paths, parallel, bottom-up parser" [Slocum et al., 1984]. It may be considered to be "some-paths" because the grammar rules are grouped into numerically indexed "levels" and the parser always applies rules at a lower level before applying rules at a higher level. Once the parser has successfully built one or more Ss at a given level, it will halt; until it discovers one or more S readings, the parser will continue to apply rules in each successive level. Extensive experimentation with the system has found that the present parser configuration is the most efficient one for METAL [Slocum et al., 1984].

Post-Editors' Evaluation

In June, 1984, the METAL system was used to translate 82.6 pages of text into English; the material varied from a sales prospectus (for a speech recognition system) through various general hardware and software system descriptions to highly technical documentation. The output was then edited by two Siemens revisors (one a member of the METAL project, one not). This section describes the revisors' objective performance and subjective reactions (including comparison with earlier versions of METAL) during this experience.
The table above summarizes the editors' revision times. They employed rather different editing techniques (editor #1 working in three passes, #2 in just two), but their times are relatively close.

Comments by Editor #1:

[The 3rd Pass] tends to be concerned with stylistic improvements, formatting changes and typing errors. The last part of this stage involves running the spelling checker on the file to eliminate remaining typing errors.

The impression of post-editing was that there have been many improvements over previous test runs. This was evidenced by the fact that on this post-editing run less than 3% of sentences were re-translated from scratch. The major task in post-editing is now changing word order, changing verb agreement and re-translating the more idiomatic usages. Considerable improvements in format made post-editing easier, although there is still room for further enhancement.

One of the greatest changes affecting post-editing was the fact that since the initial output [compared to earlier versions] of METAL was deemed to have improved, the different stages of post-editing were more clearly defined. That is to say, it was easier to produce an adequate translation during the first run through the text -- using the reformatted output on the screen and a hardcopy of the source text for reference -- than in previous tests. In the second run through a text -- using a hardcopy of the METAL output upon which preliminary post-editing has been performed -- it was easier to concentrate on polishing the translation. In the third and final post-edit stage, one was able to make a final check for stylistic weaknesses, spelling mistakes and typing errors. This was the same method as used in previous tests but one was better able to distinguish between the stages (initial technical and stylistic post-editing; polishing output; final stylistic check) and the entire process was less tiring than in the past.

Although the overall format of the output has improved...there are still [some] problem areas [with the automatic reformatting program].

As an experiment, the unformatted, interlinear [German-English] version was used for the initial post-editing phase. The text was then reformatted...and the 2nd and 3rd phases of post-editing continued as normal. The previous problems with post-editing a highly formatted text meant that whenever a textual change was made in the the text then the format had to be re-modified. The method of post-editing used in this test proved to be considerably faster and easier to handle. ...[The results] demonstrate that the time saving lies in the initial post-edit phase which is when the most changes are made and which is most time intensive with regard to re-formatting text.

Comments by Editor #2:

As compared to the last run in February 1984, the June 84 output showed considerable improvement. A greater number of sentences was useable and many required a change in word order only. Placement of the determiners has been improved. [Certain] points should be considered to improve future translations.

Future Directions

The METAL German-English configuration was released for market testing in January 1985. Current plans are to continue improvement on the present system and to branch off into other target languages, specifically Spanish and Chinese. We estimate that a German-Spanish system should be ready for testing sometime in 1986, with a German-Chinese system sometime thereafter. We have also begun working on an English-German system. If the planned work is successful, work will begin on English-Spanish and English-Chinese MT systems.

References

Slocum, J., and W. S. Bennett, "The LRC Machine Translation System: An Application of State-of-the-Art Text and Natural Language Processing Techniques to the Translation of Technical Manuals," Working Paper LRC-82-1, Linguistics Research Center, University of Texas, July 1982.

