Multilingual dictionary construction and services case study with the Fe$*$ projects

Mathieu Lafourcade

GETA - CLIPS - IMAG Campus
B.P. 53 — 38040 Grenoble cedex 9
Tél. : 04.76.51.43.80 - Fax : 04.76.51.44.05
email: Mathieu.Lafourcade@imag.fr

UTMK-USM
11800 Penang - Malaysia
email: lafourca@cs.usm.my

(submission to PACLING'97 - September 2/5 1997 - Meisei University - Ohme, Tokyo, Japan)

Summary

This paper focuses on the building methodology for multilingual dictionaries and electronic distribution issues. Two projects of French-English-Malay and French-English-Thai dictionaries are presented and analyzed under these aspects. We describe how lexicographer works can be done with a low technological profile and how distribution can be made through different categories of dictionary services and tools. We make the distinction between dedicated dictionary tools and generic approaches, but we also point out the implication of local versus remote databases on both the technical level and user experience.

Keywords

multilingual dictionaries, dictionary construction, dictionary services, tools and servers

Introduction

Two multilingual dictionary construction projects — containing respectively French-English-Malay (FeM project, achieved) and French-English-Thai (FeT project, in progress) — are the occasion to present and analyze (1) the methodology of their construction and (2) possible strategies for their electronic distribution. This task is undertaken beside the aspects concerning the printed format production and distribution. The printed FeM dictionary has been officially launched in July 1997 in Kuala Lumpur [2].

In both projects, we are targeting two kinds of dictionaries. The first one is a general dictionary of about 20.000 entries (around 50.000 word senses), containing information about the pronunciation, the word category, the gender, some meaning refinement, some phrases and locutions (more than 11.000) and some sentence examples (more than 3.000). Novice learners (Malay, Thai and even French) should be able to use these dictionaries with ease and profit. The second type of dictionary [4] is a lexicon in computer science terminology (also trilingual). It contains roughly 5.000 entries in each language, with category and pronunciation but without examples, nor locutions.

This paper first presents the methodology we adopted for the construction of these dictionaries. As (1) dictionary building is a complex collaborative project and (2) computer resources were scarce, we have deliberately taken a low technical level approach by not using specialized tools. On the other side, for the electronic distribution, we investigated (quite) high technical solutions. For this second project phase, we present different kinds of dictionary services we developed to make lexical resources accessible and exploitable to end-users.

1. Multilingual dictionary construction

a. English Centered Multilingual Fork dictionaries

The inclusion of English equivalents, as a second source language, among the source language (French in our case) and the other target languages (Malay, Thai, etc.), is crucial for lexicographers and linguists. It reduces the level of ambiguity or error risks when looking for and producing equivalent translations in target languages. Moreover, it is often quite difficult to find lexicographers having a very deep knowledge of French in countries which have not been during their history especially exposed to French (contrary to Vietnam, for instance). The presence of English translation is thus a worthy mine of information.

A fork dictionary associates to each work meaning (also called « word acceptation ») of the source language, one or several meanings in the target language. It is the usual structure of bilingual dictionary, but it is in our case extended for several target languages.

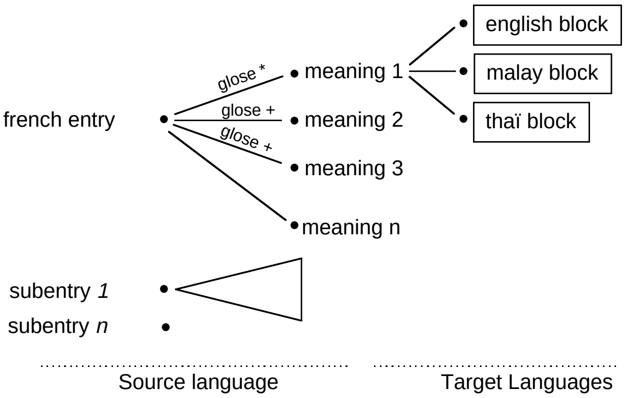


Figure 1. Fork multilingual dictionary structure. The first meaning is the more general and may not be linked to one or several glosses. Other meaning correspond to glosses, labels or different categories.

As a general schema, a French entry (the source language) is divided into several meanings by the help of glosses, labels or categories. A given entry can contain subentries (generally pronominal verbs). For a given meaning, several equivalents for each target languages and optionally some sentence examples. or locutions. In annex (figure 3), we give a sample for a French-English-Malay-Thai dictionary, with the type of each field.

