ELDIT – A Prototype of an Innovative Dictionary
Andrea ABEL and Vanessa WEBER, Bozen/Bolzano, Italy

Abstract
In the following presentation, ELDIT, a bilingual (German-Italian, Italian-German) electronic Learn(ers’) dictionary, will be introduced. The "basic vocabulary dictionary" ELDIT was conceived to contain about 3,000 entries and to encourage the users’ active use of vocabulary. The process of dictionary formation and elaboration is substantially supported by hypermedia technology. Electronic media facilitate a visual representation of lexical entities which is particularly oriented towards the learner. ELDIT is structured in such a way as to take into account a wide range of interrelations between the numerous lexical entities on a variety of levels (predominantly on the semantic, paradigmatic, and syntagmatic levels). The conception and elaboration are based on psycholinguistic studies which seem to corroborate the assumption that the human brain organises lexical information in form of paradigmatic, syntagmatic, "pragmatic" and tonal fields.

1 Introduction
During the last decade computer technology has experienced a rapid evolution which has involved the creation of multimedia software targeted at the public. This software includes multimedia and hypermedia-based language learning environments.

Lexicography has also profited from the advantages offered by computer technology. It is a valuable aid in the process of dictionary compilation, both for the extraction of examples and collocations from large corpora and for data processing and management. Recently, the market has started to offer a wide range of electronic dictionaries on CD-ROM or on the Web (WWW). Generally speaking, however, this kinds of dictionaries are electronic versions of what were originally print dictionaries and do not fully exploit the possibilities offered by the new medium.

In recent years, many lexicographical projects regarding the creation of original electronic dictionaries or lexicons have been set up. Some of them concern the storage of large amounts of lexical data of one or more languages for research or translation purposes. Others develop innovative electronic dictionaries, mainly intended for computer assisted learning of languages for special purposes. Often they are restricted to a limited number of entries and remain at a prototype stage.1 It has been proven, in fact, that the realisation of innovative electronic dictionaries requires enormous financial and human resources.

2 Presentation
The Elektronisches Lern(er)wörterbuch Deutsch-Italienisch/Dizionario Elettronico per Apprendenti Italiano-Tedesco (ELDIT) is a project for the development of an electronic learners’ dictionary for the German and Italian languages, consisting of about 3,000 entry words for each language. It is a bidirectional dictionary, addressed both to Italian and German students from beginners to intermediate level.
The dictionary has a mainly pedagogical scope: it aims to support Italian learners of German and German learners of Italian to memorise and store the basic vocabulary of their target language. Given the relatively small number of entries, ELDIT is supposed to help learners with the encoding rather than with the decoding abilities. Each entry word contains a lot of collocations and examples, and often usage notes so that the learner can see how it is employed, e.g. in which speech situations it normally occurs and with what kind of words it usually collocates.

ELDIT began as an electronic dictionary solely designed for computer use. Hypermedia technology [Hahn et al. 1998: 49+], i.e. the synthesis of Multimedia and Hypertext, is exploited in order to serve the pedagogical scope of our dictionary. Multimedia, in allowing the simultaneous use of different media, such as text, image, and sound, encourages a learning process that integrates different modes of presentation and is adaptable to different cognitive styles. Hypertext is used to organise the various pieces of lexical information in such a way as to help the user find them and arrange them mentally in an effective way.

The employment of new technologies offers also the advantage of motivating students, especially young ones who are usually very attracted by computer programs and material.

In the following paragraphs we will present some general considerations about the learning of vocabulary (3), which were our starting point for organising both the microstructure (4) and the macrostructure (5) of our dictionary. Then we will pass on to describe other devices we adopted in order to increase the comprehensibility and the usability of our dictionary (6) and we will conclude with presenting some further possible implementations of our project (7).

3 The Learning of Vocabulary – Some psycholinguistic Insights

Psycholinguistic studies show that our mental lexicon has a deep and complex representation of the words it contains. Associated to each word are its phonological and prosodic pattern and the whole range of lexical and semantic information related to its co-ordinates, hypernyms, hyponyms, synonyms, and antonyms. [Casiddu 1996: 51]

Although researchers have not yet clearly determined how our mental lexicon works and how it is organised, studies on the word-finding problems of aphasics and dyslexics, on word searches and slips of the tongue of normal people, as well as psycholinguistic experiments give us important clues. These studies have provided evidence on some important points:

- words are primarily organised in semantic and thematic fields,
- there are particularly strong connections between coordinates and collocational links,
- words which present phonological assonances or similar rhythmic patterns are clustered together.