Slocum, J., et al., "METAL: The LRC Machine Translation System," presented at the ISSCO Tutorial on Machine Translation, Lugano, Switzerland, 2-6 April 1984. Also available as Working Paper LRC-84-2, Linguistics Research Center, University of Texas, April 1984.
CSE Spracheingabe-Geräte
einführungsschrift

1 Einleitung
Die Entwicklung der Halbleitertechnik, insbesondere der Mikroprozessoren, hat in den vergangenen Jahren neue Perspektiven für die EDV eröffnet. Im Bereich der Spracheingabe wurde mit der Spracheingabe in den Computer ein langlebigerer Mutmach erfüllt. Damit steht für diesen kostenintensiven Zweig der Dateneingabe ein Verfahren zur Verfügung, das die Benutzerfreundlichkeit entscheidend verbessert, zumal die bisherigen Erfassungsmethoden, die sich der Tastatur bedienen, im Prinzip eine Anpassung des Menschen an die Maschine erforderten.
Um die Dateneingabe schneller und sicherer zu machen, mußten die Daten möglichst am Ort ihres Entstehens erfaßt werden. In einer Reihe von Anwendungen, die mit Tastaturen problematisch, wenn nicht gar unmöglich, ist dies der Fall. Wenn der Benutzer mobil sein muß, die Handen für andere Tätigkeiten frei bleiben sollen, oder die Umweltbedingungen für Tastaturen ungeeignet sind.
Die Spracheingabe bietet hier eine ideale Lösung. Gegenüber den traditionellen Erfassungsmethoden zeichnet sie sich durch folgende Eigenschaften aus:
- Optimale Anpassung an die vom Menschen als naturgemäß empfundene Kommunikationsgewohnheiten
- Geringe Einlernzeit
- Leichte Bedienung auch durch ungeübtes Personal
- Mehr Bewegungsfreiheit für den Benutzer beim Erfassen der Daten
- Vereinfachte direkte Datenerfassung
- Grossere Sicherheit der Eingabe.

Die COMPUTER GESELLSCHAFT KONSTANZ bietet folgende Spracheingabegeräte an:
- CSE 1050
- CSE 1060*

Die CSE-Geräte lassen sich an Rechner aller bekannten Hersteller anschließen.

1 Computer-Sprach-Eingabe
2 Technik der Spracherkennung
2.1 Einführung

Unter Spracheingabe verstehen wir die Eingabe von Daten per Sprache in den Computer. Dabei wird das gesprochene Wort durch Sprachabgegeräte, die den Computer vorausgeschickt, in maschinelle verarbeitbare Informationen umgesetzt. Die Entwicklungsvorgänge wird durch ein Sprachsignal zusammengesetzt

In einem Mikrofon umgesetzt. Das Sprachabgegeräte vergleicht daraufhin diese Takte mit den gespeicherten Mustern des Wortschatzes. Wird eine ausreichende Übereinstimmung mit einem der Wörter des Wortschatzes festgestellt, gilt dieses Wort als erkannt. Im andern Fall weist das Gerät die Eingabe zurück.

Der Wortschatz selbst wird zuvor durch ein Training, bei dem jedes Wort mehrfach eingesprochen ist, erarbeitet.

- Die typische Datenfluss bei der Spracheingabe sieht wie folgt aus: Bild 1: Typischer Datenfluss bei Spracheingabe

Die CSE-Spracheingabe-Geräte in Form eines vereinbarten Codes eingesprochenen Worten werden nach ihrer Erkennung vom CSE-Sprachabgegeräte in Form eines vereinbarten Codes eingesprochen an den Computer weitergegeben. Dieser führt die anwendungspezifische Verarbeitung durch. In vielen Fällen ist es dabei nützlich, den Benutzer mitzuteilen, welche Daten im Anwebogramm des Computers angenehmen sind. Diese Rückmeldung wird durch eine optische oder andere Anzeige, z.B. durch ein erleuchtetes Wort, erreicht.

