Project Number:	**25420**
Project Title:	**FASTY – Faster Typing for Disabled Persons**
Dissemination level:	**PU***

Deliverable:	**D2.5**		
Title:	**1st Edited Annual Report.**		
Contractual Date:	**2001-12-31**	**Date of Delivery:**	**2002-04-24**
Work-Package:	**WP 2**		
Nature:	**R****		
Authors:	**UU**		

FASTY Deliverable Numbering D2.5/R/PU/UU/F

Abstract:
This 1st edited annual report for publication is a summary of the FASTY project outcomes at the end of the 1st year of the 3 years project.

Deliverable Numbering:

ID	Official Deliverable ID (DX.Y)	
NATURE	Nature of Deliverable	
DLEVEL	Dissemination Level	
RESPID	Responsible Partner ID	
STATUS	Wn	Working document, version n
	Pn	PRELIMINARY complete version for discussion, version n
	Dn	DRAFT, before peer review
	F	FINAL VERSION, after peer review

***Dissemination:** INT-Project internal, CO-Confidential, RE-Restricted access
IST-IST participants, FP5-FP5 participants, PU-Public document

****Nature:** P-Prototype, R-Report, D-Demonstrator, S-Specification, T-Tool, O-Other

LIST OF AUTHORS

UU – Uppsala University, Department of Linguistics
S-75120 Uppsala
Villavägen 4
Sweden

CONTENTS

PICTURE INDEX

GLOSSARY OF TERMS AND ACRONYMS

Project partners	FASTY Principle and Assistant Contractors
Contract	FASTY Project contract with the EU
CA	FASTY Consortium Agreement

1 Introduction

Communication is a primary need for humanity. While this holds true for mankind in general, the need for communication is even more urgent for disabled people. It is also true that computers play an increasing role as a communication tool. However, in this on-going process disabled people are at a disadvantage. In particular, motor impairments make the use of standard text input devices to the computer difficult and hence slow. For instance, while non-disabled writers have a typing speed of some 200-300 characters a minute, the typing speed of a user operating the keyboard with a mouth stick is not higher than 75 – 120 characters a minute. Motor impairment often goes together with articulator deficiencies. Thus alternative methods based on speech input will not solve the problem. Providing methods for speeding up keyboard text input is a better way to go. This is the strategy that has been chosen in the FASTY project.

The project is scheduled for the period from January 2001 to December 2003. Below we will give a brief description of the aims of the project and the achievements made during the first project year.

2 Aims

The concrete goal of FASTY is the creation of a system for increasing the text generation rate of disabled persons by so-called predictive typing and dedicated advanced input devices. A prediction system attempts to predict subsequent portions of the text by analysing the text already entered and using frequency data on the vocabulary of the language. Character-by-character text entry is thus enhanced by the possibility of entering whole words and portions of words as they are proposed by the system. The selection of an alternative should be made by means of a single keystroke. Complementary to the presentation of the proposals on the screen, they will be read aloud by means of speech synthesis. The success of a system of this kind can be measured in terms of keystrokes that are saved using the predictions as compared to traditional character-by-character input. FASTY aims at a keystroke saving rate above 50%. Experiences that were made during the first project year indicate that the linguistic quality of the text will also benefit from using the prediction system.

The FASTY text prediction system should apply to four languages: Dutch, French, German, and Swedish. The future inclusion of additional languages should also be considered. The multilingual aspect is reflected in the design of the system.

User involvement in the project is strong. It is ensured by means of a user panel. The user needs are analysed subjectively by intensive interaction with the user panel. Additionally an objective User Ability Assessment Tool, UAAT, has been developed for user needs investigations. The user panel also plays an important role during verification and validation of the prototype systems. There are two kinds of users in the panel: primary end-users and secondary users such as pedagogues, therapists, carers and family members.

An important aspect of the project is the design and development of a dedicated interface adapted to the needs of the users. The user interface design and the features of the predictor program aim at a wide coverage of end users (various disabilities) and secondary users (various roles in supporting the disabled person). Self-adapting parameters and flexible configuring should ensure a high degree of usability, user friendliness and accessibility to the system. A User Simulation Tool will be used in testing the system and adapting it to different users.

Innovative and ergonomic user interfaces for various existing input methods (standard keyboard, on-screen keyboard, scanning) are developed together with the predictor thus minimising time and effort for selecting the desired word from a selection list presented on the screen. In addition, a special pressure sensitive switch/keyboard will be developed and used to improve the user interface, UI. Strategies for optimal exploitation of residual functions will be implemented.

Dissemination and Exploitation play a central role throughout the project. A Technological Implementation Plan (TIP) has been developed as a preparation for the exploitation plan. After successful finishing of the project the consortium will co-operate in order to come up with a commercial product.

3 Consortium

The FASTY consortium consists of nine partners from four countries: Austria, Belgium, Germany, and Sweden.

There are six principal contractors:

- fortec - Research Group for Rehabilitation Technology / Project co-ordinator
- ÖFAI - Österreichisches Forschungsinstitut für Artifical Intelligence
- FTB - Forschungsinstitut Technologie - Behindertenhilfe
- UU - Uppsala University, Department of Linguistics
- MULT - Multitel ASBL
- IGEL Elektronische Kommunikationshilfen GmbH

and three assistant contractors:

- ELI - Seraphisches Liebeswerk, Elisabethinum Axams
- IKuT - Ingenieurbüro für Kunst und Technik II
- FUNDP - Facultés universitaires, Notre-Dame de la Paix

fortec is the project co-ordinator and responsible for managing the project, and for system architecture and internal interfaces.

Öfai is responsible for the implementation of the language components and the provider of language resources for German. It also has the responsibility for system integration and prototyping.

FTB is responsible for user involvement including ethics, quality assurance, and user validation. It also has the responsibility for the implementation of the user components.

UU is responsible for dissemination and public relations. It is also the provider of language resources for Swedish and responsible for the grammar based prediction.

MULT is responsible for verification and redesign of the system. It is the provider of language resources for French and Dutch and of speech synthesis solutions for all languages.

IGEL is responsible for the exploitation of the system and for technical implementation. It also contributes to specifications, architecture and user involvement.

ELI's main task is to provide feedback to the developers. It also participates in user related tasks such as prototype testing, assessment of user needs etc.

IKuT is the developer of new input devices. It is associated with a local user group.

FUNDP provides an interface between the users and the developers. It participates in the analysis of user capabilities and needs, and assists disabled people in using the product.

In addition, there are a number of sub contractors.

A listing of the project partners with contact information is to be found in Chapter 10.

4 Innovative Aspects

There are a number of word predictors on the market. However, they have, typically, been developed for English, which means that they are not well suited for morphologically rich languages such as Dutch, French, German and Swedish. Simply adapting the English programs to these highly inflecting languages by replacing the English dictionaries usually implies a massive reduction of keystroke saving rate. These effects are due to the simplistic language description that is used for predicting English text and that fails to predict the correct inflectional form of a word as required by inflectional languages. The English language description is, as a rule, limited to frequency data on individual words (unigrams) and sequences of words (bigrams, trigrams). Attempts have been made in research systems for Swedish and Spanish to use a more elaborate language description, including n-grams of word classes [2] and syntactic grammars [27]. The experiences made in these projects are taken into account in the FASTY project. They do not, however, present solutions that will ensure a keystroke saving rate of above 50% for the FASTY languages. An additional problem with most of the FASTY languages is the fact that new compounds can be created on the fly, thus making it hopeless to strive for a complete lexicon. Other methods need to be employed for coping with dynamic word formation processes. Being able to cope with compounds, even if they are new, is of great importance, since compounds are usually rather long words and failing to predict them can cause a significant drop of the keystroke saving rate. Since no word prediction system currently available is able to handle new compounds, this aspect of the FASTY system is a true innovation.