For each target language, equivalents are grouped into independent blocks. We consider for each French meaning, an equivalent block for English, Malay and Thai. Each of these blocks, can refine their meaning by the use of glosses or labels in target languages. Target languages glosses are to be opposed to source language glosses which fork meaning for target languages. Target language glosses do not interfere among each other. Thus, by keeping blocks independent, we can read the dictionary from the source language to one target language (by filtering out alternative target languages).

b. Word meanings and glosses

When there are several acceptations for one French word, we tend to associate one entry to one acceptation. In fact, it is basically a lexicographer choice to split (or not) a given word into several entries if more than one meanings are recognized. Inside one given acceptation, we use glosses for refinement in the source language and in target languages.

In case of contrastive phenomena, a gloss allows one to differentiate the meaning in the target language. For example, there are three possible equivalents for « rice » in Malay : *padi* (the rice still in the field), *beras* (not husked rice) et *nasi* (cooked rice). There is only one word in French, *abats* which groups the English meanings of *offal* and *giblets.*

Labels do not directly participate to meaning refinement, but instead indicate some usage context (as*agric.* for agriculture) — or added information (as *abrév.* for abbreviation or as *vx.* for old). Labels take their values into a finite set, contrary to glosses which are open.

c. Construction methodology

A large amount of the French-English-Malay data has been created by crossing of French-English and English-Malay lexical resources. This rough material has then been corrected, completed and refined. Many word meanings have been revised, especially for many subentries which have been promoted as entries. For example, « kilométrage » (mileage) was a subentry of « kilomètre » (kilometer). We mainly kept pronominal verbs as subentries of the their non pronominal form.

Lexicographers used standard word processors (as Microsoft Word™) to complete dictionary entries. For each field type, we associate a unique typographic style. This allows us both to ease the lexicographer work and to make possible the data conversion between various formats without losing the entry structure. Also worth noting, the use of typographic attributes reduces risks of logical type confusion compared to a labeled plain text approaches.

In the case of the FeT project, Thai equivalent blocks are added next to Malay equivalent blocks. Moreover, for each Thai equivalent, we add a phonetic transcription readable for French speakers). Such transcriptions exist, but not one seems to be (1) either complete (2) nor adapted to our needs (3) nor easily readable. We are currently specifying such a transcription that will be readable for a French reader. It must also include word separators, as Thai writing does not provide for something equivalent to space character. In the working files, we will keep Malay in parallel with Thai, which is often useful for coherence checking. In the printed version, Malay will not be included, nor English, but both will still be accessible in computer versions.

During the building phase of the dictionary, we use a transcription free of all special characters and no special formatting. This allows us to freely exchange data between various systems and codings (MacOS, Windows and UNIX) through networks. For instance, the French word « dictionnaire » (dictionary) will be transcripted “^phot_ja-na=^nou-krom”. This transcription is readable for French speakers and relatively faithful to Thai pronunciation (it includes tones). We are also currently designing other non ambiguous transcriptions more adapted to English speakers (and others) and tools for automatic conversion among various transcriptions.

The primary work is done on styled Microsoft Word™ files where each logical elements (entry, basic pronunciation and variants, category, label, gloss, locutions, sentence examples, and their respective Thai equivalents, …) is associated to one logical paragraph style. File structure is nearly identical to the one used in the FeM project.