2.1 Der Erkennungsvorgang

Technisch koennen beim Vorgang der Spracherkennung zwei Schritte unterschieden werden:
- Die Verarbeitung des akustisch-phonetischen Signals
- Die Klassifizierung

CSE voice data entry devices

1 Introduction
The development of semiconductor technology, in particular the microprocessors, has opened the new prospects for EDV in the last years. In the range of data acquisition, a long-cherished wish was filled with voice data entry into the computer. With it, a method of operation which improves user convenience is available for this cost-intensive branch of data processing, because required the previous acquisition methods which handle itself of the keyboard in the principle an adaptation of the human being to the machine. In order to make data acquisition more faster and secure, the data must be registered instantly as possible at the place of its generation. In a series of applications, this is problematic with keyboards, if not indeed impossible.

This is the case, if the user must be mobile, ought to remain the hands for other activities free, or are the environmental conditions unsuitable for keyboards.

Voice data entry offers an ideal solution here. As compared to traditional acquisition methods, it distinguishes qualities/characteristics following through itself:
- Optimal adaptation to the habits of communication felt by the human being as natural
- Short/low training period
- Easy operation through inexperienced users also
- More freedom of movement for the user during gathering the data
- Simplified direct data acquisition
- Larger security of the input.

The COMPUTER GESELLSCHAFT KONSTANZ offers following voice data entry devices:
- CSE 1050
- CSE 1060*

The computer voice data entry devices allow follow the computer of all known manufacturers.

2 Technology of speech recognition
2.1 Summary

By voice data entry, we mean the input of the data per language into the computer. Therewith the spoken word is transformed by the voice data entry devices which are pre-connected to the computer into machine readable information. First the voice signal is converted during recognition procedure into a H.T. Then the voice data entry device compares this pattern then with the stored patterns of the vocabulary. If a sufficient correspondence is determined with one of the words of the vocabulary, this word is validly recognized. In andern case, the device refuses the input.

the vocabulary itself is previously through a training, with that every word and the second, to speak the word, set up.

The typical data flow during voice data entry appears as follows: Figure 1: Typical data flow during voice data entry
Das Spracherkennungsgesetz wandelt das analoge Sprachsignal in digitale Information um. Dabei wird zunächst das akustische Signal mit Hilfe eines Filters in einzelne Frequenzbereiche zerlegt. Aus diesen Bereichen werden die Merkmale in definierten Zeitsegmenten gewonnen. Sie beschreiben akustische und phonetische Eigenschaften des Signals. Diese Merkmale werden dann weiterverarbeitet als Bitmuster abgelegt. Dadurch können unterschiedliche Sprachgeschwindigkeiten ausgeschlossen werden.

Die Identifizierung dient der Zuweisung des Signals zu einem bestimmten Wort. Dies stellt den Identifizierungsvorgang im engeren Sinne dar. Das Wort wird durch Vergleich seines Musters mit den Bitmaren des definierens Wortschatzes erkannt. Bild 2 Technischer Ablauf des Spracherkennungsvorganges

3.3 Betriebsmodi

Die GSE-Geraete der COMPUTER GESELLSCHAFT KONSTANZ arbeiten in zwei Betriebsmodi:

- Trainingmodus
- Erkennungsmodus

In diesem Modus wird das Gerät zur Dateneingabe benutzt.

Die GSE-Geraete der COMPUTER GESELLSCHAFT KONSTANZ sind sprecheradaptives Einzelnervorwirkungssysteme. Die GSE-Geraete arbeiten in zwei Betriebsmodi:

- Erkennungsmodus

Einzelnervorwirkungssysteme erfordern eine erkennbare Pause zwischen zwei gesprochenen Worten. Bei den GSE-Geraeten muss diese Pause mindestens 100 msec betragen.