At a general level, the innovative aspects of the FASTY predictor are represented by

- the predictive power of the prediction engine that is based on a sophisticated language component
- a dedicated, flexible and adaptable user interface that is an integral part of the system
- new input devices

In particular, the innovative nature of the FASTY predictor is reflected in the following features:

- prediction of compounds
- prediction of proper inflectional forms based on the use of parsing

- generic algorithms to ensure cross-language portability
- dictionaries based on general language corpora and on the users' own texts
- adaptation of the dictionaries based on actual use of the predictor
- initially supported languages: Dutch, French, German, Swedish
- user interface that is an integral part of the predictor and
  - adaptable by primary and secondary users
  - capable of using different kinds of input devices
  - automatic adaptation to the performance of the user
- new input device
  - pressure sensitive switch
  - pressure sensitive keyboard

5 The FASTY System

During the first project year the architecture of the FASTY text prediction system has been specified, developer tools have been designed and implemented, and parts of the system have been implemented. The specification of the system is based on

- a market study
- a study of available technology
- a study of the users' special needs

Special care has been taken do define an open and flexible architecture that is adaptable to the users' needs. Another focal point has been to define a generic system that is not geared to any specific language. Thus the language components will be independent of the operating system and the prediction strategies will be evaluated with regard to all the FASTY languages. The system should also be portable to platforms other than MS-Windows.

The main parts of the FASTY system are

- a runtime system
  - user interface, UI
  - language component, LC
- a user adjustment tool, AT
- four developer tools
  - user ability assessment tool, UAAT
  - text collection tool, TCT
  - user simulation tool, UST
  - a simple word prediction program, SWP

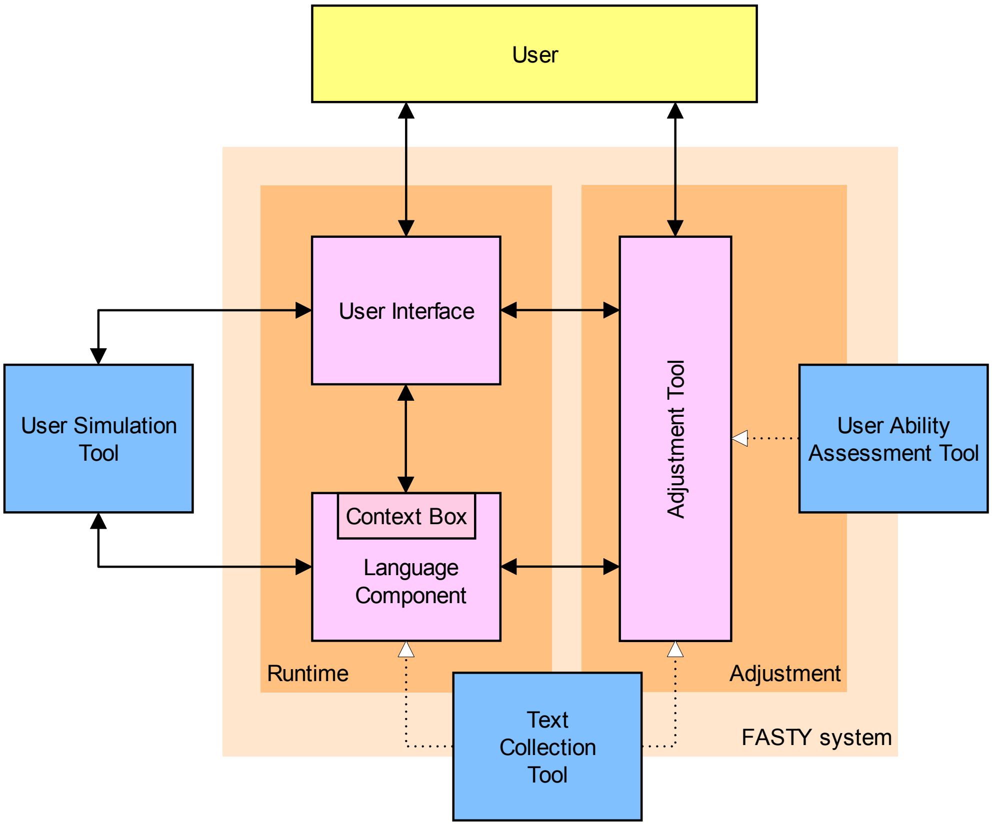

Fig. 1 General System Structure

The runtime system consists of a user interface, UI, and a language component, LC. Below we give a brief functional description of the different component.

5.1 The User Interface

5.1.1 The user centered approach

The primary user interacts with the system via the user interface, UI. A study of the products on the market gave the consortium a good starting-point for the design of the UI, and the arising ideas were built into a Rapid Prototype, RP, and first demonstrator.

The objective of the 1st Rapid Prototype (RP) was to give potential users a first impression of some aspects of the user interface.

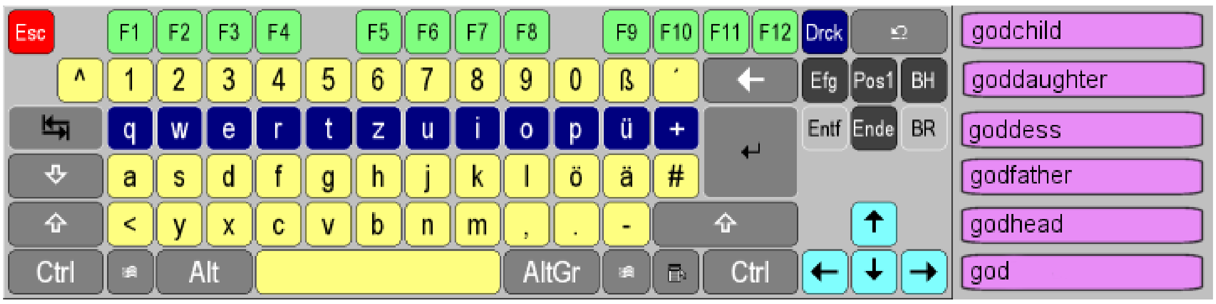

Fig. 2 Example for an OSK with prediction list
(used e.g. for a first idea collection)

The RP offers a wide variety of settings concerning the user interface combined with a rudimentary word prediction:

- a variety of docking options
- color settings
- font size settings
- scanning option
- hotkeys selection
- different sorting options

The RP has been tested by members of the user panel and their comments have been taken into account in the specification process.

5.1.2 Overall Specification

As regards, the UI was defined very open to fulfill the very wide spread needs of the different users (see Fig. 3).

In general, the User operates one or more input devices. Each input device interacts with an Input Driver, which sends standardized information to the Kernel. The Kernel interprets these input information, gives user feedback via Output Drivers and sends generated keycodes to the Keycode Filter. This module separates additional information for the Language Component (therefore it will not be passed to the operating system) from „normal" keystrokes, which will be sent to the Language Component (via the Context Box) and to the Operating System. The Context Box collects the single keystrokes and saves the left content for the Language Component. If a task-switch is recognized by the System Monitor, the Context Box switches the saved content. Coming back to this task, the previous content is restored; if a task is terminated, the System Monitor deletes the affiliated content in the Context Box.

Program Drivers will ensure the input of special information such as caret position, format information etc. if they are available. With these drivers, a better User Interface layout will be possible if the destination program has the possibility to supply the information and this program is supported (a driver is available); third party providers may write their own support-drivers in the future. Advanced information about the destination program (and which driver is needed) will be obtained by the System Monitor from the Operating System. The FASTY system will include a simple Pre-Editor for an easier handling of text-correction (deleting words, lines, aso.), that will demonstrate the idea and the power of this interface.