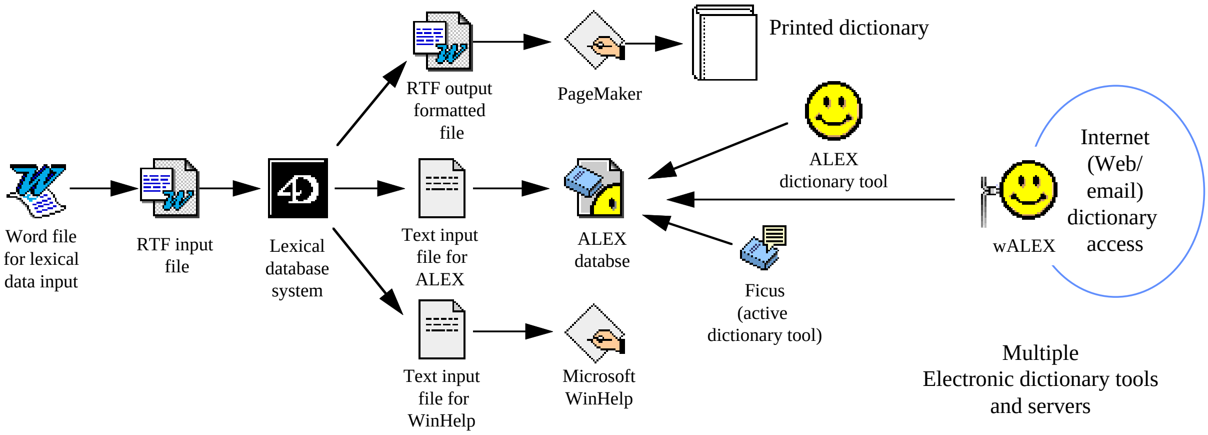

Printed dictionary
RTF output formatted file
PageMaker
ALEX dictionary tool
Internet (Web/ email) dictionary access
Word file for lexical data input
RTF input file
Lexical database system
Text input file for ALEX
ALEX databse
wALEX
Ficus (active dictionary tool)
Text input file for WinHelp
Microsoft WinHelp
Multiple Electronic dictionary tools and servers

Figure 2. Major processing stages in the Fe* dictionary production.

Saved as RTF (Rich Text Format), the working files are imported in a database system (Fourth Dimension), where some checking and counting can be easily undertaken. This project phase is done and controlled by computer scientists (such a task would not likely be taken over by lexicographers). From the database various output formats can be produced. For instance:

- a Microsoft Word™ source, dully formatted and compacted which feeds a PageMaker™ document for the printed version production (with a professional page layout);
- a Microsoft Help™ source, allowing us to produce a (very basic) electronic version of the dictionary on Windows. This dictionary can be viewed with the Microsoft Help Viewer™;
- an ALEX source allowing us to obtain an advanced electronic version of the dictionary on MacOS. This tool (see thereafter) allows the user to filter out information, to sort and count them, …). The same ALEX database can be used for making the dictionary available through the Internet (either by the Web or by e-mail).

During this project, several observations have been made. First, lexicographers prefer to work with standard software packages widely available like word processors more than Databases. Basic operations (like cutting, pasting, printing, word counting, …) seem more versatile in word processors than database engines. Also, word processors offer a high visibility of entry contexts. With or without reasons, database systems do not constitute the privileged tool of linguists. Secondly, for technical, management and relational reasons, version management packages are generally not suited to large lexicographic projects. One person should be clearly identified as the repository of the reference state of the development of dictionaries.

d. Construction costs

The global investment for the building of the FeM dictionary was of 9 man-year (108 man-month). The working experience and the data issued from this project is a good starting point for the FeT project.

The estimated work for the building of the FeT dictionary is of 2 man-years for the production of Thai equivalents. But we should not forget other task such as: (1) the production of the Thai phonetic information by automatic transcription, (2) the checking processes, (3) the lexical database production, (4) the production of the various output formats (Microsoft Word™, PageMaker™, Microsoft Help™ and ALEX), and (5) the production of the final version (printed and CD-ROM). All these tasks are estimated to one man-year. This delay can be kept relatively short thanks to the tools and procedures developed during the FeM project.

2. Electronic Dictionary Services

Services offered by computer dictionaries go well beyond what is usually expected: more or less extended cut and paste possibilities, very beneficial logical size; (entry number) vs. physical size (file size) ratio, etc. In effect, it is possible and desirable to offer services one cannot find close equivalents for unless electronically. We can already mention some of such possibilities: (1) colors and graphics, (2) hypertext links, (3) multimedia information, (4) very large amounts of information, (5) personal filtering, formatting and adaptation to user needs and desires.

a. Data formatting for end users and autonomous agent

We can consider all dictionary information which is not directly linked to either the source language or the target languages: categories, cross references, glosses, labels, etc. This is what is usually called the « dictionary metalanguage ».

It is possible to design dictionary tools and data where these information is in the native language of the reader (whatever this language could be). We can thus have a French-English dictionary with glosses in English, or a Japanese-French dictionary with a French metalanguage. In a more general way, the user can configure the dictionary tool so that that it will output the metalanguage into the language he desires irrespective of the source or target languages. To this respect, the user should be able to change these settings in a working session midcourse and to have access simultaneously to several views of the same data.

We can also consider, for a monolingual dictionary (definition dictionary), the possibility to have the definition in the language of the reader irrespective of the dictionary source language (for example, French definitions of Japanese words).

All that we considered above is related to end user experiences. But we are also investigating how to make our data available to complex NLP tasks (such as Machine Translation). We are currently evaluating some DBMT (Dialog Based MT) software architectures where autonomous agents are in charge for retrieving lexical knowledge among various remote large scale lexical databases.

For such autonomous agents, explicit labeled formats are definitely more appropriate. Such formats are more verbose than what can be produced either for printed dictionary or to end users, but it is not ambiguous and easily parsable. In that case size restriction (which is the main constraint for printed dictionary editors) is not a primary concern.

b. Dedicated dictionary reader with local and remote database.