3.2 Eigenschaften des Wortschatzes

Der Wortschatz umfasst je nach Sprachenausbau des GSE-Geraetes bis zu 370 Wörter. Er besteht aus den aus dem jeweiligen Anwender- und ausschließlich aus den aus dem jeweiligen Anwender- und des Anwender-Computer übergebenen Wörtern. Beim Anwender-Computer wird der GES- und des Anwender-Computer wird der GES-Spracherkennungsgesetz genutzt.

Die Trennung nach akustisch-phonetischem Signal und der Bedeutung dieses Signals ist eine typische Eigenschaft von sprecheradaptiven Geräten. Die Erfassung der Bedeutung und des Code oder Zeichenstrings, der an den Anwender-Computer übergeben wird.

Die Tritierung nach akustisch-phonetischem Signal und der Bedeutung dieses Signals ist eine typische Eigenschaft von sprecheradaptiven Geräten. Die Erfassung mit einem spezialisierten Programm wird jeweils für eine Anwendung festgelegt. Dies gilt auch für die Anwender-Computer zu übertragenden Code oder Zeichenstrings. Die akustisch-phonetische Signalbedeutung jedoch, also das Wort, wie es ausgeprägt wird, ist sprecheradaptiert. Dadurch spielt es auch keine Rolle, ob der Sprecher das Wort mit Dieltkaufführung oder in einer sprachlichen Ausdruck. Beim sprecheradaptiven Geräte wird aus den geschilderten Gründen von sprecheradaptiven Geräten und des Anwender-Computer abgeleitet und bei Inbetriebnahme des GSE-Geraetes in dessen Speicher geladen.

Die GSE-Geraete sind selbständig arbeitende Prozessoren, die jedoch mit anderen Anwendungsprogrammen die Funktion von Peripheriegeräten haben.

Die Verbindung zur Anwenderkonfiguration erfolgt über die genormte V.24 Schnittstelle, die den Anschluss an den Anwender aller bekannter Hersteller gewährleistet.

3.4 Quicktalk-Einrichtung

Die GSE-Sprachabfragegeräte sind selbständig arbeitende Prozessoren, die jedoch mit anderen Anwendungsprogrammen die Funktion von Peripheriegeräten haben.

Die Verbindung zur Anwenderkonfiguration erfolgt über die genormte V.24 Schnittstelle, die den Anschluss an den Anwender aller bekannter Hersteller gewährleistet.

3.4 Quicktalk-Einrichtung

Die GSE-Sprachabfragegeräte sind selbständig arbeitende Prozessoren, die jedoch mit anderen Anwendungsprogrammen die Funktion von Peripheriegeräten haben.

Die Verbindung zur Anwenderkonfiguration erfolgt über die genormte V.24 Schnittstelle, die den Anschluss an den Anwender aller bekannter Hersteller gewährleistet.

Die Sprachabfragegeräte sind selbständig arbeitende Prozessoren, die jedoch mit anderen Anwendungsprogrammen die Funktion von Peripheriegeräten haben.

Die Verbindung zur Anwenderkonfiguration erfolgt über die genormte V.24 Schnittstelle, die den Anschluss an den Anwender aller bekannter Hersteller gewährleistet.

Die Sprachabfragegeräte sind selbständig arbeitende Prozessoren, die jedoch mit anderen Anwendungsprogrammen die Funktion von Peripheriegeräten haben.

Die Verbindung zur Anwenderkonfiguration erfolgt über die genormte V.24 Schnittstelle, die den Anschluss an den Anwender aller bekannter Hersteller gewährleistet.

Die Sprachabfragegeräte sind selbständig arbeitende Prozessoren, die jedoch mit anderen Anwendungsprogrammen die Funktion von Peripheriegeräten haben.

