

# Preliminary Results about VITIPI Evaluation Efficiency

Philippe Boissière, Daniel Dours

IRIT – UPS – CNRS, 118 Route de Narbonne, 31 062 TOULOUSE Cedex France  
{boissier, dours}@irit.fr

**Abstract.** We propose preliminary results about VITIPI efficiency. These results were obtained with 2 corpora and writing simulation.

## 1 Introduction

VITIPI is a writing assistance system. VITIPI's aim is to increase speed text acquisition in all computer applications. Unlike the other systems, it doesn't display lists of unexpected words but when the user types one letter, the system displays either the ending of the word, or a part of it. As long as the word remains incomplete, the user provides letters to the system depending on the context. It takes into account up to 10 previous words for prediction. As soon as new sentences are entered, they are integrated into the system, and can be used for further predictions. In this poster, we will first explain the methodology by showing evaluation background, giving efficiency notion, describing writing simulation and corpora composition. Secondly, we will expose testing results with known and unknown sentences. Finally, we will draw conclusions of these results.

## 2 Methodology Evaluation

Rigorously, an evaluation is made up of five ergonomic criteria: utility, usability, efficiency, no dangerousness and user's satisfaction [1]. In the field of writing assistance some of these criteria are trivial. An evaluation framework for such systems has been proposed [2] that take into account *system features* and *system performances*. In this paper we will just focus on the system performances by computing efficiency criterion which is the most important one according to user's point of view. To evaluate it, we have computed the ratio of input keys typed by the user compared to the ratio of output characters provided by the system output. This defines the ratio time saving (*R.T.S*) that is similar to Zagler keystroke saving [3].

If the user has to use a writing assistance system, it means that he has difficulties to write. So it seems impossible to ask him to write a great amount of text just for system evaluation. So, a testing program has been realized to write in user's place. The testing program gives letters to the writing assistant system that provides the

ending of word or a part of it. User's keys and system output symbols are computed to evaluate the R.T.S. formula.

Evaluations were made with 2 different corpora. The first one deals with weather forecasting French sentences collected during 24 days. It is compounded of 439 sentences, 7,830 words and 985 words for vocabulary size. The second one contains 30 letters wrote by a French disabled association (for entertainment meetings). It is compounded of 499 sentences, 7,579 words and 1,275 words for vocabulary size and was divided into 3 sub corpora. An additional standard lexicon made up with 5,930 useful French words can optionally be set.

### 3 Testing Results

Various parameters were analyzed [2]. We have got to outline that for all corpora, if the standard lexicon option is set, R.T.S is lesser than the one obtained with unset option. For example, when first corpus is used and faced to unknown sentences, taking into account 3 previous words, the R.T.S reaches 45 % with standard lexicon against 47 % without it. When all sentences belong to corpus, taking into account 10 previous words, R.T.S. climbs up to 78.9 % without standard lexicon. Another testing has been made with the second corpus without standard lexicon. At the very beginning, system does not contains any sentence or word. When first sub corpus is typed with testing program (134 sentences), the R.T.S. reached 28 %. After that, first sub corpus is known by system, but the second sub corpus is unknown. When it is typed by testing program (159 sentences), R.T.S goes up to 38.8 %. Finally, first and second sub corpus are known by system, when the third one is typed by testing program (206 sentences), R.T.S climbs up to 48.2 % of predicted output.

### 4 Conclusion

It has been shown that with very specialized documents with some stereotyped sentences such as weather forecasting, results are high. With ordinary texts such as office letters, reports, ... , results seem promising. Other experiences should be done.

### 5 References

1. Sperandio, J.C., *Critères Ergonomiques d'Assistance Technologique aux Opérateurs*, JIM'2001 : Interaction Homme-Machine & Assistance pour Grand Public in Systèmes Complexes / Handicap, p. 30-37 Metz, 4-6 juillet 2001.
2. Boissière, Ph., Dours, D., *A Proposal of an Evaluation Framework for Writing Assistance Systems: Application to VITIPI*. In Computers Helping People with Special Needs, Proceedings of the 8<sup>th</sup> ICCHP'2002. LNCS 2398, pp 276-278 Springer-Verlag 2002.
3. Zagler, W.L Seisenbacher, G. *German Language Predictive Typing : Results from a Feasibility Investigation*. in Proceedings of the ICCHP'2000, pp.771-779, Karlsruhe, 17-21 July 2000.