Adaptation of AAC to the Context Communication: A Real Improvement for the User Illustration through the VITIPI Word Completion

Philippe Boissière, Nadine Vigouroux, Mustapha Mojahid, and Frédéric Vella

Université Paul Sabatier, Institut de Recherche en Informatique de Toulouse,
UMR CNRS 5505, 118, Route de Narbonne
F-31062 Toulouse Cedex 9, France
{boissier,vigourou,mojahid,vella}@irit.fr

Abstract. This paper describes the performance of the VITIPI word completion system through a text input simulation. The aim of this simulation is to estimate the impact of the linguistic knowledge base size through two metrics: the KeyStroke Ratio (KSR) and the KeyStroke Per Character (KPC). Our study shows that the performance of a word completion is depending of the % of words not available and the size of the lexicon.

Keywords: AAC Word completion system, KSR and KSPC metric, AAC.

1 Introduction

Many researches on text entry on the area of assistive and augmentative communication have been conducted (for an overview, see [6]). Text prediction systems were initially designed to help people with a low text composition rate. This includes people with severe speech and motor disabilities (cerebrally and physically disabled persons, Locked-in syndrome, cerebral palsy, etc.). The main aim of a predictor system is both to reduce the effort required and the message composition time. To reduce the effort, it is necessary to decrease the number of keystrokes needed for composing a message by anticipating the next block of characters (letters, syllables, words, sentences, according to the predictor system nature).

To facilitate the text entry, many solutions consisting in displaying a candidate word list have been designed and tested. The scientific community has clearly identified some theoretical improvements brought by the prediction system (for instance, the low keystroke number and the accelerated speed rate), while the use studies shows that the prediction system are rarely actually used.

This paper will present the new VITIPI version, a word completion component of Augmentative and Alternative Communication system designed in the framework of the PALLIACOM project: `recherche.telecom-bretagne.eu/palliacom`.

After our review of related works, we will introduce the principles of word completion algorithm. Then, we will describe our experiment to demonstrate the impact of

K. Miesenberger et al. (Eds.): ICCHP 2012, Part II, LNCS 7383, pp. 451–458, 2012.
© Springer-Verlag Berlin Heidelberg 2012

size corpora on completion system's performance trough the KSR and KSPC parameters. Finally, results will be presented and discussed.

2 State of the Art

Most Augmentative Alternative Communication (AAC) systems include method to speed up communication by trying to save the number of keystrokes needed for the composition of a message. There are two main methods: word completion and word prediction. Both are basic technologies for text entry as well as an important component in augmentative language technology tools. Their main aim is to reduce the number of keystrokes typed during text entry and, consequently the tiredness of the person. We will illustrate through an example the differences between then.

Suppose that the lexicon contains the following list of French words:

- directeur
- direction
- directrice
- égalisateur
- égalisation
- égalisatrice

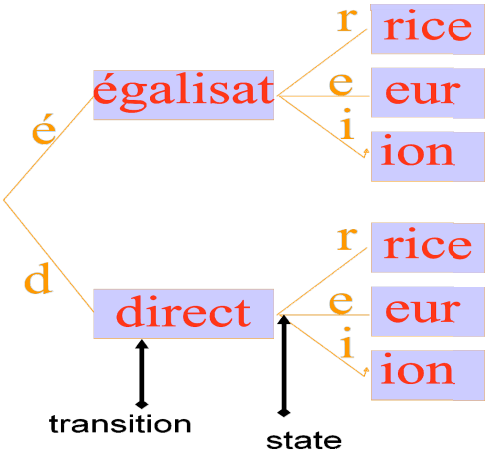

Fig. 1. Linguistic representation through a transducer

2.1 Principles of Completion Mode

When the user has typed the letter "*d*", the system automatically displays "*irect*" until there is no ambiguity for the generation process; then the user can enter another letter ("*e*" "*i*", "*r*") and the system completes by the appropriate sequence of letters "*eur*", "*ion*", "*rice*" according the linguistic representation (Fig. 1). In this mode no word list was presented to the user.

The completion system has only added the letters that could be completed automatically. However the user can refute them if they do not suit it. The most popular system is the T9, used on the standard 12-key keypad on the majority of mobile phones [10].