Die Verbindung zur Anwenderkonfiguration erfolgt über die genormte V.24 Schnittstelle, die den Anschluss an den Anwender aller bekannter Hersteller gewährleistet.
3.5 Weitere Eigenschaften der CSE-Spracheingabegeräte

Der Benutzer kann bestimmen, dass aktuell nur eine Unterzeile des Wort-
schatzes in den Erkennungsvorgang einbezogen werden soll. Durch diese
Strukturierung des Wortschatzes wird die Dialogführung des Benutzers unter-
stützt, da nur die der augenblickliche Phase des Dialoges als aktiv
zugelassenen Woerter vom Gerat angenommen werden.
Zusätzlich wird beim CSE 1060 die Möglichkeit der Datenausgabe geboten.
Am Puffer ist das CSE-Spracheingabegerät mehrere Eingabe auf der Puffer kann
von der Benutzer Korrekturen untersuchen (z.B. Losen einer oder mehrerer
Eingaben).

4. CSE-Gerätekonfiguration

4.1 Spracheingabegerät CSE 1050

Das CSE 1050-Gerät besteht aus
- Mikrofon
- Operateurkonsole
- Vorverstärker
- Sprachprozessor
- Anzeigezeug

Das Mikrofon ist einen leichten Kopfbügel befestigt (Headset-Mikrofon).
Hintergrundgerauschen werden weitergeleitet undroben, so dass die Spracheingabe
auch in geräuschvolker Arbeitsumgebung möglich ist.
Die Operateurkonsole dient hauptsächlich der Identifizierung des Sprachers
bei der Tafel und des Trainingsablaufes und der Aktualisierung des
Wortschatzes.
Mit Hilfe des Vorverstärkers kann das CSE-Gerät an die Lautsprecher des
Sprachers angepasst werden.
Der Sprachprozessor ist die zentrale Komponente des CSE-Gerätes. Er führt
die Spracherkennung durch. Zur Kommunikation mit dem Anwendungscomputer
besteht ein E/A-Interface eine V.24-Schnittstelle.
Mit Hilfe des 16- bzw. 32-Zeichen-Anzeigegerätes werden dem Benutzer
Informationen zur Unterstützung des Dialogs gegeben (Eingabeaufforderungen,
Kontrollmeldungen, Rückschreiben). Der Umgang der Kommunikation richtet
sich nach den Bedürfnissen der jeweiligen Anwendung. Die Steuerung der An-
zeige geschickt durch das Programm im Anwendungs-Computer.
Der Informationsaustausch zwischen CSE-1050-Gerät und Anwendungsprogramm
folgt über eine definierte Programmschnittstelle.
Über die werden sowohl Steuerbefehle als auch Daten übertragen. Die Über-
tragung findet zeichenschwarz im ASCII-Code oder über den Transport des Wort-
schatzes in Binärform statt.
Das Anwendungsprogramm kann folgende Funktionen des CSE-1050-Gerätes auf-
rufen:
- Start und Beenden des Trainings
- Anfordern von Referenzmustern nach Trainingsabschluss zur Ablage
  auf einem Hintergrundspender
- Normalierung des Gerätes
- Laden eines neuen Wortschatzes
- Bestimmung des aktiven Teilung des Wortschatzes

Im Erkennungsmodus meldet das CSE-1050-Gerät die zum erkannten Wort
verwaltete Wörter mit einer Zeichen mit der Bedeutung "Eingegangenes
Wort nicht erkannt". Diese Meldungen müssen vom Anwendungs-Computer quit-
tiert werden.

4.2 Spracheingabegerät CSE 1060

Das CSE 1060-Gerät besteht aus
- Mikrofon
- Vorverstärker
- Sprachprozessor
- Magnetbandkassettengerät
- Lichtsignal

3.5 Further qualities/characteristics of the CSE voice data

entry devices

The user can specify that only a subset of the vocabulary should be included
for the present time in recognition procedure. The dialog/prompting of the
user is supported by this structuring of the vocabulary, since only the words
authorized for the immediate/current phase of the dialog as active are
accepted by the device.

The possibility is offered additionally with the CSE 1060 to data buffering.
There with CSE voice data entry device buffers several inputs. In buffer
could the user of corrections perform (e.g. deleting/clearing of one
or several inputs).