The various lexical networks or fields are closely connected and form a kind of multi-dimensional word-web. Within the human word-web, each word is linked to other words or groups of words in a variety of ways and can be retrieved from different starting points. The larger and more complex the mental lexicon, the greater the layering of the semantic networks.
and the possibilities of arranging the words in multiple and diverse ways. As far as the learning of vocabulary is concerned, this means that the more words we know the easier it becomes to learn new ones [Kielhöfer 1996:11].

Foreign language learners, especially beginners, have a quantitatively poor and qualitatively badly organised vocabulary which hinders the storage of new words and the retrieval of those already known. Therefore, they need to get a lot of information on the basic words they come across in order to help them create the missing links. Thus, their lexicon, even though limited to a few thousand words will become richer and more compact. Consequently, it will be easier for the students to store and memorise words and to retrieve them quickly when needed.

4 The Microstructure – Components and Organisation

For the organisation of the lexical material of our dictionary we kept the above mentioned psycholinguistic principles in mind and tried – with the help of information technology – to offer a wide range of information for each item and to connect the various pieces of information on different levels.

Each entry is considered under three main aspects:

- semantic area,
- syntagmatic relations,
- paradigmatic relations,

and offers other supplementary information

4.1 The Explanation of the Meaning of the Words

The users get their first contact with the word through the definition. For each entry the main meanings arranged according to a pragmatic criterion are given: more general and frequent meanings are listed before specific and less frequently used ones. The definition is written in L2 as simply as possible and in a linear and easy syntax: brief sentences, few subordinate clauses, explicit rather than implicit ways of conveying information. The defining vocabulary is limited to the about 3,000 entry words plus structure words contained in the dictionary itself. In this way, we enable the learners to understand each word given in the definition and we spare them the frustrating experience of finding a definition more difficult and obscure than the word looked up. Great attention is paid to avoiding, for as long as possible, the defects of circularity – especially direct circularity – which is banal and at the same time irritating for the user. As far as the defining style is concerned, we drew our inspiration from the Cobuild style of defining [Hanks 1987; Sinclair 1991] and adapted it to our needs.

Each meaning definition is followed by an example. This has two main functions:

- it shows a typical context in which the entry word can be used and at the same time it gives a first choice of words that typically occur with it;
it completes the role of the definition by conveying supplementary semantic information that helps the user to understand the word meaning. In this case we speak of "forcing examples." [Zöfgen 1994:188]

We generally use made up examples although followers of new trends in lexicography prefer to use authentic ones. We have two main reasons for this. On the one hand, unlike the English language, German and Italian do not have at their disposal large and balanced corpora from which we can extract a great range of good authentic examples. On the other hand, we do not think that authentic examples best serve the needs of our target group, i.e. beginners and intermediate learners [Humble 1998]. By creating made up examples we can make sure that we use the words - in our case basic words - in their core meaning and set them in typical, even if simplified situations. At the same time we can control that the words we employ are not too far beyond our learners' comprehension capacity which would render the examples useless.

On the right-hand side of each meaning definition there is an extra column which contains one or more semantic equivalents in the learner's language. The role of semantic equivalents is to give the learners a quick hint about the meaning they are looking for in order to guide them in their consultation. The semantic equivalents should be considered as a help for the learners, and not as a translation in a strict sense, since we do not want to suggest that there is a 1:1 relation between the two languages.

The semantic equivalent functions also as a bridge between the two languages, because, by clicking on it, we can pass from the Italian version to the German one and vice versa.

4.2 The Syntagmatic Relations

Although each word has a potentially infinite range of possible combinations with other words, actually only a restricted number of combinations is normally used, whereas others are rare or not even admitted. Syntagmatic relations are partly ruled by grammar (verbal and nominal constructions) and partly determined by lexical combinations (collocations) or metaphorical use (idiomatic expressions). Our dictionary offers detailed information about verb construction (valency), a list of frequent word combinations and collocations, and the most frequently used idiomatic expressions.