It is possible to distribute lexical data under a compact and protected (by encryption) format and to made them exploitable by a « reader » (in the same spirit of Portable Digital Documents (PDF) and associated readers distributed by Adobe). From this point, we can consider two research and development directions.

First, data are available locally (on the end-user computer) and a dictionary tool (the reader) allows one to navigate among dictionary entries. The user can open several dictionaries, create and modify personal dictionaries, copy and paste (formatted) data in other applications, etc. We actually developed such a tool — ALEX. The ALEX format dictionary may be encrypted, making the reverse engineering of distributed dictionaries a difficult job.

Secondly, it is also possible to consider a dictionary tool which remotely fetches data on remote bases. The user has only the dictionary tool (and not the bases) to install on his computer. It is possible to have access to a virtually unlimited number of dictionaries without cobbling the computer resources. Dictionary updates become transparent to the user. We also developed such a tool, as a new version of ALEX (“ ALEX remote ”). Experience shows that, the functionality number and complexity should be kept reduced compared to locally based dictionary tools. This is practically due to bandwidth and server response time which has nothing to be compared with local computation.

We consider this second solution more adapted to autonomous agents dealing with no real time tasks (such a updates of local lexical database or merging of lexical information issued from multiple sources — as general dictionaries, specialized terminology banks, etc.).

c. Generic dictionary readers: Web access and e-mail

The FeM dictionary is available through the Web. The main advantage compared to a local dictionary tool, is the use of a Web browser (as Internet Explorer™ or Nescape Navigator™ for instance) as a generic application widely available. Request results are generated by the dictionary server and are not pre-computed. The form presented to the user through the Web browser interface can offer (or hide) several parameters allowing one to control the formatting, the range of information detail, information level, the languages (interface, metalanguage, sources, targets), the number and nature of dictionaries looked upon, and so on.

One striking feature of the Web is not only the ease of use, but also the reduced time of development needed to achieve query interfaces. The investment in time and effort (at least for rapid prototyping) is tremendously reduced comparatively to classical database systems. Moreover, this solution is reasonably portable among OSes — which is a very crucial aspect for research laboratories.

Web approach seems to be well accepted by end users but (1) people having access to the Internet, while increasing, do still represent a privileged minority. There is many more people having only email access ; (2) we desire also to make our dictionaries accessible and exploitable to unsupervised computer tasks (or to use a more hyped term, to « autonomous agents »). For example, email access to large scale lexical resources allows the regular batch updating of smaller linguistic local bases. Such an approach, although asynchronous, offers numerous advantages among which robustness. An email based request allows to specify with a very fine granularity what is looked for — but generally it goes well beyond what an end user would accept to do.

d. *Passive and active dictionary tools*

Approaches presented above, means that the user has to « go and fetch » the data. In effect, he has to activate the dictionary tool and key in the word to look up (at least). These tools are passive. Active tools initiate by themselves the lookup upon the monitoring of the user activity. For example, such a tool (as the « FICUS » prototype we implemented) will scrutinize selected words on the current application (cf. figure 8) and make the lookup. If the word candidate is found in the open dictionaries, the corresponding information is formatted and proposed to the user in a no intrusive way. The user is then free to consider this information and to manipulate it. As the user is not solicited, this kind of approach leaves the user free to concentrate on his work. This last aspect is crucial for usability.

From a technical point of view, a preprocessing (such a lemmatization, or morphological analysis) is needed before the lookup. Furthermore, the potential ambiguity (several lemmas issued from the word candidate) should be handled. For instance, in French « avoir » (to have) et « avion » (plane) pour « avions » (we had/planes). In English, we can find the noun « book » and the verb « to book » from « books ».

e. *Unique engine and common database*

ALEX is a simple dictionary tool with a « plug in » mechanism allowing one to convert it as a Web/Email dictionary server. ALEX is used as a dictionary engine, which retrieves information associated to a given entry and formats it. The multiple format architecture is extensible. Thus for each dictionary it is possible to dynamically manage several types of representation.

Conclusion

It is valuable to produce multilingual dictionaries and to be able to successfully distribute them. We consider that a low technical level approach for the building of multilingual dictionaries can be a main success factor. The systematic inclusion of English equivalents is made necessary for their building and for the sake of their quality. Production methodology should integrate computer tools mastered by lexicographers, which are essentially word processors. It is thus possible to have a precise management of project which does not disrupt lexicographer work. Whatever, the approach taken, a dictionary project cannot be considered as a small and easy task.