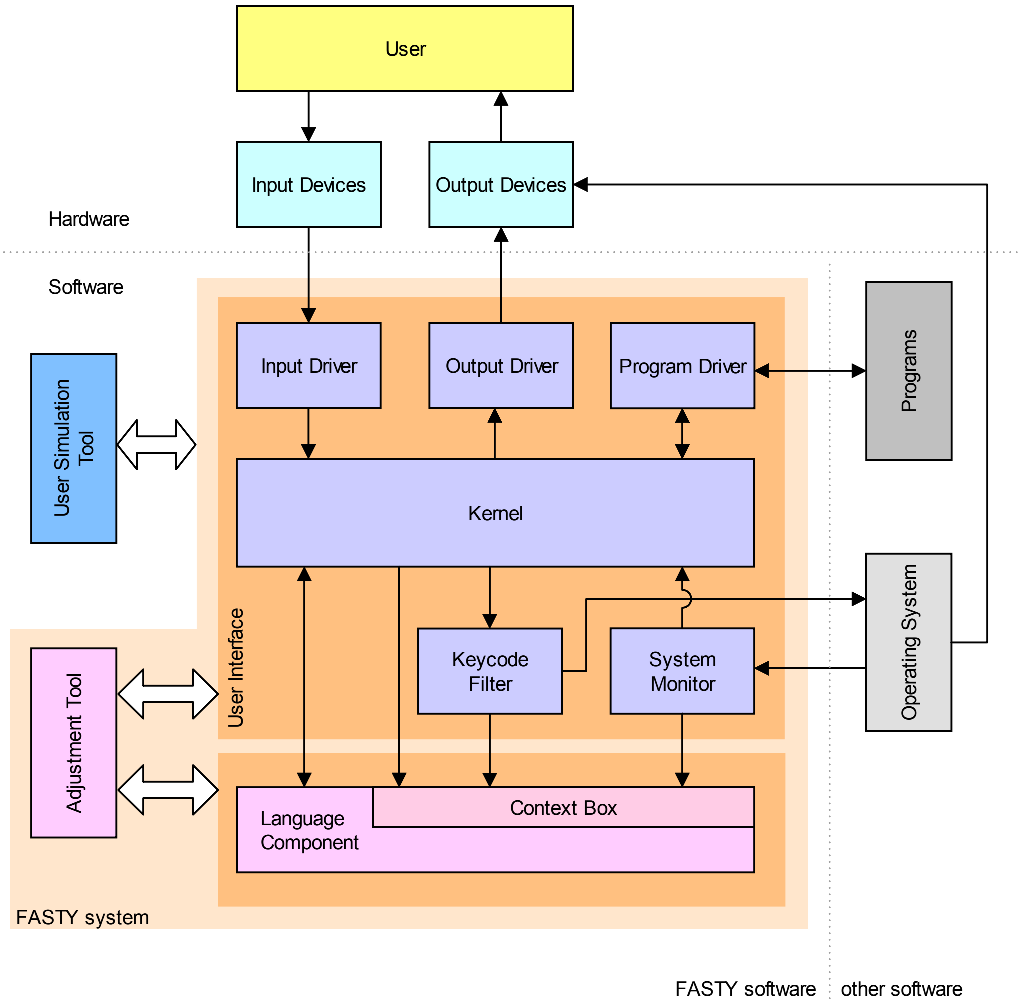

Fig. 3 General structure of the UI

For more details, especially on the technical issues, see [8][10].

5.2 *The Language Component*

Below we focus on the functionality of the language component. For further detail, see [11].

5.2.1 Overall Specification

FASTY's core functionality will be provided by a statistic language model based on n-grams. These statistics will be collected not only from standard corpora, but also from texts generated by disabled users in various communication scenarios. In addition, the use of topic specific lexica and word statistics will have to be considered. It is a well known fact, that word probability is not independent of context. Word n-grams yield only a rough approximation of this variation, there are also lexico-semantic and topic-specific factors influencing word distribution. So-called 'recency' adjustment is a special case of this phenomenon. Different approaches, such as collocation analysis or trigger pair identification will be explored to collect statistics that may help in finding the most probable predictions. The use of these statistics and their integration in a prediction system is another innovative aspect of the FASTY system.

The variety of languages to be supported and methods to be integrated into the FASTY system demands a modular architecture. The combination and integration of prediction components needs to be handled in a very flexible way, for a variety of reasons:

- Different languages may put different emphasis on different modules, so it must be possible to arrange these modules in a different way.
- The effects of the different prediction methods are not yet known precisely, experimental adjustments as well as parameter tuning have to be possible.
- Different application scenarios (varying from text writing up to spontaneous dialogs) may require different combination and weighting of components.
- Adaptation to users with different degrees and types of disabilities will also be required.

Thus the backbone of the linguistic prediction component of the FASTY system will be a controller that is flexibly driving the different prediction modules and combining their results. Thus it will be easy to optimise the overall prediction behaviour and also adaptation of FASTY to another language without modifying the whole system will be made possible.

5.2.2 Methods used in existing prediction systems

When one considers methods for saving keystrokes in text typing, one has to differentiate between keystroke-saving methods in the UI and methods involving linguistic or statistical prediction of words, word sequences and phrases. Here we will put our emphasis on the latter. However, for the sake of completeness a short listing of methods belonging to the User Interface side will be given.

- automatically inserting a space after every predicted word accepted by the user. This method compensates for the extra keystroke needed for selecting the prediction.
- automatically removing preceding whitespace immediately before punctuation characters (and inserting the appropriate amount of spaces afterwards). This method complements the previous one, as the need for the user to backspace the automatically inserted whitespace is alleviated.
- auto-capitalisation. This method in fact also needs at least some linguistic knowledge, it is listed here just because of the requirement to be able to change characters the user already has typed. Auto-capitalisation may occur after sentence ending periods, on words recognised as proper names or (in some languages, e.g. German) on nouns in general.

5.2.2.1 String-based statistical methods

All systems on the market that we are aware of use some kind of frequency statistics on words and (sometimes) word combinations. Given a prefix of a word, with a frequency annotated lexicon the most probable continuation(s) of that word can be retrieved easily. Sometimes, not only word-based frequency counts are maintained (1-grams), but also bigrams and even trigrams are used for enhancing prediction accuracy. N-gram language models are widely used in speech recognition systems, and their benefits are also exploited in some predictive typing systems. The key observation behind this kind of models is, that the probability of a word occurring in the text depends on the context.

5.2.2.2 Syntactically motivated statistics

The superiority of n-gram based predictions over simple frequency lexicons stems from the fact, that n-grams are able to capture some of the syntactic and semantic regularities intrinsic to language. However, a severe drawback of word-based n-grams is, that, even with very large training texts, the data still is rather sparse, and thus in many actual cases during prediction no information is available. The usual technique to cope with syntactic regularities uses class-based n-grams (usually n=3), the classes being defined by the part-of-speech information of a tagged corpus. Copestake [4] reports on an improvement in KSR of 2.7 percent points by just taking PoS-bigrams into account. A good description on the integration of PoS trigrams into a statistical word prediction system for Swedish is given in [2].

5.2.2.3 Capturing semantics with statistics

For a human language user it is obvious that in a given context some words are more probable than others just because of their semantic content. Factors influencing word probability due to semantics are (among others):

- the user and the type and topic of the text s/he writes (global factors)
- constraints due to the lexical semantics of words (e.g. subcategorisation requirements);

these are local factors that mostly operate on sentence level.

Collocation analysis (in a broader sense, not reduced to idioms only) can reveal some of these dependencies. However, very large corpora are needed. Rosenfeld [29] uses the concept of „trigger pairs” to capture these relationships statistically (basically these are bigrams of words occurring together in a window of a certain size in a corpus). If a word that has been recently entered occurs in such a trigger pair the probability of the other word of the pair should be increased. Recency, as implemented as a heuristic in some prediction systems, can be seen as a self trigger and is a (rather crude) measure to exploit semantic or topical appropriateness of a word.

5.2.2.4 Rule-based approaches

Several methods of integrating grammar rules into statistics based prediction have been tried, but none of them had made it into a commercially successful product. Such an integration, however, is seen as a major challenge in the FASTY system.