4. Computer voice data entry device

configuration

4.1 Voice data entry device CSE 1050

CSE 1050-device consists of
- Microphone
- Operator console
- Preamplifier
- Voice processor
- Display.

the microphone is a lightweight headset attaches (headset microphone).
Background noises are suppressed to a large extent so that also voice data
entry is possible in noisy working environment.
The operator console is used mainly for the identification of the speaker
during loading as well as the training for the format and for the updating of
the vocabulary.

With the aid of the preamplifier, the computer voice data entry device can be
adapted to the volume of the speaker.
The voice processor is the central component of the computer voice data
entry device. It operates the speech
recognition technique for the communication the host computer
offers it as I/O interface to the computer.

with the aid of 16 or 32-character display are the user information
for support the dialog (prompting messages, control messages.
rejections). The extent of the communication depends on the
requirements of the respective application. The controller of the display
occurs per program in the host computer.

The exchange of information between CSE-1050 device and user program occurs
via a defined software interface.
Control character as well as data also are transmitted via them. The
transmission occurs character by character in ASCII code or for the
transportation/transportation of the vocabulary in binary form.
Application program can call the following functions of the CSE-1050 device:

- Starting and termination of training
- Requesting reference patterns after training for the file on a
  secondary storage device
- Reseting of the device
- Loading of a new vocabulary
- Regulation/definition of the active part of the
  vocabulary
- Output of text 16 or 32-character display.

In recognition mode, the CSE-1050 device reports the word number determined
for the recognized word and/or a control character with meaning/important "input word not erkannt". These messages must be acknowledged by the host
computer.

4.2 Voice data entry device CSE 1060

CSE 1060-device consists of
- Microphone
- Preamplifier
- Voice processor
- Magnetic tape cassette device
- Display device.
Mikrofon und Vorverstärker entsprechen denen des CSE 1050-Gerätes. Der Sprachprozessor enthält über die Funktionen des CSE 1050-Gerätes hinaus noch das Trainingsprogramm, so dass das Training zur Erstellung des Wortesatzes unabhängig vom Anwendungs-Computer durchgeführt werden kann.

Das Magnetbandkassettenetager setzt den Sichern des Wortschatzes nach dem Training und seinem Laden von der Magnetbandkassette in den Sprachprozessor im Erkennungsbetrieb. Uber das Sichtgerät, das mit einer Testatur ausgerüstet ist, werden die Begriffe des Wortschatzes und die zugehörigen Ausgangszeichenstrings vereinbar.

Die Programmschnittstelle zum Anwendungs-Computer ist beim CSE 1050-Gerät einschließlich mit einem Benutzerfreundlichem, durch Erkennen eines Wortes wird der versehene Zeichenstring vom CSE-Gerät übergeben. Die Übergabe des Anwendungsprogramms erfolgt wie bei einer Eingabe z.B. durch ein Daten-Sichtgerät.

5. Betrieb der CSE-Gerate
Die CSE-Gerate bieten eine sehr hohe Erkennungsgenauigkeit und damit eine große Sicherheit bei der Datenerfassung gegen ubereinmündige Benutzungsverhaltungen.

Bei "kooperativen" Verhalten des Sprachers wird eine Erkennungsausgabe von etwa 99% erreicht. Die folgenden Benutzungsregeln dienen diesem Ziel:

- richtige Auswahl der Wörter des Wortschatzes
- sorgfältiges Training
- deutliches Sprechen im Erkennungsbeginn.

