DOCUMENT RESUME

ED 466 159 IR 021 230

AUTHOR Gamper, Johann; Knapp, Judith
TITLE New Media in the Design of a Learners' Dictionary.
PUB DATE 2001-06-00
NOTE 7p.; In: ED-MEDIA 2001 World Conference on Educational Multimedia, Hypermedia & Telecommunications. Proceedings (13th, Tampere, Finland, June 25-30, 2001); see IR 021 194.
PUB TYPE Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Computer Software Development; *Dictionaries; Educational Media; Foreign Countries; *German; *Hypermedia; Information Systems; *Italian; *Multimedia Materials; *Vocabulary; World Wide Web

ABSTRACT

This paper presents an interdisciplinary research project that aims at developing an electronic vocabulary acquisition system for the German and the Italian language called ELDIT (Elektronisches Lern(er)worterbuch Deutsch Italienisch). The approach for studying and practicing the vocabulary. To ensure maximum effectiveness of the learning process, modern psycholinguistic methods are supplied alongside with new media and technologies including adoptive hypermedia. The system is implemented as an adaptive hyperbook and runs on the World Wide Web. The paper is structured as follows. In the second section, the concept of learners' dictionaries is introduced. In section three, the ELDIT dictionary is described in detail. Section four provides an overview of how we intend to enlarge the basic dictionary to an adaptive vocabulary acquisition system. Section five discusses related work. (Contains 16 references.) (Author/AEF)

New Media in the Design of a Learners' Dictionary

Johann Gamper
Free University of Bozen
Mustergasse 4
39100 Bozen, Italy
johann.gamper@unibz.it

Judith Knapp
European Academy of Bozen
Weggensteinstr. 12/A
39100 Bozen, Italy
judith.knapp@eurac.edu

Abstract: In this paper we present an interdisciplinary research project which aims at developing an electronic vocabulary acquisition system for the German and the Italian language (ELDIT). The basic dictionary will be enlarged with text and exercise units in order to support a systematic approach for studying and practicing the vocabulary. To ensure maximum effectiveness of the learning process, modern psycholinguistic methods are applied alongside with new media and technologies including adaptive hypermedia. The system is implemented as an adaptive hyperbook and runs on the WWW.

1. Introduction

Recent research in the field of computer assisted language learning (CALL) shows that multimedia and hypermedia teachware seems to be effective and motivating for language learning (Egert 00). Such technologies open new doors for the presentation, maintenance, and dissemination of complex knowledge and information units. However, students who are less experienced in the use of information technologies may get into trouble and lose control over the system, especially if they are expected to work independently of a teacher or a tutor. Adaptive systems cope with this problem by maintaining a model of each individual user, which helps the system to prepare individually designed pages and to guide the users according to their personal learning goals.

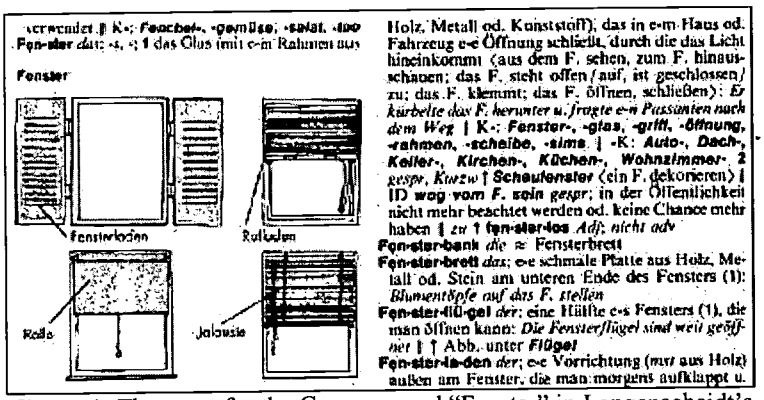
Until now, such systems have mainly been developed for teaching natural science and computer science (Nejdl et al. 99, De Bra et al. 98). We are currently investigating the use of adaptive systems in the domain of foreign language learning. In the ongoing research project ELDIT we are developing an electronic vocabulary acquisition tool for the German and the Italian language. Modern linguistic and didactic issues are considered in the design of the system. It will further contain a model of each user, adapt the content to the individual needs and preferences of each user, and guide the user through a systematic but individually shaped vocabulary acquisition process.

The paper is structured as follows: In section 2 the concept of learners' dictionaries is introduced. In section 3 the ELDIT dictionary is described in detail. Section 4 provides an overview of how we intend to enlarge the basic dictionary to an adaptive vocabulary acquisition system. In section 5 we discuss related work.