2.2 Principles of Prediction Mode

When the user has typed the letter "*d*", the prediction system proposes the list of words beginning with this letter: "*directrice*", "*directeur*", "*direction*" for the example. The word list can be ordered by decreasing probability or alphabetically according the system characteristics [6]. This list may be truncated to the first n elements (generally between 5 and 7).

2.3 Background Studies

These two modes were investigated through numerous studies to improve text entry: A list of words is provided: for instance, Dasher [15], [2], FASTY [12], SIBYLLE [14], etc. A survey of text prediction systems can be found in [6]. These studies reported encouraging user experiences.

However [14] mentioned that some users did not select the intended word even though it is clearly present in the prediction list. [5], [9], [8] explained that user's cognitive load is due to more attention and visual efforts. In fact they have to continuously look between the word list, the keyboard and the text entry focus. [1] has proposed a new distribution of the keyboard around a word list to avoid this problem.

A possible solution to reduce the cognitive load could be to implement direct "word completion" [3], [13] by proposing the most probable termination of the current word immediately after the latest typed letter by the writer.

3 Principles of the VITIPI Completion System

The VITIPI system principles consist of completing (automatically typing by the computer) the word, in part or in whole, while the user is inputting it. After each character typed by the user, the VITIPI system proposes a string (which may be empty), which extends or completes the word. The VITIPI linguistic kernel is based on a transducer [3] without probability (Fig. 1). VITIPI models the previous context through a N-gram model like.

4 Experiment

The aim of this experiment is to estimate the evolution of VITIPI performance by taking into account different representations of linguistic structures (words, sentences) through the transducer by means of a simulation system.

4.1 Simulation Principles

To value from a theoretical point of view the text entry performance a simulation program was developed. The text that the writer could have written is given to the simulation system: each character is entered one after one without orthographic fault. At each character typing step, the completion system is trying to propose, as soon as

possible, a completion string. Then a comparison string algorithm compares the completion string to the word part string expected. Two assessment states are possible:

- Strings are identical: the completion is counted as right; the completion string is automatically inserted ; then a new character is entering;
- Strings are not identical: the completion is counted as false; the completion string is removed from the string; then the current character is entering.

The process is iterative until the end of word. A log program captures the following data to compute the letter number to be entered, the letter number proposed by the simulation system and the number of false completion list.

4.2 Corpora

The topic of the corpora is weather forecast. We have chosen to evaluate the VITIPI system on the field of meteorology because its lexical coverage is relatively small (see description below). This choice relies on the following hypothesis: the word completion must be adapted to the application field (for instance, personal mail, written communication topic, etc.); then the word completion is more efficient; in consequence this treatment reduces the subject's fatigue for long text entry [2].

Each sub corpus (coded from 1 to 33) corresponds to the text of a daily weather. The set of the 33 sub corpora is composed of 457 sentences including 8 436 words. The lexicon size corresponds to 948 words. The context size retained for the VITIPI system is 3-gram.

4.3 Simulation Task of Typing Text

The test has consisted to simulate the text entry of the sub corpus (daily j) with the transducer which models the set of the (j-1) sub corpora. Two serials of test were run according two modes: chronological and randomly mode for the construction of the transducer.

4.4 The Metrics

Two metrics were used as evaluation parameters:

- The Keystroke Saving Rate (KSR) [4] which is an estimation of the percentage of letters typed by the user. The KSR estimates the percentage of letters automatically written by the system.

$$KSR = \left(1 - \frac{Nb\_char\_typed + Nb\_function\_keys}{Length\_of\_text}\right) * 100 \quad (1)$$

where:

- *Nb_char_typed* is the number of characters typed by the user to write his/her text,

- *Nb_function_keys*, the number of function keys used by him/her to correct erroneous word completion;
- and *Length_of_text* is the text length to be typed.

— The KSPC (KeyStroke Per Character) generally used by the text entry community [11] to compare the performance between prediction systems.

4.5 Results

We have measured the effect of the transducer enrichment, daily corpus by daily corpus, by means of the KSR parameter. Two measure sets have been done according the two modes of the transducer construction.