Beside the printed distribution, we have presented in this paper several dictionary services. Contrary to the construction task, dictionary services can be highly technical (but should remain simple to use). The typology of such services can be divided among the following criteria: local or remote database dictionary, dedicated or specific dictionary reader and tools, passive or active dictionary tools.

We hope that this work will inspire new dictionary projects and make flourish such services and tools, both usable for autonomous agents and end users.

References

[1] Gaschler, J. and M. Lafourcade (1994) *Manipulating human-oriented dictionaries with very simple tools.* Proc. COLING-94, August 5-9 1994, Makoto Nagao & ICCL, vol. 1/2, pp 283-286.

[2] Gut Y (1996)., R. Puteri, Z. Yusoff, C. K. Choy, S. A. Samat, Ch. Boitet, N. Nedeau, M. Lafourcade, J. Gaschler, D. Levenbach (1996) *Kamus Perancis Melayu Dewan - dictionnaire français-malais.* Dewan Bahasa dan Pustaka, Kuala Lumpur, 667 p

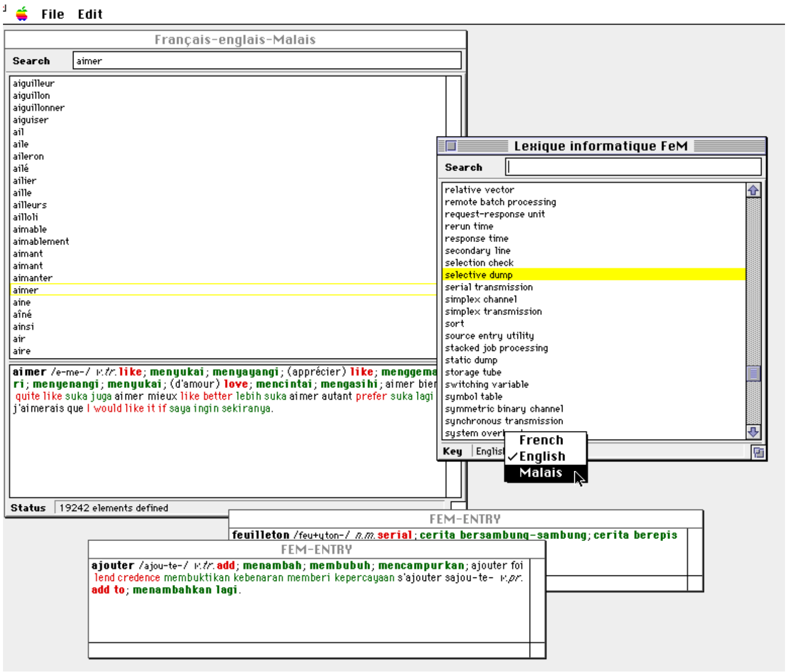
[3] Lafourcade, M. and G. Sérasset (1993) *DOP (Dictionary Object Protocol).* GETA-IMAG, Grenoble, Common Lisp Object System (MCL - CLOS), Apple Macintosh, version 2.0.

[4] Unit Terjemahan Melalui Komputer (1996) *Lexique de terminologie informatique français-anglais-malais.* 166 p.

Annexes — Figures

abandonner	entrée
/ABAN-DONE-/	prononciation
v.tr.	catégorie (verbe transitif)
(laisser)	• glose (1er sens)
to leave	équivalent anglais
meninggalkan	équivalent malais
∑È	équivalent thaï
∑Õ¥∑È	équivalent thaï
cette mère a abandonné ses enfants	phrase exemple française
this mother left her children	phrase exemple traduction anglaise
ibu ini telah meninggalkan anak-anaknya	phrase exemple traduction malai se
·¡§ππÈ¥∑Õ¥∑È≈ŸÊ¢Õß‡∏Õ	phrase exemple traduction thai
(déserter)	• glose (2eme sens)
(milit.)	étiquette
to desert	équivalent anglais
meninggalkan tugas	équivalent malais
Àπ∑À“√	équivalent thaï
(renoncer à)	• glose (3eme sens)
to give up	équivalent anglais
to abandon	équivalent anglais
melepaskan	équivalent malais
membiarkan	équivalent malais
¬°‡≈°	équivalent thaï
≈â‡≈°	équivalent thaï
≈–∑È	équivalent thaï
il a abandonné son projet	phrase exemple française
he had gave up his project	phrase exemple traduction anglaise
dia telah membiarkan projeknya	phrase exemple traduction malaise
‡¢“‰¥â≈â‡≈°‚§√ß°“√¢Õß‡¢“	phrase exemple traduction thai
(céder)	• glose (4eme sens)
to give in to	équivalent anglais
mengalah	équivalent malais
menyerah	équivalent malais
¬Õ¡	équivalent thaï
s'abandonner	• sous-entrée (forme pronominale)
/SABAN-DONE-/	prononciation
v.pr.	catégorie
s'abandonner à	locution
to give oneself up to	locution équivalente en anglais
menyerah kpd	locution équivalente en malais
¬Õ¡„Àâƒ—	locution équivalente en thai
¬Õ¡√ƒ—	locution équivalente en thai
elle s'est abandonnée au désespoir	phrase exemple française
she has given up to despair	phrase exemple traduction anglaise
dia telah menyerah kpd kekecewaan	phrase exemple traduction malaise
‡∏Õ¬Õ¡√ƒ—§““¡º¥À“ß—	phrase exemple traduction thai