2. Learners' Dictionaries

Vocabulary acquisition is an important part of foreign language learning. To support this difficult task, lexicographers designed a special kind of dictionary, so-called *learners' dictionaries*. A learners' dictionary serves both as a reference book to decode what the learner does not understand and as an instrument which supports text production (Aarts 99). Word definitions are simpler and might be supported by pictures. Carefully selected examples are very important for the illustration of the meaning of words, as well as typical lexico-grammatical patterns, etc.

Figure 1 shows the entry of the word „Fenster" (English *window*) in Langenscheidt's learners' dictionary „Deutsch als Fremdsprache". In a printed dictionary space is limited, information is structured in a linear manner, and a lot of abbreviations are used. Therefore it is difficult to get an overview of the provided information, to understand which information is provided in detail, and where particular information can be found. Moreover, it is impossible to directly access semantically related words, except derivatives which, according to their usual lexicographic order, are physically close to the word under consideration.



Holz, Metall od. Kunststoff), das in e-m Haus od. Fahrzeug e-e Öffnung schließt, durch die das Licht hineinkommt (aus dem F. sehen, zum F. hinausschauen; das F. steht offen/auf, ist geschlossen/zu; das F. klemmt; das F. öffnen, schließen): *Er kurbelte das F. herunter u. fragte e-n Passanten nach dem Weg* | K-: **Fenster-, -glas, -griff, -öffnung, -rahmen, -scheibe, -sims** | -K: **Auto-, Dach-, Keller-, Kirchen-, Küchen-, Wohnzimmer-** 2 *gespr, Kurzw* ↑ **Schaufenster** (ein F. dekorieren) | ID **weg vom F. sein** *gespr*; in der Öffentlichkeit nicht mehr beachtet werden od. keine Chance mehr haben | *zu* ↑ **fen-ster-los** *Adj; nicht adv*

Fen-ster-bank *die* ≈ Fensterbrett

Fen-ster-brett *das*; e-e schmale Platte aus Holz, Metall od. Stein am unteren Ende des Fensters (1): *Blumentöpfe auf das F. stellen*

Fen-ster-flü-gel *der*; eine Hälfte e-s Fensters (1), die man öffnen kann: *Die Fensterflügel sind weit geöffnet* | ↑ Abb. unter **Flügel**

Fen-ster-la-den *der*; e-e Vorrichtung (nur aus Holz) außen am Fenster, die man morgens aufklappt u.

Figure 1: The entry for the German word "Fenster" in Langenscheidt's learners' dictionary "Deutsch als Fremdsprache"

Modern hypermedia and multimedia technologies allow to address a lot of these problems by providing new ways for structuring and presenting complex information. An overview of English learners' dictionaries available on CD-ROM is given in (Nesi 99). Cambridge Dictionaries Online provides web-access to a range of English learners' dictionaries (Harley 00).

3. The ELDIT Dictionary

At the European Academy Bolzano we are currently developing an electronic learners' dictionary called ELDIT (Elektronisches Lern(er)wörterbuch Deutsch ITalienisch). ELDIT can be used both as a monolingual and as a bilingual dictionary. Currently we are implementing ELDIT for the German and Italian languages, but the software could easily be adapted to different language pairs. The following sections describe the system in more detail by concentrating on different aspects of content and design.

3.1 A Sample Dictionary Entry

Figure 2 shows a screenshot of the dictionary entry for the German word "Haus" (English *house*). The information is presented in two different frames. The left-hand frame shows the lemma "Haus", a loudspeaker-button which on activation plays a sound file with the pronunciation of the word, and different meanings of the word. The right-hand frame shows various pieces of information about the word, which are collected in different tabs. Some of these tabs can only be activated once a specific meaning of the word is selected. In figure 2 the first meaning of the word "Haus" is actually selected, and the right-hand frame shows the collocation tab, which contains a list of the most frequent collocations together with their translation and an illustrative example. ELDIT contains a large number of patterns for word usage (collocations, idiomatic expressions, constructions) in combination with a translation of the patterns as well as illustrative examples. These data are presented in the tabs „Verwendu

According to Aitchinson (Aitchinson 94), people not only remember words in a possible context, but group the words in their minds into multidimensional word nets. Several word nets are visualized in ELDIT, among them the semantic field which groups words by synonymy, hyperonymy, antonymy, etc. and can be accessed by activating the tab „Verwandte Wörter".

Chapelle stresses the importance of drawing the users' attention to linguistic characteristics of a specific language (Chapelle 98). ELDIT uses footnotes to inform learners about particular differences between the two languages. Clicking on a footnote opens a small window, where linguistic difficulties are explained (see figure 2). A

summary of all difficulties related to a specific word is shown in the „N.B.“ tab. For a more comprehensive description of these features in ELDIT see (Abel 00).