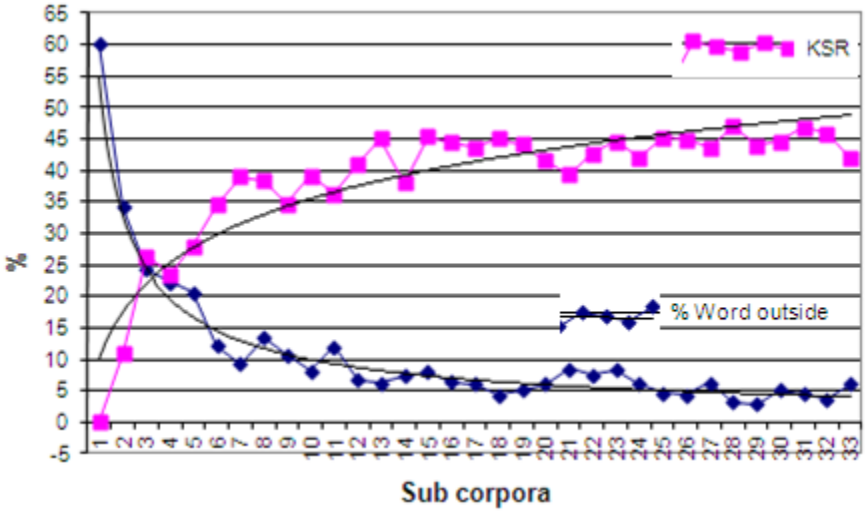

Fig. 2. KSR and rate of words outside the lexicon (chronological adding of sub corpora)

Firstly, the figures (Fig. 2 & Fig. 3) show that there is no effect of the adding mode of the sub corpora in the KSR performance. Secondly, we note that for the sub corpus 1 (Fig. 2) and for the sub corpus 23 (Fig. 3) the KSR is respectively null and negative (-0.85) because the transducer has still no representation of words. These two figures demonstrate also the great importance of the learning phase to update the lexicon and the linguistic representation with the sentence structure. KSR is increasing very quickly and then move around 45.75% for the transducer modeling the set of 15 sub corpora (Fig. 2).The adding of new sub corpora (from 16 until 33) has quite no effect on the KSR. The same effect is observed on the (Fig. 3) (after the consideration of 16 sub-corpora the KSR is around 44,5).

The figures (Fig. 2 & Fig. 3) demonstrate also that for a percentage of words not available in the lexicon lower than 5% the KSR is quite stable. This suggests us to formulate the hypothesis that for a lexicon of 1 000 words, it would be possible to obtain a stable KSR average around 45.

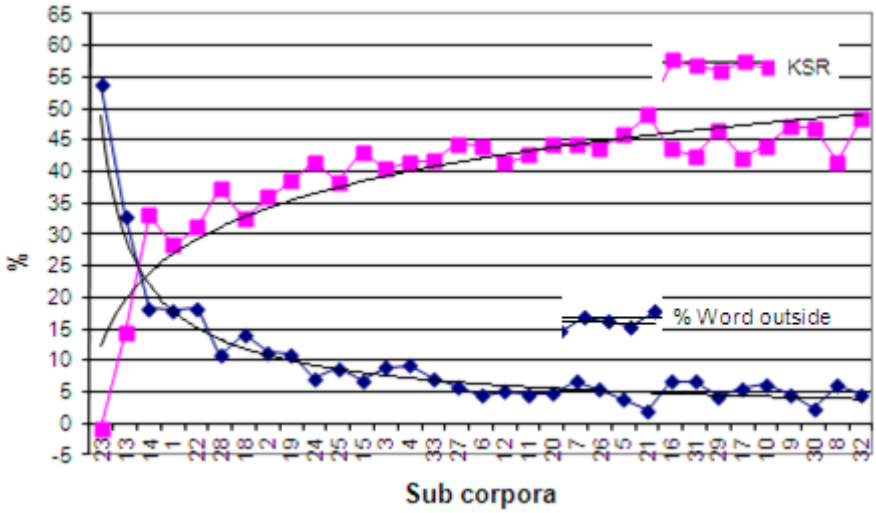

Fig. 3. KSR and rate of words outside the lexicon (random adding of sub corpora)