5.2.3 Linguistic components and resources for text prediction

Basic to our approach is the modular architecture of our system. In addition to the flexibility such an approach provides for the adaptation to different languages, application scenarios and users – as described in the introduction – it also ensures robustness and graceful degradation in the case one module should be missing or fail. Furthermore, this type of architecture allows for the possibility of exploring various more advanced – and albeit more risky – methods without endangering the successful implementation of the language component in case some of these methods should not prove successful.

The core of the system will be a module based on the prediction of word forms due to their absolute frequency and the probability of their associated PoS. Such a module is state-of-the-art and guarantees a basic performance. A number of other methods to improve prediction quality will be investigated. All methods will be evaluated with respect to their performance for different target languages and language specific phenomena (e.g., compounding). Those that prove to be successful for one or more of the target languages will be integrated with the core component – either alone or in combination with others.

5.2.3.1 General word n-gram-based Prediction

As mentioned previously, n-gram prediction (for n > 1) is superior to the use of a simple frequency-ranked lexicon, and will be used in the FASTY language model as a base. The probability of occurrence of a list of word bigrams will be the main reference since it is expected to contribute most to the appropriate choice of predicted items. Word n-grams of longer lengths, including so-called lexicalised phrases, will also be taken into account. Because most word n-grams for n > 2 have low probability in comparison to word bigrams, methods will be developed to ascertain when and how to predict them. These word n-grams will be accessed for both word prediction, when a complete word is chosen to follow an already-typed word, and for word completion, when the user has already begun to type a word. Text corpora for the acquisition of word n-grams will be collected from material belonging to the project partners and from texts on the Internet that are freely available for use.

5.2.3.2 User- and Topic-specific n-gram-based Prediction

If the probabilities of user- and topic-specific words are to be consistent with the probabilities of the word n-grams which are derived from corpora containing millions of words, the texts from which these words are taken would have to be of the same order. Since this is not possible, especially in the case of user-specific words, a factor to adapt the measured probability of words and longer expressions in the user- and topic-specific lexicons will be determined by experimentation. It will be possible to generate lexicons of both user-specific and topic-specific words and expressions from previously written, and electronically readable, texts. User-specific words and expressions will also be stored during text composition. The use of several user- and topic-specific lexicons at the same time will be allowed and all activated lexicons will be searched. Words and expressions with highest probabilities (naturally occurring as in the case of word bigrams, or adjusted as in the case of longer expressions and words coming from user- and topic-specific lexicons) will be offered in a prediction list.

User-specific word n-grams may be collected automatically by the word prediction program while the user is writing texts. They may also be included in the database by generating entries from previously-written texts while running the program in an automatic mode. This same method may be used for any computer-available texts in particular subject areas.

5.2.3.3 Part-of-Speech n-gram-based Prediction:

A statistical model will be devised in which probabilities for tag trigrams play an important role. A factor determining the relative importance of this information will be derived experimentally. Tagged text corpora or previously developed taggers will be acquired from researchers or research institutions for the acquisition of word class n-grams where this is found to be possible.

5.2.3.4 Morphological processing and Backup Lexicon:

The morphological component has three modes of operation:

- analysis: given a word form, morphological analysis returns the lemma, the part of speech and a set of morpho-syntactic features.
- generation: given a lemma and morphosyntactic features, all wordforms compatible with the given information will be returned.
- completion: given a prefix, morphological analysis returns all the wordforms contained in the lexicon having this prefix, plus all the other information returned by regular analysis. Care has to be taken that the prefix is selective enough, otherwise huge amounts of data will be returned. Probably the number of solutions returned should be limited.

Morphological processing will be implemented with finite-state transducers. A prototype implementation is already available as an ÖFAI background resource. How the needed large coverage morphological lexica will be acquired is not yet clear. Several possibilities exist: one could start from wordlists either given or collected from corpora and use existing morphology engines to come up with the analysis of these words. Another way could be a learning approach, such as [15]. The resulting annotated wordlist can then be converted into a transducer along the lines of [12].

5.2.3.5 Abbreviation Expansion

The following basic functionality should be foreseen:

- given a prefix, all abbreviation codes starting with that prefix should be returned.
- given an abbreviation, the expansion string should be returned.
- given an abbreviation code and its expansion the system should store that abbreviation into the abbreviation table.

Abbreviations are entered by the user or by its care-person. The possibility of automatically scanning the user's texts for frequently occurring word combinations and suggesting abbreviation codes for them can be explored.

5.2.3.6 Collocation-based prediction

Collocation-based prediction should deliver a list of correlated words given a list of content words (from a fixed window of the left context). Input as well as output will be lemma-based, as morphological variation is supposed not to contribute significantly to the semantic relation, and lemma-based counts will be less sparse. Consequently, the collocation component will have to interact with the morphological processor.

Trigger pairs will have to be collected from rather large corpora, which need not have any annotation. Morphological analysis will be needed during collection, since the triggers should be lemma-based. A publicly available tool for efficiently extracting trigger pairs from large corpora has been written by Adam Berger and can be downloaded from http://www.cs.cmu.edu/~aberger/software/.

5.2.3.7 Grammar-based prediction and prediction ranking

The syntactic predictions will be based on partial parsing. The grammar will not be complete. It will include rules for selected constructions that are identified as crucial from a prediction point of view for the individual languages of the projects. Typically, they will be constructions with a fairly fixed word order and feature constraints. Examples of such constructions are nominal phrases, prepositional phrases, and verbal clusters. Sentence initial position and the placement of the finite verb are candidates for additional focal points considered by the grammar. The grammar-based predictions will be scored according to frequency, internal syntactic structure, and, if appropriate, the position of the phrase in the sentence. In terms of these scores, the predictions generated by the syntactic component will compete with those generated by the other modules of the system. Sometimes, they will coincide.

It is planned to use the extensive work of the partner at Uppsala University in chart-based grammar checking as a basis for the syntactic component (see [31][32][33]). A concrete result of this work is a chart-based parser for partial parsing written in C. It goes with a grammar for Swedish, but grammars for the other languages of FASTY remain to be written. The parser provides a basis for generating prefabs in terms of syntactic phrases, provided that a full-form dictionary is available. Tools for analysing compounds and other words outside the dictionary will be shared with other components of the system.

5.2.3.8 Compound prediction

Presumably, compound prediction is useful only, if the first part of the compound is already known. The following factors will most probably influence the prediction of compounds:

- The first part of the compound
- the n-gram predictions taken as if the first part of the compound hadn't been entered. These words could possibly be compound heads.
- mutual information statistics (as with the collocation-based prediction described above), as the parts of compounds are usually semantically related.
- the results of the collocation-based prediction, as newly formed compounds usually relate to the previous context and are often created to avoid repetition of words and phrases.

Taken these factors into account, the component will deliver compound predictions. Morphological compound formation rules have to be taken into account. Since compound prediction is a true innovation in word prediction systems, the way how to infer new compounds from the input evidence is not precisely known yet and subject to further research within the project.

The morphological compound formation rules have to be created manually by language experts, alternatively a learning approach along the lines of [16] can be explored. Acquisition of the mutual information statistics will be similar (if not identical) to the one needed for collocation-based prediction. In addition, compounds existing in the lexicon will be analysed and contribute also to the mutual information statistics.

5.2.4 Interaction of components and control structure

The operation of the Language Component is driven by the User Interface. Depending on the request type the LC returns one or more values as a response to the UI (e.g. the n most likely predictions). A central part of the interface between LC and UI is the Context Box. Its most important part is a string buffer serving as a communication area between. It contains he left context of the current text entry point.