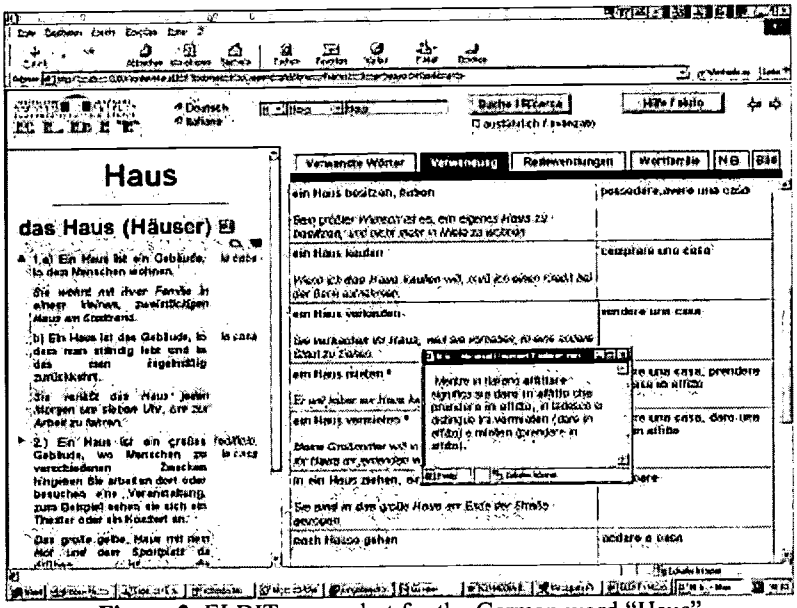


Figure 2: ELDIT screenshot for the German word "Haus"

3.2 Vocabulary Coverage

The ELDIT system should support learners of the German and the Italian language to prepare for the exam of bilingualism which is required for all employees at public institutions in South-Tyrol, a bilingual region in the North of Italy. For the selection of the vocabulary included in ELDIT the following considerations have been taken into account. The so-called basic vocabulary of a language consists of approximately 3,000 words and covers about 95% of the words in a normal text (Lewandowski 90). For both the Italian and the German languages an intersection of different basic vocabularies has been built. The two resulting vocabularies have been adjusted to each other in order to avoid big differences between the two languages. Finally, the two vocabularies have been enlarged by some frequently used words in South-Tyrol such as "farmer", „wine“, etc. The resulting vocabularies cover standard German and Italian with focus on the language variants spoken in South-Tyrol.

3.3 Searching the Dictionary

One of the most obvious advantages of electronic media over print media is their possibility to provide fast and efficient search capabilities. In ELDIT, the user can search a lemma directly or by one of its declined or conjugated forms, e.g. „ging“ leads to the lemma „gehen“. The learner can also search more complex expressions such as collocations and idiomatic expressions, e.g. „aus dem Fenster schauen“, by simply typing the whole expression into the search field. Furthermore, the user is allowed to omit some problematic parts of an expression and to replace it by wildcards, e.g. „Geb*de e*ichten“ results in the collocation „ein Gebäude errichten“. Finally, ELDIT can detect spelling errors in the search expression and provide alternative propositions.

3.4 Data Model

Figure 3 shows a simplified version of our basic data model represented in an extended entity-relationship notation with generalisation/specialisation. The main objects in our domain are word entries, which are further classified into nouns, verbs, adjectives, and structure words (articles, prepositions, etc.). Every word entry is associated with various pieces of information. A distinction is made between information concerning the word itself (mainly lexical information) and information concerning a particular meaning of the word (mainly semantic information). The former class includes the lemma (e.g. "Haus"), morphological information (a word completed with article and plural form, e.g. "das Haus, die Häuser"), derivations (e.g. "das Häuschen" and "häuslich"),

idiomatic expressions (e.g. "jemandem ins Haus platzen"), and footnotes. The latter class concerns information about a specific word meaning and includes compound words (e.g. "Hausfrau"), collocations (e.g. "in einem Haus wohnen"), a list of adjectives which typically go with this noun (e.g. "baufällig", „renovierungsbedürftig"), constructions which describe the use of a verb together with a noun or a preposition, and a picture. Words will be grouped into associative and semantically related word fields. These fields will be visualized as interactive graphs, where nodes represent the words (word entries) and edges are annotated with information explaining the difference between related words.