5.1 Auswahl der Wörter für den Wortschatz
Die richtige Auswahl der Wörter für den Wortschatz verhindert die Gefahr von Verwechslungen (Substitutionen) im Erkennungsbeginn. Die Stimme eines Sprechers kann sich leicht durch verschiedene Einfüsse ändern. Diese Schwankungen müssen jedoch vom Sprachregebnisgestet aufgefangen werden, ohne dass das Training notwendig wird. Deshalb ist es sehr wichtig, dass die Auswahl der Wörter des Wortschatzes darauf abzielt, dass sie sich in ihrer Aussprache deutlich voneinander unterscheiden. Beispielsweise sollten die Wörter "immer" und "immer" nicht gleich klingen. Diese Schwankungen müssen durch den Sprecher abgeleitet werden, um die Wörter des Wortschatzes korrekt zu erkennen.

5.2 Gestaltung des Trainings
Sprechadaperative Geräte erfordern ein Training durch den jeweiligen Sprecher für dessen spezifischen Wortschatz. Das Training soll unter denselben Bedingungen erfolgen, wie sie im späteren Erkennungsbeginn anzutreffen sind.

Dies gilt auch für die Arbeitsausstellung als auch für die von der Person abhängige Erkennung. Der Sprecher sollte es brauchen, dass beim Training jedes Wort mehrfach eingegeben wird, um die Aussprache z.B. hinsichtlich der Übung und der Betonung zu erweitern, so dass es möglichst viele Varianten für den späteren praktischen Betrieb erfasst.

5.3 Sprachverhalten im Erkennungsbeginn
Die Kriterien für das richtige Sprachverhalten sind die Eingabeaussprech-heit und die Deutlichkeit der Aussprache. Unterschiedliche Lautstärke und Geschwindigkeit beim Sprechen des einzelnen Worts gleicht das CSE-Gerät aus. Die Eingabeaussprechheit hängt von der auf- einanderfolgenden Eingabe der Wörter ab, die in Sprechtempo gegeben werden, die auf das Eingabebereitschaftssignal ansetzen müssen.

5.4 Microphone and preamplifier devices correspond to those CSE of 1050-device. The voice processing system includes the training program on the computer and its loading of the magnetic tape cassette into the voice processor in recognition mode.

- The voice recognition system can be trained independently of the host computer.

- The magnetic tape cassette device delivers the data to the preamplifier and the device is used for recording the voice after training and its loading onto the magnetic tape cassette into the voice processor in recognition mode.

- The software interface to the host computer, 1050-device is with CSE 1050-device and it is very user-friendly, after recognizing a word is the magnetic tape cassette device deliver data to the voice processor in recognition mode. The transfer through application program occurs as during an input e.g. through a data display terminal.

5.5 Operation of the computer voice data entry devices offer a very high recognition accuracy and with it a larger security during data acquisition as compared to conventional methods.

- With the "cooperative" behaviour of the speaker is a reading accuracy of more than 99%.

5.6 Selection of the words for the vocabulary
The proper/correct selection of the words for the vocabulary prevents the danger of the substitutions (substitutions) in recognition mode. The voice of a speaker having the same name but spelling it differently must be compensated by the voice data entry device, without further training is necessary. Since it is meaningful/user-friendly to the speaker to recognize the "im" words of the vocabulary on it that they differ clearly in pronunciation from one-other/from-each other.

- The speaker should have pronounced them separate being from those "three", this is then intelligible in front of each other, even if they are coming from different recognition procedures, especially with high value of information is attributed to the vowels in the words.

5.7 Organization of training
Speaker-dependent strategies require a training through the respective speaker for whose specific vocabulary training should occur among the same conditions/requirements, as those in the later recognition mode to finds.

This applies for working environment as well as for peculiarities of speech recognition. Differences in volume and speed of pronunciation: In the training, every word can be entered repeatedly, modify around the pronunciation after the example with regard to accentuation to, so that inasmuch as many variants for the last practical operation are registered.

With operation of the device, it can result that the individual words with regard to the substitution are especially critical. In these cases, it is not necessary to train the entire vocabulary again. It suffices in a training update to input that or the individual critical words again.

5.8 Speaking in recognition mode
The criteria for proper/correct speaking are the input rate and the clarity of pronunciation. Different volume and speed of pronunciation, the individual word equals the computer voice data entry device on the input rate, however, under that is speed to understand the consecutive input of several words, is specified of the speaker.