The Controller receives requests from the User Interface and is responsible for

- Extracting the input data required for the different prediction components from the Context Box
- Selecting which prediction components to use (depending on the current parameter
- settings)
- Feeding the different prediction components with the appropriate input

Possibly enriching the Context Box with data returned from some components (e.g. part-of-speech information)

The Prediction Generator receives the predictions made by the different components, together with their probabilities, and combines them to a prediction list which is delivered to the User Interface. How the Prediction Generator comes up with the combined prediction depends on

- Parameter settings, which may be user and language specific. E.g. Compound prediction may not be useful in French, or every user may want to use his or her own abbreviation table.
- Interpolation weights that have to be determined empirically.

Each of the components relies on language specific resources, some are shared between different components. Also the possibility exists, that a component uses the results of other components, e.g. Grammar-based prediction uses compound analysis and morphological processing.

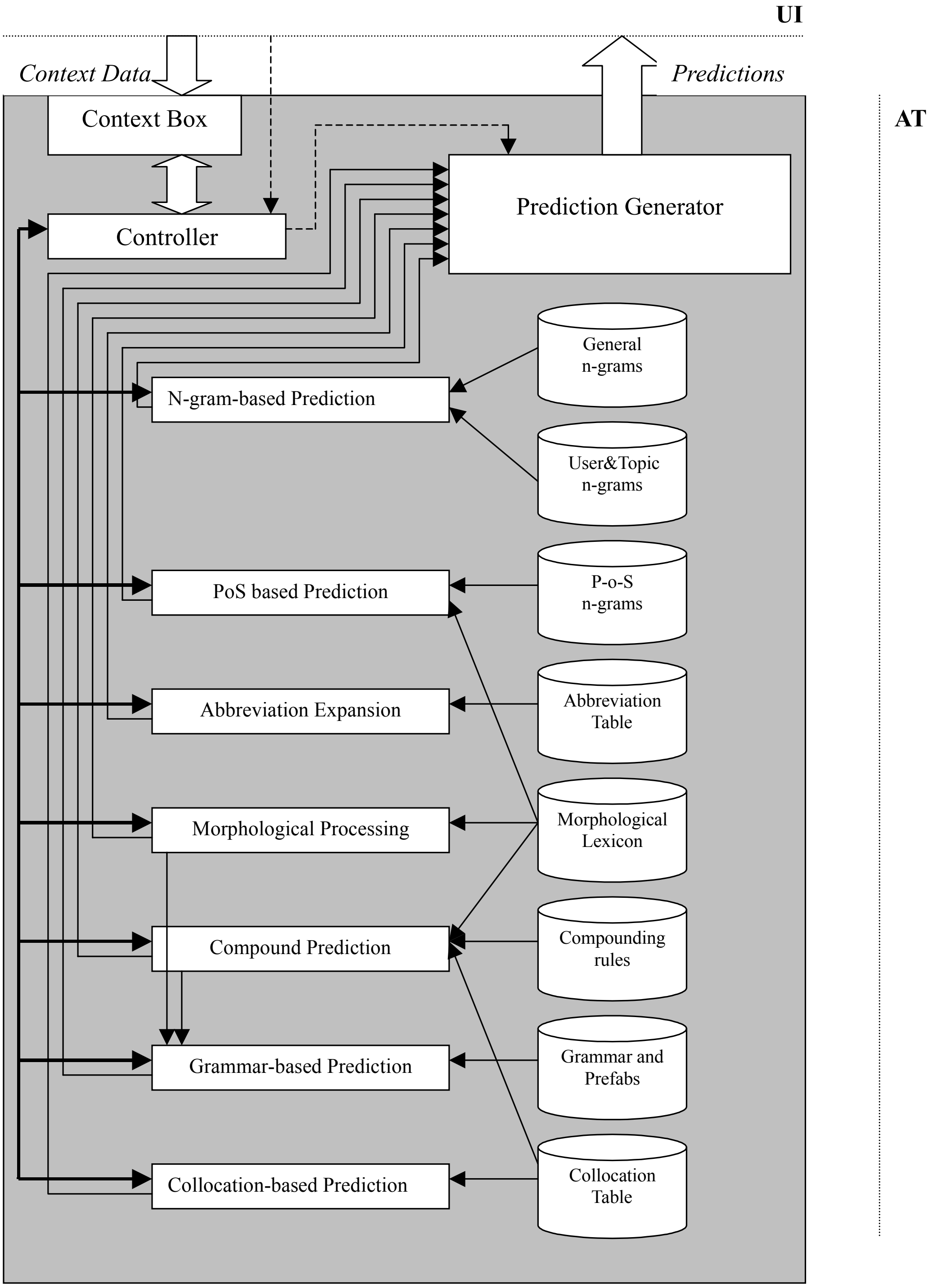

Fig. 4 Architecture of the Language Component

5.2.5 Speech Synthesis

Speech synthesis can play an important role in a word prediction system. For users who have reading problems including persons with dyslexia, it is very helpful to have words in the prediction list read aloud. These words are very similar after the first letter or two of a word has been typed, and this can be quite confusing for someone who often recognises words just from their first letters. Results that were obtained during the user's survey reported in „Report on user abilities, preferences and needs“ [6] confirm this. Users point out that an audible feedback of a selected word is a very attractive option. Nowadays most of the text prediction software on the market include a speech synthesiser. The FASTY system will follow this line. Additional advantages provided by a speech synthesiser can be given by allowing all menu items and messages to be readable by means of speech synthesis. Hereby all users will be able to participate fully in the use of the program.

The speech synthesiser should easily handle several languages and, in the first place, all the project languages. The speech synthesiser that we intend to use will be a combination of a „grapheme (letter) to phoneme" converter and a „phonetiser" (phoneme to sound converter). The „grapheme to phoneme" converter will be based on a so-called induction decision tree, ID3 [26][28][17]. It learns „grapheme to phoneme" conversion rules automatically from a phonetised dictionary. Thus no extra linguistic knowledge need to be generated provided that a large phonetised dictionary is available. This is an important aspect in the multilingual setting of the FASTY project.

The phonetiser that we propose to use in FASTY project is the award winning (IT European award 1996) MBROLA synthesiser available through Multitel partner [14]. MBROLA is a speech synthesizer based on the concatenation of diphones. It takes a list of phonemes as input, together with prosodic information (duration of phonemes and a piecewise linear description of pitch), and produces speech samples that may be played in the earphone of the computer. More information on MBROLA synthesiser algorithm may be found at http://tcts.fpms.ac.be/synthesis/mbrola. MBROLA currently handles 26 languages and the procedure to add a new one is easy and well defined. Note, however, that not all the languages of the FASTY project are available yet.

The language component delivers a list of predictions to the UI and the UI transfers them to the speech synthesiser to be played by the earphone. The predictions must be supplied with information about part of speech for the speech synthesis to work well. The speech module should drive any sound card compatible with Microsoft Windows standards (Sound Blaster compatible).

5.3 User adjustment tool

Before a user starts working with the FASTY system, it has to be adjusted to her specific needs and situation. The adjustment consists in the setting of a large number of parameters and options. This is not a trivial task, and it will be carried out in co-operation between the primary user and the carer. The Adjustment Tool will offer support in this process. It provides three modes: a simple guided one for beginners, and a more sophisticated one for advanced users and an expert mode with all options.