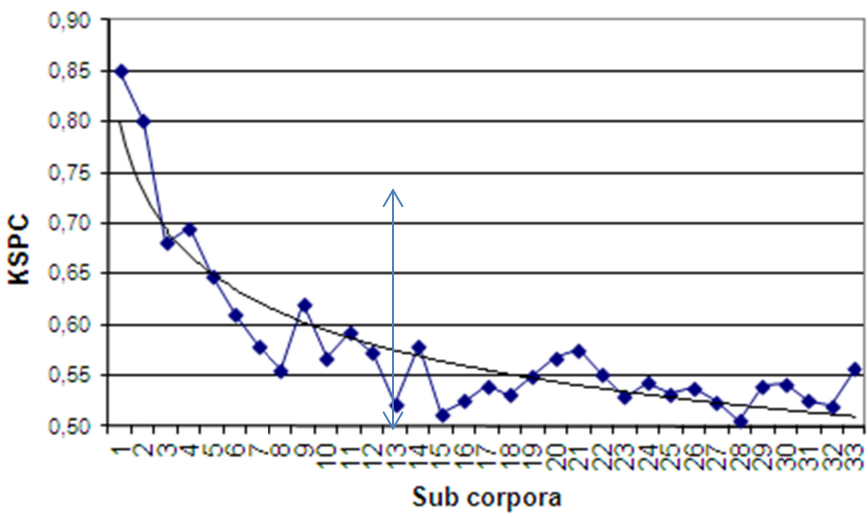

Fig. 4. KSPC (chronological adding of sub corpora)

The KSPC parameter has an opposite variation to the KSR parameter (Fig. 4 & Fig. 5). It strongly decreases until a set of 13 (Fig. 4) or of 10 (Fig. 5) sub corpora are taken into account in the transducer representation. Then it is in the range [0.50, 0.57] with few little variations.

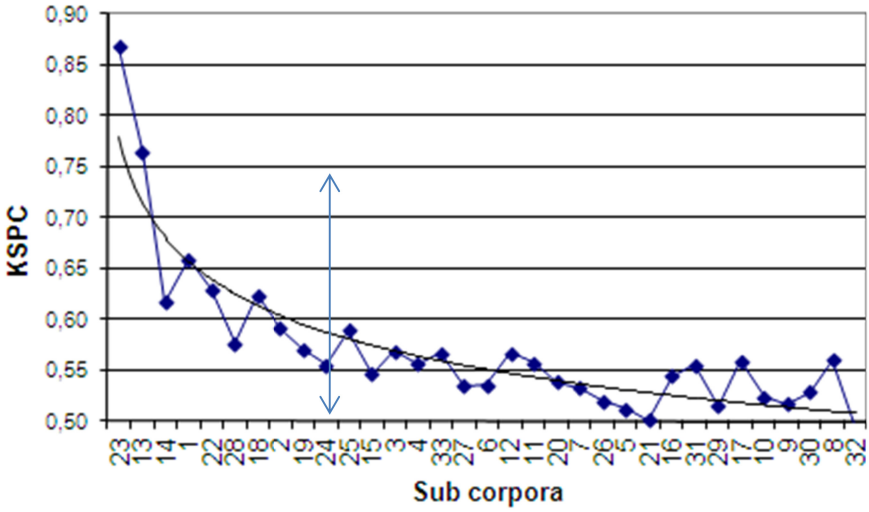

Fig. 5. KSPC (randomly adding of sub corpora)

4.6 Discussion

When the KSR is compared to the percentage of new words from the lexicon size (Fig. 2 & Fig. 3), we observe that these two curve variations are antagonist: growth corresponds to a decline of the other and vice versa. The experiment reports that the highest is the number of entries in the lexicon, the lesser the VITIPI system completes the user's current input. And vice versa, the more the lexicon will be reduced, the more the VITIPI system will perform

New words is also strongly penalizing for the KSR. When the word is unknown of the VITIPI linguistic representation, the writer has to type all letters. Moreover for the weather domains, the new words of the lexicon are often proper nouns: Consequently, it is impossible to use inference process of the VITIPI system. These results postulate that it extremely important to built quickly representative transducer.