Figure 3. Entry sample of the French-English-Malay-Thai dictionary.



Search: aimer

aiguilleur
aiguillon
aiguillonner
aiguiser
ail
aile
aileron
ailé
ailier
aille
ailleurs
ailloli
aimable
aimablement
aimant
aimant
aimanter
aimer
aine
aîné
ainsi
air
aire

aimer /e-me-/ *v.tr.* **like**; **menyukai**; **menyayangi**; (apprécier) **like**; **menggema ri**; **menyenangi**; **menyukai**; (d'amour) **love**; **mencintai**; **mengasihi**; aimer bien quite like suka juga aimer mieux like better lebih suka aimer autant prefer suka lagi j'aimerais que I would like it if saya ingin sekiranya.

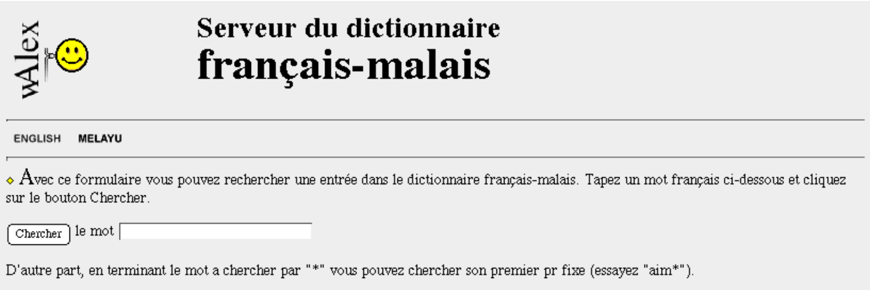
Status: 19242 elements defined

relative vector
remote batch processing
request-response unit
rerun time
response time
secondary line
selection check
selective dump
serial transmission
simplex channel
simplex transmission
sort
source entry utility
stacked job processing
static dump
storage tube
switching variable
symbol table
symmetric binary channel
synchronous transmission
system overh

feuilleton /feu+yton-/ *n.m.* **serial**; **cerita bersambung-sambung**; **cerita berepis**

ajouter /ajou-te-/ *v.tr.* **add**; **menambah**; **membubuh**; **mencampurkan**; ajouter foi lend credence membuktikan kebenaran memberi kepercayaan s'ajouter sajou-te- *v.pr.* **add to**; **menambahkan lagi**.

Figure 4. User interface of the ALEX dictionary tool. The user can look for a word and navigate through the entry list. Data cannot be modified for distributed dictionaries, but only copied as a formatted text. One can filter out some fields.



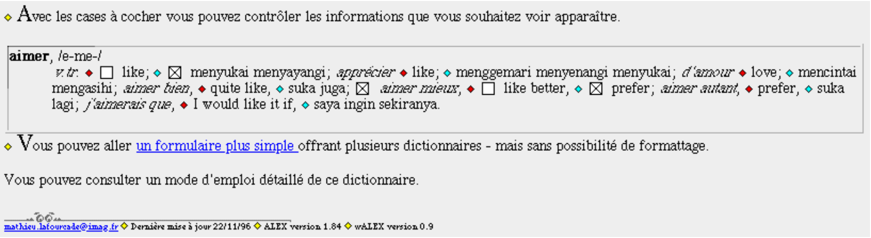
Serveur du dictionnaire
français-malais

◇ Avec ce formulaire vous pouvez rechercher une entrée dans le dictionnaire français-malais. Tapez un mot français ci-dessous et cliquez sur le bouton Chercher.

le mot

D'autre part, en terminant le mot a chercher par "*" vous pouvez chercher son premier pr fixe (essayez "aim*").

Figure 5. A specific Web form for the French-English-Malay dictionary (with French interface)



◇ Avec les cases à cocher vous pouvez contrôler les informations que vous souhaitez voir apparaître.