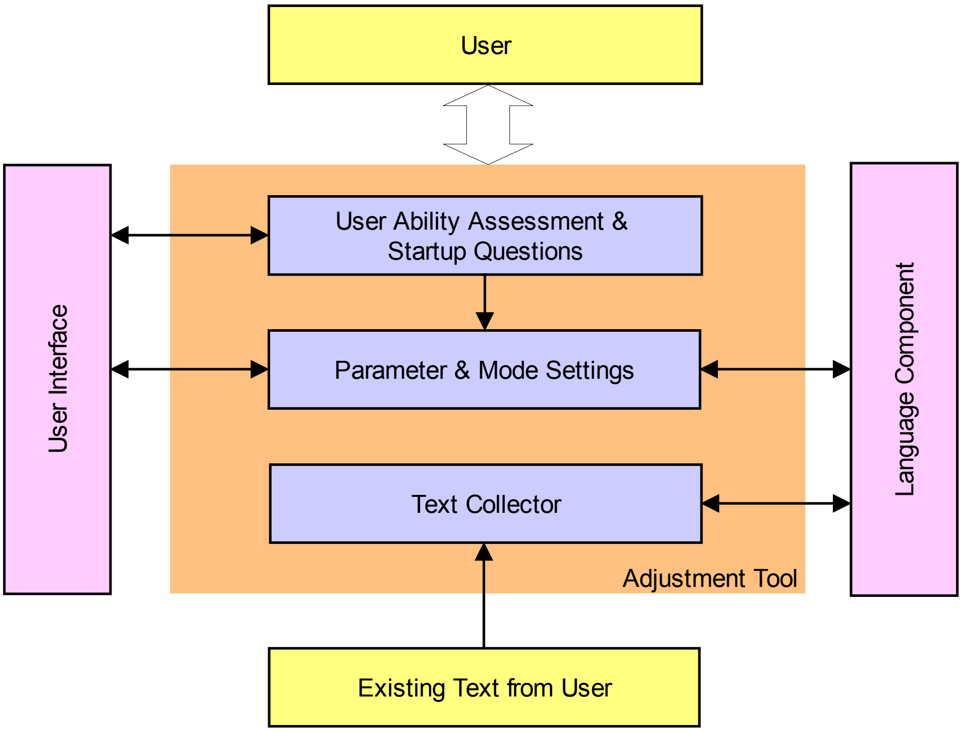

Fig. 5 General Structure of the Adjustment Tool

The Adjustment Tool will contain some functionality based on functions of the User Ability Assessment Tool, but it will also contain extended features. The user runs a set of tests and answers some questions, which are intended to provide a picture of his abilities. The results are used to make settings in a first run. Those settings can be fine-tuned manually with increasing practice and experience. The user-presentation of the parameters will be divided into some user-modes like wizard (beginner), advanced and expert. This model ensures, that all parameters may be changed. At the same time, inexperienced users will not be bothered by them.

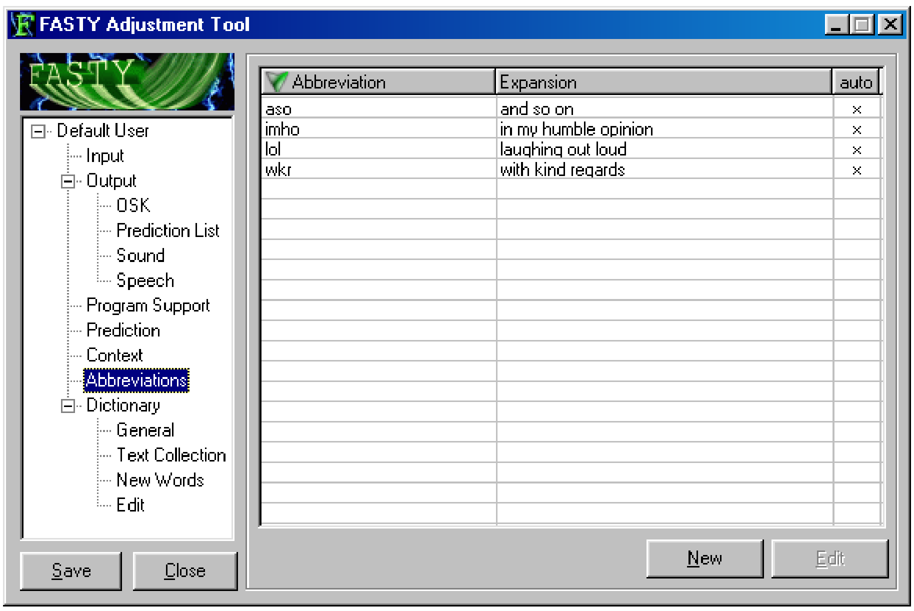

Fig. 6 Study of a possible UI for advanced users

5.4 *Developer tools*

As mentioned above, the FASTY system comprises four developer tools, i.e.

- a user ability assessment tool, UAAT
- a text collection tool, TCT
- a user simulation tool, UST
- a simple word prediction tool, SWP, for the developers of the language component

5.4.1 User Ability Assessment Tool (UAAT)

A User Ability Assessment Tool (UAAT) was developed to collect data about basic user performance, such as typing speed, or reaction time. The program supports standard input devices as well as special input devices connected via serial port and the prototype of the pressure sensitive keyboard developed by IKuT. The collected information is used to determine the applicability of further tests to the situation of the user. It is also used to get an impression of the users current hardware and software status and the way text input is written.

The results from the questionnaires and the User Ability Assessment Tool do not show a uniform picture of the potential users. There is a wide variety of abilities and demands. Although FASTY will not contain communication assistance, the desire for an appropriate support was expressed.

5.4.2 Text Collection Tool (TCT)

The language component requires big amounts of text for building user-adapted dictionaries. The TCT was built with this aim in mind. It will be used during the development phase for the collection of text from the Internet and from the users. Later versions will be integrated in the Adjustment tool and will be used to adjust the general dictionaries to the special needs of a user; this adjustment will be part of the installation process. An important feature of the TCT is its ability to anonymise the texts so that they cannot be traced back to the author.

5.4.3 User Simulation Tool (UST)

A user simulation tool will help the UI and LC developers to evaluate different algorithms and settings. Later commercial versions may help carers to find the best settings for the primary users in an iterative way without burdening the primary user himself with the testing of all possible settings and options.

5.4.4 Word Prediction Tool for Developers (SWP)

A word prediction tool has been developed to support the continuous development of the language component. Currently, it includes the core langue functionality of the FASTY system (n-gram prediction), the automatic acquisition of n-grams from corpora, and the morphology module. The tool is in daily use at the three language development sites (ÖFAI - German, MULT – Dutch and French, UU - Swedish). By means of SWP, baselines for the core functionality for the four languages have been set in terms of keystroke savings for different parameter settings of the program. As new functionality is added to the SWP, it will be possible to study how the predictive power and precision increases with the integration of new functionality with respect to the four languages.

6 User involvement

6.1 User abilities, preferences and needs

The aim of the report on „User Ability, Acceptability, Preference and Need Assessment“ is an in-depth analysis of the abilities, needs and preferences of potential users. In particular to investigate the potential users knowledge and needs with respect to word prediction and related tasks are important information when defining the functional specification of the FASTY system.

In the beginning of the project a potential user pool consisting of 58 primary and secondary users was set up by the partners FTB, ELI, IKuT, FUNDP, IGEL, MULT and FORTEC.

Two questionnaires were developed. A User Description Sheet (UDS) to be able to choose appropriate participants for this investigation. An extensive User Questionnaire with regards to the preferences and user needs in view of word prediction but also sections about the handicap, used software and hardware, PC usage and typing, and the readiness to be involved and to support FASTY.

The results of the user questionnaires are summarised below

- Nearly 75% of the primary users have mobility problems, some of them have extreme difficulties in PC operations. Half of the users have difficulties in communication.
- More than a third of the users do not use a standard mouse or a standard keyboard. Nearly half of the users do not operate the keyboard with both hands.
- Nearly all users will use the word prediction with MS Word or another word processing software. A third will use it also for Internet applications.
- Approximately a third of the users are still working with Windows 95. The remaining group is working with Windows 98 or higher. The primary users are working with older and slower hardware than the secondary users.
- 70-80% of the users are ready to be involved in or to support FASTY. Only half of the users are willing to supply private or office text for text collection and text collecting tools.
- About half of the primary users know word prediction programs even if they do not use them. Only a few users have some concrete ideas about word prediction.
- The majority of the users thought that a program with word prediction should be useful for them. They expect to write or type faster, to be able to write and communicate and to write correct sentences.
- Users want to have prediction with an insertion as automatically as possible. It is important that the predictor inserts more than a word, e.g. a space after a predicted word and after punctuations, capital letters after punctuations and a u after q. Abbreviation Expansion is also desirable.