- It is supported by an optical display which signals availability to it. This aids speaking approaches which are free of experience for the speaker very rapidly to reach optimal input rate.

- During use of the quicktalk feature, several words can be entered in sequence, without having to pay attention to the ready signal.
Ein nicht mit ausreichender Deutlichkeit ausgesprochener Wort wird vom CSE-Gerät zurückgewiesen. Die Rücksprache wird optisch und ggf. akustisch angezeigt, so dass der Benutzer eine erneute Eingabe veranlassen kann. 

6. Einsatzzwecke der CSE-Geräte

Neben ökonomischen Gesichtspunkte spielen bei der Spracherkennung auch ergonomische Überlegungen eine wichtige Rolle. So gewinnt das Argument der Anpassung der technischen Geräte an die menschlichen Kommunikationsgewohnheiten in der Diskussion um die Wahl der geeigneten Datenerfassungsmethoden zunehmend an Bedeutung.

Die Sprachaufnahme bringt viele Vorteile gegenüber den traditionellen Eingabemethoden. Da das gesprochene Wort direkt zur Daten- oder Kommando-Eingabe verwendet werden kann werden Tastaturen des Ableseens und der wiederholten Übertragung von Datens mit ihren negativen Effekten der Ermüdung und der nachlassenden Konzentration beseitigt. Dadurch wird die Fehlerrate auch bei ungeschultem Personal herabgesetzt. Tastaturen, bei denen beide Hände benutzt werden sollen, der Sichtkontakt zur Datenquelle erhalten bleiben muss oder Arbeiten in schmutziger Umgebung stehen einer Dateneingabe nicht mehr hindernd im Wege. Die Erfahrungen beim bisherigen Einsatz der Geräte haben gezeigt, dass es eine Vielzahl von Einsatzzwecken in verschiedenen Anwendungsbereichen gibt:

- Qualitätstechnik
- Lagerverwaltung
- Inventarverwaltung
- Warenwirtschaft
- ausgangskontrolle
- Betriebsdatenerfassung
- NC-Programmierung
- Anlagensteuerung
- Steuerung von Industrirobotern
- Hilfe für Behinderte
- Steuerung von Geräten
- Grafischer Arbeitsplatz
- Erfassung kartografischer Daten
- CAD
- Photogrammetrie
- Labordatenerfassung
- Röntgenbildanalyse
- Flugzeugenortung
- Paketsortierung
- Militärische Euerungssysteme
- Nachrichtentechnik

A word pronounced not with sufficient clearness is refused by the computer voice data entry device. Rejection is indicated optically and where appropriate acoustically so that the user can initiate a further input. 6. possible applications for the computer voice data entry devices

Next to economic aspects, also ergonomic considerations play an important role during speech recognition. Then the argument of the adaptation of the technical devices becomes important increasingly at the human habits of communication in the discussion at the choice of the suitable data acquisition methods.

Voice data entry has advantages. Since the spoken word can be used directly for data or command input, the activities of reading and the repeated transmission of the data with their negative effects of tiredness and the declining concentration are corrected. Thereby the error rate is reduced also with inexperienced users. The activities with which both hands must be used whose sight maintained had to be for the data source, or working in dirty environment are a data input no longer hindering in the way/path. The experiences during the previous application of the devices have indicated that there is a large number of the possible applications in the different fields for voice data entry:

- Quality control
- Warehouse management
- Stock-taking
- Stock control
- Production data acquisition
- NC-programming
- System control
- controller of industrial robots
- help/said for disabled person
- controller of devices
- Programming console for disabled persons
- Graphic work station
- Acquisition of cartographic data
- CAD
- Photogrammetrie
- Acquisition of laboratory data
- Interpretation of radiography
- Flight luggage sorting
- Parcel sorting
- Military management information systems
- Communications