As far as nouns are concerned, for example, we have two separate sections which deal with syntagmatic relations: One considers word combinations (1), the other idiomatic expressions (2).

1. In the file Verwendung/Combinazioni frequenti we give a list of both free combinations and restricted collocations (see Figure 1). Although we are of the opinion that learners should pay more attention to restricted collocations, since they are non-predictable and their misuse is considered a lexical mistake, we think that for our target group it is important to have a wider range of examples that show the use of the words in context. Each word combination or collocation is treated as an item and is supplied with a translation and an example. The combinations are given according to the following order: noun + verb or verb + noun, noun + (prep.) + noun, noun + adjective, where the entry word, in this case a noun, is the base. [Heid 1997]
2. In the file *Redewendungen/Espressioni idiomatiche* we include idiomatic expressions, proverbs, fixed phrases, and similes, i.e. all those kinds of expressions which present a different degree of idiomaticity, but are perceived as stable units. These items are accompanied by a translation when it is possible to find an equivalent expression in the other language and/or an explanation to make clear the meaning of opaque expressions and the origin of their figurative sense. The explanations are given in the learner’s mother tongue. At the end, an example shows a typical situation in which the idiomatic expression can occur.

In these two sections, dedicated to syntagmatic relations, the learner can find a large number of examples which are useful for illustrating different contexts in which the entry word can be used and are important tools for encoding purposes. At the same time they help the learner to build a reference frame that contains actions (verbs), objects (nouns), and features (adjectives and adverbs) belonging to typical situations, scenes of everyday life or particular themes.
4.3 The Paradigmatic Relations

In order to help the learner to organise and memorise the words, we think it is useful to build connections between words based on sense relationships.

Since it is not our purpose to go into the details of the different theories on the semantic field, we assume the notion of field in a broad sense for didactic purposes only. In other words, we do not claim to construct complex lexical systems, perfectly coherent from a theoretical point of view, nor to mirror a psychological reality.

The word field which is constructed around the entry word considers only words of the same class as the entry itself and is based on the three following types of meaning relations:

- similarity of meaning and concepts (synonyms and co-ordinates)
- contrast of meaning and concepts (gradable and nongradable antonyms)
- contiguity of concepts (converse terms) [Blank 1999: 204].

Practically, the section Verwandte Wörter/Campo semanticco is represented by computer graphics - which are automatically generated - organised on three main levels, in accordance with the following lexical hierarchy:

- The central level is the area that contains the entry word examined under one specific meaning. This entry word is both the core and the starting point of the graphics, since all other words are described in opposition to it. The central level includes synonyms, quasi-synonymous words, antonyms, and co-hyponyms, differentiated by various colours.

- Above this central area hypernyms or generic words are represented. The relationship between the upper and the central level might not always be one of perfect inclusion, since a word can be a hypernym for some semantic traits considered in the field, but not for all of them. For example, the Italian word casa, under the meaning "building where people live", has two main generics: edificio and abitazione. The word appartamento, which belongs to the same Campo semanticco, is a co-hyponym of casa if referred to the superordinate term abitazione, but it is not if referred to the term edificio.

- The lower level includes the hyponyms of the entry word only, i.e. those words which are a specific "kind of" the entry word. It is not always easy to trace a boundary between the central and the lower level of our graphics. Should we consider villa a co-hyponym of casa or a "kind of" casa? Since we do not think that students are really bothered with problems of classification, nor that this problem is crucial for their learning process, we take a rather arbitrary decision and consider villa as a co-hyponym of casa.

As it is possible to create different kinds of relationships between words, and as not all kinds of relationships appear in each word field, the structure of our graphics can both vary according to the different parts of lexicon examined and to their complexity. We also have to consider that semantic relations differ in accordance to the word class taken into consideration [Aitchinson 1997:130].

812
As far as paradigmatic relations are concerned, another important area to work on for didactic purposes is synonymy. It is not easy for language learners to discriminate between the slight stylistic, semantic, and/or contextual differences that distinguish two synonyms. Furthermore, dictionaries often neglect the explanation of such differences. ELDIT points out the distinctions between synonyms in an extra window which can be activated by clicking onto a specific button in between the two words.