- The user wants to choose the predicted selection with function keys or with the input device. The selection list should contain at least 1 or 5 predictions but no more than 10. The number of words in a prediction string should be adjustable.
- Grammar checking is desirable for a great majority of the users.
- A third of the users wanted opportunities of audible feedback.
- It is important to the users that new words are added automatically to the dictionary which should be editable.

6.2 Ethical Aspects

The Report on Ethical Aspects describes the actions and efforts that have been and will be taken to assure an appropriate ethical standard in the FASTY project and will serve as a guideline. The Community Directive on the protection of individuals with regard to the processing of personal data and on the free movement of such data (Directive 95/46/EC) together with the opinions of the European Committee of Ethic (ECE) have been the basis for this report.

The variety of subjects includes the behaviour towards the other partners involved in the project as well as the handling of collected data such as user questionnaires, user ability assessment data, and text samples. Since sensitive and confidential data is involved special care has to be taken in protection of the privacy of the participating disabled persons. Particularly persons suffering from any kind of communication disorder, which is one of the target groups in the FASTY project, deserve in general a high degree of protection.

All data should be transformed and kept in a form, which avoids a trace back to a specific user or permits identification of a specific user for no longer than it is necessary. This was a guideline in the development of the Text Collection Tool that was mentioned above.

6.3 Text Collection

Text was collected from users with the Text Collection Tool. As described before this text material is needed for the prediction part in the developing phase. Special care was taken to fulfil the ethical guidelines.

6.4 User Panel

The User Panel is in close contact with the developers and is asked for approval of user specific documents. Thus, a user centred design is achieved.

7 Technological Implementation and Quality Assurance

7.1 Technological implementation plan

The Technological Implementation Plan summarises the results that have and will be achieved in the project. A draft plan for FASTY is available.

By the end of year 1 of the three-year plan, the following was achieved:

- The prototype of a User Ability Assessment Tool (UAAT) has been developed for testing purposes.
- The prospective users have been contacted and the user panel has been founded.

- A FASTY web site in 5 languages has been created at http://www.fortec.tuwien.ac.at/fasty.
- A prototype of a pressure sensitive keyboard is available for testing and gaining first data.
- A Rapid Prototype showing a variety of prediction windows and settings has been developed.

Broad dissemination and use intentions for the expected outputs:

- The composition of the consortium ensures a strong commercial participation. Focus will be given to a medium time-to-market of 3-5 years.

7.2 Quality Assurance Plan

The Quality Assurance Plan should lead to procedures, techniques and tools that serve to ensure that a product meets or exceeds pre-defined or implicitly assumed standards during its life cycle. This means that responsibility for the product's quality does not end with project's completion date but lasts beyond that, according to the exploitation plan.

Quality control and assurance in the FASTY project have two aspects:

- Quality of the work - to be achieved by process standards. Process standards describe the organizational basic conditions.
- Quality of the final results - to be achieved by product standards. Product standards describe the process that leads to the quality of the product itself.

8 Dissemination and PR Activities

During the first year of the project, FASTY partners have carried out several activities for the dissemination of the expected results of the project and the benefits of the future use of its results.

8.1 Project presentation on the web

There is a presentation of the FASTY project on the web at the following site:

http://www.fortec.tuwien.ac.at/fasty

It is available in English, German, French, Dutch and Swedish.

Fig. 7 screenshot of the start-page

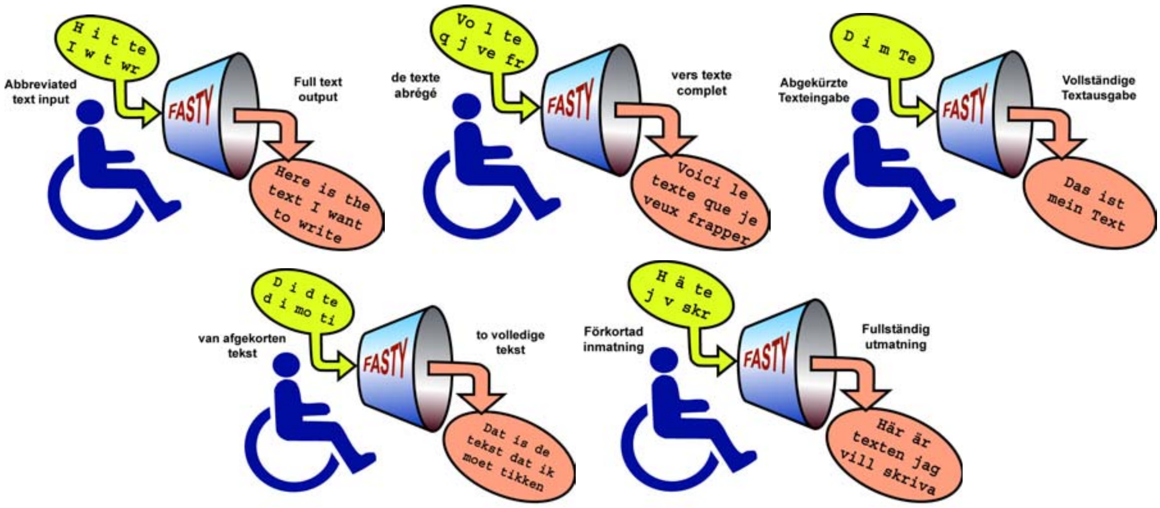

Fig. 8 one of the logos in all supported languages

8.2 Project Folder

The Project Folder, with an overview of the objectives and features of the project, has been issued in English, German, French, Dutch and Swedish and distributed at meetings, conferences and workshops. It is also available for download at the FASTY project pages.

8.3 Project presentations

The FASTY partners have participated in several conferences and other events, as detailed below:

- IST Concertation meeting, 1 March 2001.
- ScienceWeek@Austria, 11–20 May 2001.
- Medicon 2001 IX Mediterranean Conference on Medical And Biological Engineering and Computing, 12-15 June 2001.
- TALN 2001, Tours, France, 2-5 July 2001.
- AAATE, Ljubljana, Croatia, 3-6 September 2001.
- Regional ISAAC conference, Dortmund, Germany, 15 September 2001.

The FASTY partners have participated in the following workshops, seminars or courses:

- IST Multivision Workshop, Budapest, Hungary, 20 April 2001.
- FASTY Workshop, Germany, 12 June 2001.
- „9. Wiener Seminar Rehabilitationstechnik und Schulung im Dienste des Behinderten und seiner Umwelt" (Seminar on Rehabilitation Technology and Training, AKH Vienna, 13 June 2001.
- FASTY U1 Workshop
- Workshop, Creth, Belgium, 21 August 2001.
- RehaCare Fair, Düsseldorf, Germany, 5 October 2001.
- Fachtagung „Computer helfen heilen", Berlin, Germany, 9-10 November 2001.

- Workshop, Creth, Belgium, 4 December 2001.

In the following countries, articles about FASTY have been published in newspapers and also an interview on the radio has been carried out.

Austria:

- Two articles in ÖGAI (Austrian national scientific journal), July 2001.
- Radio broadcast feature on the Ö1 channel of the Austrian broadcast cooperation, 18 August 2001.

France:

- TALN paper for conference, July 2001.

United Kingdom:

- Article „Getting up to speed with FASTY" in issue 40 of Ability magazine of the British Computer Society, November 2001.

8.4 Presentation of FASTY to different interest groups

The FASTY project has also been presented to different interest groups. We detail hereinafter the following countries as concrete channels where some contacts were made.

Belgium:

- Presentation to Créneau.
- Presentation to Modem.

9 References

[1] Aurora. http://www.aurora-systems.com/

[2] Carlberger J. (1998): Design and Implementation of a Probabilistic Word Prediction Program, Royal Institute of Technology (KTH).