To validate the robustness of these results it is necessary to study the correlation between the KSPC and KSR parameter for more hug corpora including words which are not specific of the domain.

5 Conclusion and Perspectives

We have demonstrated the variation in performance of a word completion system. This change is depending of the % of words not available and the size of the lexicon. This result demonstrates that "finalized" corpora might be defined according to the language register of the user. These results are extremely encouraging to adapt AAC to the user. However, these results must be consolidated on longer simulation trials and other topics (daily communication for instance).

Acknowledgement. This work is partially funded by the ANR TECSAN (PALLIACOM Project).

References

1. Badr, G., Raynal, M.L.: CentraList, AAATE'2011, Everyday Technology for Independence and Care. In: Gelderblom, G.J., et al. (eds.), pp. 944–951. IOS Press (2011)
2. Bérard, C., Neimeijer, D.: Evaluating effort reduction through different word prediction systems. In: Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, La Haye NL, vol. 3, pp. 2658–2663 (2004)
3. Boissière, P., Dours, D.: From a specialized writing interface created for disabled, to an universal interface for all: The VITIPI System. In: 1st International UAHCI 2001 Conference (Universal Access for Human Computer Interfaces), News-Orleans, August 5-10, pp. 895–899 (2001)
4. Boissière, P., Schadle, I., Antoine, J.-Y.: A methodological framework for writing assistance systems: applications to Sibylle and VITIPI systems. Edited by Association for the Advancement of Modeling & Simulation Techniques in Enterprise, AMSE-Journals, Modeling C 67 (supp. Handicap 2006), 167–176 (2006)
5. Dubus, N.: Evaluation de l'interface intelligente d'aide à la saisie informatique, VITIPI au ly-cée Le Parc Saint-Agne. Journal d'Ergothérapie, 95–100 (Mars 1996)
6. Garay-Vitoria, N., Abascal, J.: Text prediction systems: a survey. International Journal on Universal Access in the Information Society 4(3), 88–103 (2006)
7. Garay-Vitoria, N., Abascal, J.: Modelling text prediction systems in low- and high-inflected languages. Computer Speech and Language 24, 117–135 (2010)
8. Goodman, J., Venolia, G., Steury, K., Parker, C.: Language modeling for soft keyboards. In: IUI 2002: Proceedings of the 7th International Conference on Intelligent User Interfaces, pp. 194–195. ACM, New York (2002)
9. Koester, H.H., Simon, P., Levine, S.P.: Modeling the Speed of Text Entry with a Word Prediction Interface. IEEE Transactions on Rehabilitation Engineering
10. MacKenzie, I.S., Kober, H., Smith, D., Jones, T., Skepner, E.: LetterWise: Pre_x-Based Disambiguation For Mobile Text Input. In: UIST 2001: Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology, pp. 111–120. ACM, New York (2001)
11. Soukoreff, R.W., MacKenzie, S.I.: Metrics for text entry research: an evaluation of MSD and KSPC, and a new unified error metric. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (2003), pp. 113–120 (2003)
12. Trost, H., Matiasek, J., Baroni, M.: The Language Component of the FASTY Text Prediction System. Applied Artificial Intelligence 19(8), 743–781 (2005)
13. van den Bosch, A.: Effects of Context and Recency in Scaled Word Completion. Computational Linguistics in the Netherlands Journal 1, 79–94 (2011)
14. Wandmacher, T., Antoine, J.-Y., Poirier, F., Departe, J.-P.: SIBYLLE, an Assistive Commu-nication System Adapting to the Context and its User. ACM Transactions on Accessible Computing 1(1), Article 6, p. 6.1-6.30 (2008)
15. Ward, D., Blackwell, A., Mckay, D.: Dasher: A Data Entry Interface Using Continuous Gestures and Language Models. In: Proceedings of the 13th Annual ACM Symposium on User Interface Software and Technology, UIST 2002, San Diego, CA, pp. 129-137 (2000)