The words in the graphics are hyperlinked to the respective entry words, so that the learner can easily consult their definitions. This is rarely the case with words on the lower level, which normally do not belong to our basic vocabulary. For these words, which are not examined in greater detail, a translation into L1 is available if required.

### 4.4 Other Elements

Besides information about the meaning of the word, meaning relations and syntagmatic relations discussed in the previous paragraphs, ELDIT offers also information concerning the following aspects:

- **Pronunciation:** by clicking onto the loudspeaker icon, it is possible to listen to the pronunciation of the word.

- **Grammar:** general grammatical indications (word class and inflected forms) are supplied immediately after the entry word in an implicit and intuitive form. Peculiarities or restric-
tions of usage are always pointed out in a note written in L1. For verbs we also reserve a separate section dedicated to valency.

- Word family: in the section Famiglia lessicale/ Wortfamilie we list the main derivatives and compounds of the entry word. We hereby show the learner how to build new words through affixation or compounding and which are the most productive prefixes and suffixes. The part dedicated to compounds is particularly rich for the German language.

- Remarks: in the file N.B. we list particularities, general remarks, stylistic nuances and false friends belonging to the entry word and its use. These remarks are seen both in an intralingual and a contrastive perspective and have the function of error prevention. They are written in the learner’s mother tongue.

- Images: concrete words are illustrated by images. We are also studying ways of exploiting images in a more creative and effective way, but this remains future work.

4.5 Presenting the Microstructure – The Screen Organisation

One of the advantages of electronic dictionaries is that they are not subject to space problems. They can contain much more information than paper dictionaries without making the structure and size of the dictionary too bulky and without requiring more efforts of the user for consultation, both in terms of time and complexity. Nonetheless, given that electronic dictionaries contain more lexical material, it is very important to structure the different pieces of information in a clear and effective way to help users to find their way through the several screens and windows.

Figure 3: The screen organisation: 1. The entry word and definition frame, 2. The orientation frame, 3. The additional information frame
The screen of the entry word is divided into three main frames:

1. The entry word and definition frame
   The entry word and definition frame is on the left-hand side of the screen. It contains the entry word, the general grammatical indications, the pronunciation and the meaning definition/s with the respective example/s and semantic equivalent/s. It may also include visible notes giving restrictions and particularities (for example when a word is used only in the plural form).

2. The orientation frame
   The orientation frame is on the highest part of the screen and is divided into an upper and a lower part. The upper part acts as a connection between the entry and the macrostructure, since the function Suche/cerca allows the user to go to another entry word. By clicking onto the flashes, users can retrace their way back to the beginning of their search and then move forward again. If they get lost they can ask help by clicking onto Hilfe/aiuto. The lower part displays the tabs which give access to the additional information (in the following order: wordfield, collocations/combinations, idiomatic expressions, word family, remarks, image). The titles of the tabs are not given in technical terms, but in terms which can be understood intuitively by the user.

3. The additional information frame
   The additional information frame occupies the right-hand side of the screen and appears as an empty space. This space is filled whenever users recall information by clicking onto the various tabs. Some of the tabs (idiomatic expressions, wordfamily, remarks, image) can be opened at any moment, since they contain information that do not strictly depend on a particular word meaning. The other tabs can be activated only after the users have chosen the particular word meaning they are interested in.

5 The Macrostructure

ELDIT contains about 3,000 headwords for each language which include nouns, verbs, and adjectives. Structure words will be treated separately on the basis of other criteria of presentation and description. The words are grouped according to an onomasiological/thematic order, but an alphabetical search order is also available. The thematic units group words of different word classes, which belong to the same thematic area and/or which can be referred to within the same pragmatic context. This kind of organisation is particularly useful for the development of the learners’ language production and for the widening of their active vocabulary, since it allows the learners to work and concentrate on small and structured portions of vocabulary and help them establish connections between words.