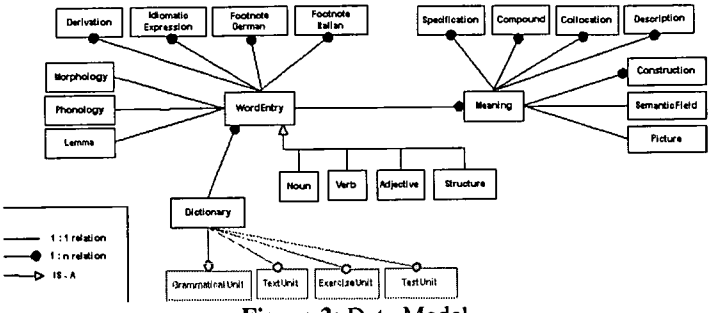


Figure 3: Data Model

Most of the above mentioned pieces of information are complex units each of which might be composed of a definition, a pattern, a translation, examples, explanations, and comments. For instance, a collocation consists of a pattern, a translation, and an example (see figure 2).

The basic data model will be extended by several new classes of information units: extensive grammar units will provide grammatical background; exercises, tests, and text material will allow the learner to practice the acquired vocabulary and to train both reading and writing skills (gray part in figure 3)

4. Towards an Adaptive Vocabulary Acquisition System

In this section we discuss two extensions of the basic ELDIT system: adaptation to individual users and systematic vocabulary acquisition. We are currently implementing these features, an evaluation remains future work.

4.1 Adaptation of Information Presentation

A great advantage of hypermedia systems is their ability to be combined with adaptation techniques. This allows to adapt the system to individual users and to provide each user with the content which best fits his/her needs and wishes. Various user features have been identified to which the content presentation of the ELDIT dictionary should be adapted.

First of all, the user's language skills and knowledge should be considered. A novice should not be faced with too many different word meanings; collocations and difficulties concerning specific words are much more helpful. An advanced learner wishes a more comprehensive picture of the language and needs more word meanings as well as a precise definition thereof, free combinations, and lots of idiomatic expressions.

The second idea of adaptation concerns language for special purposes. The current version of ELDIT covers standard German and Italian, which is sufficient in many cases. However, people in South Tyrol are more and more required to use the second language at work, which means that professional terminology becomes increasingly important. Therefore, ELDIT should support the learning of languages for special purposes, e.g. by adapting the examples to the individual user's background, occupation, business, or interests.

There are several meaningful ways to adapt the presentation of the information to user preferences. Different users might focus on different pieces of information when they access a new word. ELDIT can record each user's actions. If certain preferences are observed, the system shows the preferred information immediately, e.g. if a user almost always listens to the pronunciation of a word, the sound-file can be activated automatically every time the user accesses a dictionary entry.

Another user feature for adaptation is the possibility to use the ELDIT system both as a monolingual and as a bilingual dictionary.

4.2 Systematic Vocabulary Acquisition

We will now describe our ideas concerning the extension of the basic dictionary to a tool for systematic vocabulary acquisition. In order to retain the acquired vocabulary, „the learner needs to have opportunities to produce target language output" (Chapelle 98). ELDIT will be enlarged with text units for studying and practicing the vocabulary. The words in the text units will be linked to the corresponding dictionary entry, and the vocabulary to be trained will be highlighted, as it was successfully done in Ridder's study about reading and incidental vocabulary acquisition (De Ridder 00). Each text contains a couple of questions. The learner has to answer these questions in his/her foreign language, which serves as an indication whether he/she has understood the text and whether he/she is able to use the vocabulary.

Furthermore, all word entries in ELDIT will be categorized in various groups. Such groups are, for example, the words belonging to a particular text or a specific domain, semantically related words, or the associative field. The ELDIT system will then support the following approaches:

- *Vocabulary acquisition by word groups*: The user can study a group of words and afterwards practice the acquired knowledge with a text.
- *Vocabulary acquisition by texts*: The user can choose a text, read it, check the meaning of unknown vocabulary in this particular context, try to answer the questions, and finally systematically explore and memorize the meanings, the usage, and the context of the unknown vocabulary.
- *Repetition of vocabulary*: The user can decide to repeat the formerly acquired vocabulary and practice it with a text.

In every situation the user can ask the system to choose an appropriate vocabulary group or a text that matches best. When choosing a text the system takes into account the frequency of known and unknown words within the chosen text as well as additional data stored in the user model.

5. Related Work

A large number of CD-ROMs and web-based tools for language learning are available, but very often these tools are just one-to-one copies of printed textbooks. Only a few systems exploit the potential of modern hypermedia technologies in a meaningful way (Gamper et al. 01). Closest to our work are the vocabulary acquisition system CAVOCA (Groot 00), the dictionary Alexia (Selva et al. 97), and the adaptive language learning program CASTLE (Murphy et al. 97).