[3] Chelba C., Jelinek F. (1998): Exploiting Syntactic Structure for Language Modeling, in Proc. of 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics, Universite de Montreal,

[4] Copestake A. (1996): Applying Natural Language Processing Techniques to Speech Prostheses, in Working Notes of the 1996 AAAI Fall Symposium on Developing Assistive Technology for People with Disabilities.pp.225-231.

[5] D4.1 Quality Assurance Plan

[6] D4.2 User Abilities, Preferences and Needs – FASTY 2001

[7] D4.3 Report on Ethical Aspects

[8] D5.4 Overall Specifications and Specifications of Interfaces – FASTY 2002

[9] D6.1 User Ability Assessment Tool and Documentation – FASTY 2001

[10] D6.2 Rapid Prototype and first Demonstrator for User Interface – FASTY 2001

[11] D7.1 Definition of Language Components – FASTY 2001

[12] Daciuk J., Mihov S., Watson B., Watson R. (2000): Incremental construction of minimal acyclic finite-state automata, Computational Linguistics, 26(1).

[13] Darragh J.J. and Witten I.H. (1992): The reactive keyboard. Cambridge University Press, Cambridge, England.

[14] Dutoit T., Pagel V., Pierret N., Van der Vreken O., Bataille F. (1996): The MBROLA Project: Towards a Set of High-Quality Speech Synthesizers Free of Use for Non-Commercial Purposes, Proc. ICSLP 96, Philadelphia. Also on website: http://tcts.fpms.ac.be/publications.html

[15] Goldsmith J. (2000): Unsupervised Learning of the Morphology of a Natural Language,unpublished manuscript, University of Chicago. Available online at http://humanities.uchicago.edu/faculty/goldsmith/Linguistica2000/pdf/paper.pdf

[16] Goldsmith J., Reutter T.(1998): Automatic Collection and Analysis of German Compounds, in Busa F., et al.(eds.), The Computational Treatment of Nominals, Universite de Montreal, Canada, pp.61-69.

[17] http://www.cse.unsw.edu.au/~cs9414/notes/ml/06prop/id3/id3.html

[18] Hunnicutt, S. and Carlberger, J. (in press). Markov models and heuristic methods improve a word prediction program. Augmentative and Alternative Communication

[19] Information about doc2Txt http://plaza27.mbn.or.jp/~satomii/soft/cui/doc2txt.html

[20] Information about DOM http://www.w3.org/DOM

[21] Information about XERCES http://xml.apache.org

[22] Jelinek, F. (1990). Self-organized language modelling for speech recognition. In A. Waibel & K.F. Lee (Eds.), Readings in speech recognition (pp. 450-506). San Mateo, CA: Morgan Kaufmann.

[23] Jelinek, F. (1991). Up from trigrams! The struggle for improved language models. In Proceedings of Eurospeech: Vol. 3, (pp. 1037-1040). Geneva: ESCA

[24] McCoy K.F., Demasco P. (1995): Some Applications of Natural Processing to the Field of Augmentative and Communication, Proc. of IJCAI-95 Workshop on Developing Applications for People with Disabilities, Montreal, Canada.

[25] Microsoft speech development kit (SAPI) http://www.microsoft.com/speech/

[26] Pagel V., Lenzo K., Black A.W. (1998): Letter to Sound Rules for Accented Lexicon Compression, ICSLP98, Sydney. Also on website : http://tcts.fpms.ac.be/publications.html

[27] Palazuelos-Cagigas S. (2001): Contribution to word prediction in Spanish and its integration in technical aids for people with physical disabilities. PhD Thesis, Universidad Politecnica de Madrid.

[28] Quinlan J.R (1993). C4.5: Programs for machine learning, San Mateo, CA Morgan Kaufman,1993

[29] Rosenfeld R. (1996): A Maximum Entropy Approach to Adaptive Statistical Language Modeling, Computer Speech and Language, 10, pp. 187-228.

[30] Rosenfeld R. (2000): Incorporating Linguistic Structure into Statistical Language Models, Philosophical Transactions of the Royal Society, Series A, 358(1769), pp.1311-1324.

[31] Sågvall Hein A. (1999): A Grammar Checking Module for Swedish. In: Working Papersin Computational Linguistics & Language Engineering 12. Department of Linguistics, Uppsala University. 22 p.

[32] Sågvall Hein A., Starbäck P. (1999): A Test Version of a Grammar Checker for Swedish. In: Working Papers in Computational Linguistics & Language Engineering 12. Department of Linguistics, Uppsala University. 43p.

[33] Weijnitz P. (1999) Uppsala Chart Parser Light. System Documentation. In: Working Papers in Computational Linguistics & Language Engineering 12. Department of Linguistics, Uppsala University. 24 p.

[34] Wood M.E.J. (1996): Syntactic Pre-Processing in Single-Word Prediction for Disabled People, Department of Computer Science, University Bristol.

10 Appendix

10.1 Project Coordinator:

Institute of Industrial Electronics and Material Science, fortec
Research Group for Rehabilitation Technology
Favoritenstrasse 11/366-1B
A-1040 Wien
Contact: Dr. Wolfgang Zagler, G. Seisenbacher, C. Beck
Email: fasty-fortec@fortec.tuwien.ac.at
Tel. +43-1-58801-76611
http://www.fortec.tuwien.ac.at/fasty

10.2 Project Partners:

ÖFAI:

Oesterreichisches Forschungsinstitut fuer Artifical Intelligence
Contact: Dr. Ernst Buchberger, Prof. Dr. Harald Trost
Email: fasty-ofai@fortec.tuwien.ac.at

FTB:

Forschungsinstitut Technologie-Behindertenhilfe der Evangelischen Stiftung Volmarstein
Contact: Dr. Helmut Heck
Email: fasty-ftb@fortec.tuwien.ac.at

UU:

Uppsala University Department of Linguistics
Contact: Prof. Anna Sågvall Hein
Email: fasty-uu@fortec.tuwien.ac.at

MULT:

Multitel ABSL
Contact: Dr. Ir. Stéphane Deketelaere
Email: fasty-mult@fortec.tuwien.ac.at

IGEL:

IGEL Elektronische Kommunikationshilfen GmbH
Contact: Dipl.-Ing. Holger Neumann
Email: fasty-igel@fortec.tuwien.ac.at

ELI:

Seraphisches Liebeswerk für Tirol und Salzburg,Elisabethinum Axams
Contact: Bernhard Frischmann, Stefan Mina
Email: fasty-eli@fortec.tuwien.ac.at

IKuT:

Ingenieurbüro für Kunst und Technik II
Contact: Ing. Jörg-Michael Lindemann
Email: fasty-ikut@fortec.tuwien.ac.at

FUNDP:

Facultés universitaires Notre-Dame de la Paix
Contact : Geneviève Bazier, Jean Pierre Peters, Bruno Plumat
Email : fasty-fundp@fortec.tuwien.ac.at

Peer Review Report *Jiri J. Vasa,* ***Queen's University, CA***

Deliverable: D2.5 – 1st Edited Annual Report

Rating (tick)	**Excellent**	**Good**	Satisfac-tory	**Poor**	**N/A**
Aims and goals: clear and justified	X				
Methods of investigation: described clearly and well thought-out	X				
Analysis: described clearly and well thought-out	X				
Results: complete and presented clearly	X				
Discussions and Conclusions: valid and justified by the results	X				
Structure and presentation: length, layout, figures, style…	X				
State of text: does it still need editing?	X				

Rating (tick)	**Excellent**	**Good**	Sufficient	**Moderate**	**Insufficient**
Summary	X				

<u>Comment:</u>

This is a good document. While some of the above criteria apply only loosely to an Annual Report, the overall quality of the document is as stated.

Signature: J.J. V a s a   Date: 23 April 2002