The learners who need specific information about a word and want to access to a particular entry word rapidly can consult the alphabetical order, or they can directly enter their request. This last search possibility also allows them to find collocations, compounds, and the corresponding entry word of inflected forms.
6 A Semi-bilingual Dictionary

Our type of dictionary lies between a monolingual and a bilingual dictionary. As in a monolingual dictionary, the two language systems are described from the inside, using similar criteria of description and presentation, and not as opposed to one another. Unlike bilingual dictionaries, the two languages are not presented in a symmetrical way and it is not suggested that there is a 1:1 relationship between the words and the structure of the two languages. Nevertheless, in contrast to monolingual dictionaries, a second language – corresponding to the users’ mother tongue - is employed as a resource to facilitate learners’ understanding and to stimulate their linguistic and metalinguistic reflection. In other words, ELDIT is a kind of adapted monolingual dictionary, made easier to meet the needs of those learners who are not sufficiently advanced (or who have not acquired enough skills in dictionary use) to profit from the use of a monolingual dictionary, albeit a learners’ one. In fact, although language teachers generally assume that it is better to use monolingual dictionaries [Zöfgen 1991: 2888] and encourage their students to do so, empirical studies show [e.g. Laufer 1994] that only advanced students profit from the advantages offered by monolingual dictionaries. Beginners, especially if unskilled dictionary users, benefit more from a bilingual dictionary. For this reason, bilinguilised (or semi-bilingual) dictionaries such as ELDIT are valuable instruments to enable users to pass gradually from the use of bilingual to the use of monolingual dictionaries and are particularly suitable for learners.

Since ELDIT can be seen as the sum of two semi-bilingual dictionaries, the dictionary is perfectly bidirectional, i.e. it serves German and Italian students equally. It was conceived to help German students to encode in Italian, and Italian students to encode in German.

In order to meet the needs of our target group, i.e. beginners and intermediate learners, we try to convey the different pieces of information in as transparent a way as possible. Therefore we limit the use of meta-language, symbols and abbreviations. In fact, these are the elements that make it difficult for the users to consult dictionaries and discourage the less scrupulous ones. Grammatical information must be conveyed in an easy and understandable form. For this reason we hyperlink grammar notes and the unavoidable abbreviations to specific grammar cards explained in the learners’ mother tongue.

Great attention is paid to the contrastive aspects of vocabulary and lexical use in the German and Italian languages with the purpose of preventing interference. Differences between the two language systems regarding the entry words are collected and highlighted in the N.B. file.

7 Conclusions and Further Prospects

Although we are perfectly conscious that learners do not acquire a second language vocabulary only with the help of a dictionary, we believe that ELDIT could be a valid instrument for the development and the deepening of the basic vocabulary. Due to the variety of interconnections made possible by multimedia technology, ELDIT has been structured in such a way as to help learners store and memorise vocabulary in an effective way. Moreover, it offers a wide range of examples which show how words are used in context. However, for learning purposes, ELDIT can be advantageously integrated into a more extensive learning environment. A further stage of our project provides for the combination of the dictionary with texts, grouped according to
BILINGUAL LEARNERS' DICTIONARIES

themes, and exercises. The dictionary would then become an integral part of a computer assisted language learning program.

ELDIT was conceived as an electronic learners' dictionary, but the model developed for its realisation can be adapted to different aims and implemented in a number of ways.

For example, ELDIT could be transformed from a basic dictionary to an unabridged one, by increasing the number of entries. At the same time it would be possible to develop the different sections (semantic field, idiomatic expression, word family) and create a sum of dictionary types. Our model could also be adapted to dictionaries for special purposes.

Notes

1 For the first kind of projects we cite, for example, WordNet, an on-line lexical database elaborated at the University of Princeton (http://www.cogsci.princeton.edu/wn/), the corresponding German and Italian versions, GermaNet, developed at the University of Tübingen (http://www.sfs.nphil.uni-tuebingen.de/lsd/Contents.html), and Italian WordNet developed at ITC/IRST of Trento (http://ecate.itc.it:1024/projects/wordnet), and LOGOS (http://www.logos.it/index.html); for the second kind of projects see DAFA Dictionnaire d'apprentissage du fran français des affaires developed at the Université Catholique de Louvain [Binon, Jean/Verlinde, Serge 1998], and Alexia, developed by Selva T., Chanier T. [Selva T., Chanier T. 1998]

2 For more details on the technical implementation of ELDIT see [Gamper/Knapp 2000].

3 Our notion of word field derives from Lutzeier [1995: 4+] who refers to this term as a valid instrument to cover parts of the lexicon. According to Lutzeier word fields are structured portions of the lexicon located between the single lexical item and the whole lexicon.

References