ELDIT provides more possibilities to apply, practice, and repeat the acquired vocabulary than CAVOCA. With approximately 3000 words our system contains far more dictionary entries than the monolingual system Alexia, which contains about 200 word entries. Regarding the design, both Alexia and ELDIT visualize related words as interactive graphs. ELDIT is a bilingual system and covers two languages not yet covered by the other two systems. Furthermore, the inclusion of adaptive components is planned.

CASTLE (Murphy et al. 97) is one of the few adaptive language learning programs. The traditional grammatical approach is combined with a more functional communicative approach. ELDIT mainly differs from this system in its systematic approach to vocabulary acquisition, communicative components are provided as short dialogues which are included as illustrative examples.

6. Conclusion

In this paper we presented an ongoing research project concerning the development of an adaptive vocabulary acquisition system. The system consists of a learners' dictionary which contains about 3,000 word entries for the Italian and the German languages. The design of the dictionary is motivated by modern psycholinguistic theories - such as the mental lexicon – which help the learner to remember more easily the acquired information. The basic dictionary will be extended with texts and exercises, and a systematic vocabulary acquisition process based on individual user interests and guided by the system will be provided.

New media are explored in various ways in order to support information access and presentation. Hypertext and multimedia features allow new ways for structuring, linking and presenting complex knowledge and

information. Word entries as well as the various pieces of information composing a word entry can be organized, connected, and grouped in a non-linear way, which brings different views to the world of words in a dictionary. The presentation of the information can be tailored to the individual user needs. Visual and acoustic elements support knowledge transmission. Electronic search features provide a fast and efficient access to the information, which is a very important issue in dictionaries.

Future work includes the extension of the basic dictionary with text units and grammatical units as well as a systematic and comprehensive evaluation of the system.

References

Aarts, F. (1999). Syntactic information in OALD5, LDOCE3, COBUILD2 and CIDE. In *The Perfect Learners' Dictionary*, Lexicographica: Series Maior; 95, page 16, Tübingen, Niemeyer.

Abel, A & Weber, V. (2000). ELDIT, prototype of an innovative dictionary. *In Proceedings of the 9th EURALEX International Congress on Lexicography (EURALEX'00).*

Aitchison, J. (1994). *Words in the Mind: An Introduction to the Mental Lexicon.* Blackwell Publishers Ltd, Oxford, UK.

De Bra, P. & Calvi, L. (1998). AHA: a generic adaptive hypermedia system. *In Proceedings of the 2nd Workshop on Adaptive Hypertext and Hypermedia.*

Chapelle, C. A. (1998). Multimedia CALL: Lessons to be learned from research on instructed SLA. *Language Learning & Technology*, 2(1):22–34.

Egert, C. (2000). Language learning across campuses. *Computer Assisted Language Learning*, 13(3):271–280.

Gamper, J. & Knapp, J. (2001). A review of CALL systems in foreign language instruction. *In Proceedings of the 10th International Conference on Artificial Intelligence in Education (AIED 2001)*, submitted for publication.

Groot P. J.M.(2000). *Computer assisted second language vocabulary acquisition.* Language Learning & Technology, 4(1):60–81.

Harley, A.(2000). Cambridge dictionaries online. In *Proceedings of the 9th EURALEX International Congress on Lexicography (EURALEX'00)*, see http://dictionary.cambridge.org.

Jones, C. (1999). Contextualise & personalise: Key strategies for vocabulary acquisition. *ReCALL*, 11(3):34–40.

Murphy, M. & McTear, M. (1997). Learner modeling for intelligent CALL. *In Proceedings of the 6th International Conference on User Modeling (UM97)*, pages 301–312. Springer Wien New York.

Nejdl, W. & Wolpers, M (1999). KBS hyperbook–a data-driven information system on the web. *In Proceedings of the 8th International World Wide Web Conference (WWW8),* Toronto.

Nesi, H.(1999). A user's guide to electronic dictionaries for language learners. *International Journal of Lexicography*, 12(1).

De Ridder, I.(2000). Are we conditioned to follow links? Highlights in CALL materials and their impacts on the reading process. *Computer Assisted Language Learning*, 13(2):183–195.

Selva T., Issac F., Chanier T., & Fouquer'e C. (1997). Lexical comprehension and production in the Alexia system. In *Proceedings of the Conference on Language Teaching and Language Technology (LTLT'97).*

Lewandowski, T.(1990). *Linguistisches Wörterbuch*, volume 3 of *UTB-Taschenbücher*. Quelle & Meyer, Heidelberg–Wiesbaden, 5 edition, 390–391.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

☒ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").