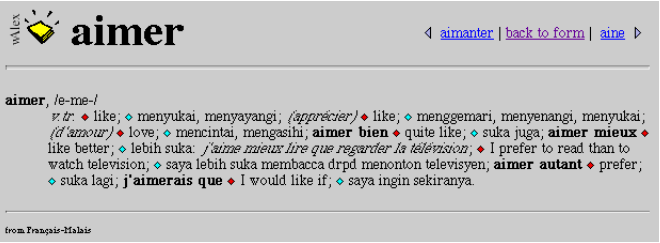
aimer, /e-me-/
v.tr. ◆ □ like; ◇ ☒ menyukai menyayangi; *apprécier* ◆ like; ◇ menggemari menyenangi menyukai; *d'amour* ◆ love; ◇ mencintai mengasihi; *aimer bien*, ◆ quite like, ◇ suka juga; ☒ *aimer mieux*, ◆ □ like better, ◇ ☒ prefer; *aimer autant*, ◆ prefer, ◇ suka lagi; *j'aimerais que*, ◆ I would like it if, ◇ saya ingin sekiranya.

◇ Vous pouvez aller un formulaire plus simple offrant plusieurs dictionnaires - mais sans possibilité de formattage.

Vous pouvez consulter un mode d'emploi détaillé de ce dictionnaire.

mathieu.lafourcade@imag.fr ◇ Dernière mise à jour 22/11/96 ◇ ALEX version 1.84 ◇ wALEX version 0.9

Figure 6. The user can select the kind of information he needs.

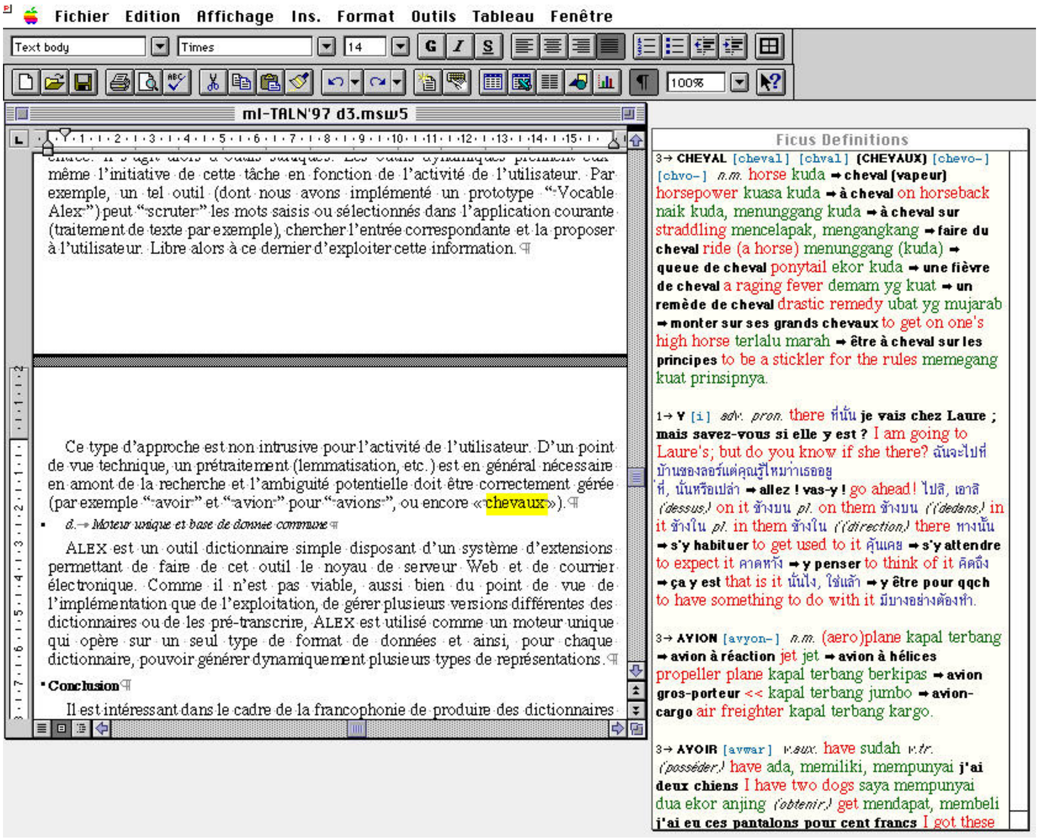


aimer

aimer, /e-me-/
v.tr. ◆ like; ◇ menyukai, menyayangi; *(apprécier)* ◆ like; ◇ menggemari, menyenangi, menyukai; *(d'amour)* ◆ love; ◇ mencintai, mengasihi; **aimer bien** ◆ quite like; ◇ suka juga; **aimer mieux** ◆ like better; ◇ lebih suka: *j'aime mieux lire que regarder la télévision*; ◆ I prefer to read than to watch television; ◇ saya lebih suka membacca drpd menonton televisyen; **aimer autant** ◆ prefer; ◇ suka lagi; **j'aimerais que** ◆ I would like if; ◇ saya ingin sekiranya.

from Français-Malais

Figure 7. A Web query answer for the French word « aimer » (like, love) with English equivalents. Here, glosses are in French, but the interface is in English.



même l'initiative de cette tâche en fonction de l'activité de l'utilisateur. Par exemple, un tel outil (dont nous avons implémenté un prototype "Vocable Alex") peut "scruter" les mots saisis ou sélectionnés dans l'application courante (traitement de texte par exemple), chercher l'entrée correspondante et la proposer à l'utilisateur. Libre alors à ce dernier d'exploiter cette information. ¶

Ce type d'approche est non intrusive pour l'activité de l'utilisateur. D'un point de vue technique, un prétraitement (lemmatisation, etc.) est en général nécessaire en amont de la recherche et l'ambiguité potentielle doit être correctement gérée (par exemple "avoir" et "avion" pour "avions", ou encore «chevaux»). ¶

- *d. → Moteur unique et base de donnée commune* ¶

ALEX est un outil dictionnaire simple disposant d'un système d'extensions permettant de faire de cet outil le noyau de serveur Web et de courrier électronique. Comme il n'est pas viable, aussi bien du point de vue de l'implémentation que de l'exploitation, de gérer plusieurs versions différentes des dictionnaires ou de les pré-transcrire, ALEX est utilisé comme un moteur unique qui opère sur un seul type de format de données et ainsi, pour chaque dictionnaire, pouvoir générer dynamiquement plusieurs types de représentations. ¶

- **Conclusion** ¶

Il est intéressant dans le cadre de la francophonie de produire des dictionnaires

Ficus Definitions

3→ **CHEVAL** [cheval] [chval] **(CHEVAUX)** [chevo-] [chvo-] *n.m.* horse kuda ➡ **cheval (vapeur)** horsepower kuasa kuda ➡ **à cheval** on horseback naik kuda, menunggang kuda ➡ **à cheval sur** straddling mencelapak, mengangkang ➡ **faire du cheval** ride (a horse) menunggang (kuda) ➡ **queue de cheval** ponytail ekor kuda ➡ **une fièvre de cheval** a raging fever demam yg kuat ➡ **un remède de cheval** drastic remedy ubat yg mujarab ➡ **monter sur ses grands chevaux** to get on one's high horse terlalu marah ➡ **être à cheval sur les principes** to be a stickler for the rules memegang kuat prinsipnya.

1→ **Y** [i] *adv. pron.* there ที่นั่น **je vais chez Laure ; mais savez-vous si elle y est ?** I am going to Laure's; but do you know if she there? ฉันจะไปที่บ้านของลอร์แต่คุณรู้ไหมว่าเธออยู่ที่, นั่นหรือเปล่า ➡ **allez ! vas-y !** go ahead! ไปสิ, เอาสิ *(dessus.)* on it ข้างบน *pl.* on them ข้างบน *((dedans.)* in it ข้างใน *pl.* in them ข้างใน *((direction.)* there ทางนั้น ➡ **s'y habituer** to get used to it คุ้นเคย ➡ **s'y attendre** to expect it คาดหวัง ➡ **y penser** to think of it คิดถึง ➡ **ça y est** that is it นั่นไง, ใช่แล้ว ➡ **y être pour qqch** to have something to do with it มีบางอย่างต้องทำ.

3→ **AVION** [avyon-] *n.m.* (aero)plane kapal terbang ➡ **avion à réaction** jet jet ➡ **avion à hélices** propeller plane kapal terbang berkipas ➡ **avion gros-porteur** << kapal terbang jumbo ➡ **avion-cargo** air freighter kapal terbang kargo.

3→ **AVOIR** [avwar] *v.aux.* have sudah *v.tr.* *(posséder.)* have ada, memiliki, mempunyai **j'ai deux chiens** I have two dogs saya mempunyai dua ekor anjing *(obtenir.)* get mendapat, membeli **j'ai eu ces pantalons pour cent francs** I got these

Figure 8. A Ficus window (on the right) window behind a common word processor. The user just select words and some dictionary lookup is automatically undertaken. The information found is then proposed to the user in a non intrusive way. Ficus is a generic dictionary service which can cooperate with various applications.

○○○○○○○○○○○○○○○○○○○○