Henning Bergenholtz y Heidi Agerbo

English and Chinese e-tools on biomechanics and sports science

1. First considerations for a concept for e-tools on biomechanics and sports science

A good story about dictionary projects describes an investigation with the finished result. However, most dictionary stories are without the desired ending; only 25% of all started dictionary projects are actually finished. In addition to this, the projects that have been finished have taken on average four times as much time as was accounted for in the original dictionary plan. A majority of these dictionary projects were planned quite poorly. In the case of the ones that were started but never finished, we can talk about lexicographic disasters. But we typically never hear the stories about planned dictionary projects that never came into any kind of realization. Such projects cannot be called disasters. On the contrary, in such cases we often see that good ideas were not transformed into practice because the project did not get the needed funding or because the lexicographers after the planning phase discovered that their plan was unrealistic for example in terms of time and available staff.

The following project description is perhaps a project of this latter kind, but we do not yet know. In November 2014, we were asked by Yifang Fan (Fujian Normal University, Fuijan, China) and Zhiyu Li (Jinan University, Guangzhou, China) to make a plan for an information tool about biomechanics and sports science. According to their proposal, this plan will be included in a large Chinese research project. If this project receives financial support, and if it will include the mentioned plan, it will result in a lexicographic project with real planning and preparation. If not, we still have some conceptual proposals for future meta-lexicographic discussions or for future concrete lexicographic projects about information tools, e-tools and/or paper tools.

The starting point in a dictionary project is the intended function(s) of the information tool(s); this means choosing the intended user type(s) and the intended user situation(s). In the case of the user type for a specialized field like biomechanics and sports science, the main difference is one between degrees of knowledge about the field. There is, of course, not a clear distinction between the three levels we apply, but in every instance of an information need, the user of an information tool should be able to make a selection of an information tool for laymen, semi-experts or experts. As shown in Bergenholtz & Kaufmann (1997), a semi-expert or an expert of molecular biology needs a very different kind of explanation than a layman. Below, this is illustrated with two different definitions of the lemma *gene*, the first one for laymen and the second one for semi-experts:

> **gene** The basic unit of inheritance transmitted from parent to offspring. An organism contains many genes – in humans approximately more than 100,000. Each gene has a specific characteristic, e.g. one out of the potential blood groups. In chemical terms genes are small sections of big complex molecules, the nucleic acids. In bacteria these are coiled aggregates and in higher organisms they are constituents of chromosomes.

This article was first printed in: *Estudios de Lexicografía*. Revista Mensual del Grupo de Las Dos Vidas de Las Palabras. Número 4, Junio de 2015: 121-136.

> **gene** A gene is a DNA sequence encoding a mRNA (protein), tRNA or rRNA. For eukaryotes a gene can also be defined as a transcribed DNA sequence or transcription unit. In prokaryotes two or more proteins are often encoded in the same transcription unit, and such a transcription unit plus its associated regulatory sequences is termed an operon.

Real experts do not need to look up a definition, but sometimes they need to be sure that they use a term correctly. The definition that experts are given could be the same as the one written for semi-experts as experts will be able to understand everything that a semi-expert understands and also since the definition written for semi-experts will be very similar to one written for experts as these two user types are able to understand almost the same. The definition for semi-experts/experts should not be longer than the one written for laymen, but this assumption is based on the definition of terms used in molecular biology.

In information tools, it is also possible to make different entries for different degrees of knowledge of the language, e.g. being a native speaker or a non-native speaker. In our case of biomechanics and sports science, we will not transfer this possible distinction into the concept of the planned information tools because the difference between native speakers and non-native speakers does not play a major role in specialized fields: A native speaker will often have the same communicative, interpretative, operative and cognitive problems as a non-native speaker based on the assumption that the non-native speaker will be provided with data in a language in which he is strong.

The function theory (see Bergenholtz & Tarp, 2003; Tarp, 2008) distinguishes between four main user situations:

A. communicative situations
1. reading a text in Lx
2. writing a text in Lx
3. proofreading in Lx
4. translating a text from Lx to Lz

B. cognitive situations
1. knowledge about a single phenomenon
2. knowledge about a complex topic
3. systematic learning
4. knowledge about new scientific tendencies
5. ...

C. operative situations
1. knowledge and/or skill in order to act in a certain situation
2. knowledge and/or skill in learning situations
3. ...

This article was first printed in: *Estudios de Lexicografía*. Revista Mensual del Grupo de Las Dos Vidas de Las Palabras. Número 4, Junio de 2015: 121-136.

D. interpretative situations
1. knowledge needed in order to understand a certain non-linguistic sign
2. ...

In situation A, a person could be reading a newspaper article about a basketball match. He comes across the word *dunk* in the sentence, "he dribbled the full length of the court and *dunked* the ball on the other end," but it is not clear to the reader what the player is actually doing in this situation. Therefore, he uses an information tool in order to find an explanation of the word that can help him understand its meaning in the given context. In situation B, a person is interested in gaining knowledge about the type of sport called basketball. This might be because he intends to start playing the game or is simply curious about this branch of sport. In either case, he needs an information tool to learn as much as possible about this branch of sport. It may be argued that this does not necessarily call for the use of an information tool as the user could simply search the Internet via a search engine such as Google. However, the Internet is filled with data that has not been structured; that a person has to spend time finding; that is not intended for a specific user type; and that has not (at least not all of it) been written by an authority who makes sure that the content is correct. Therefore, an information tool produced to be used in a cognitive situation is much more useful. In situation C, a person has just started playing basketball. He is not sure how penalty free throws are made (e.g. where to stand during the shot and how to position his feet) and therefore he looks *free throw* up in an information tool. In this situation, the user is looking for data about *how* he should actually make the throw and not the typical definition of a free throw that explains *what* it is and *why* it is awarded to a player, e.g. from the Oxford British & World English Dictionary, "An unimpeded attempt at a goal awarded to a player following a foul or other infringement". This definition is more suitable for a reception situation (though it could be argued that it is too superficial for this function). In situation D, a person is watching a basketball match and is not familiar with all the rules. The referee makes a movement to indicate a certain foul, but the person watching the match is not familiar with this signal. Therefore, he needs an information tool that explains what this signal means in basketball and why the players act the way they do according to this signal.

2. Selecting data fields for a biomechanics and sports science database

First of all, we have to mention a banal fact: A database is not a dictionary. From one and the same database, we can in fact extract and present different monofunctional or polyfunctional dictionaries, especially as e-dictionaries though every e-dictionary can also be printed. We propose a database with data fields written in English and Chinese, and from this database, we will be able to extract at least 17 information tools (dictionaries, lexicons and handbooks) for laymen, semi-experts and experts. These information tools are intended to help potential users who have communicative, cognitive, operative and interpretative needs. We plan to publish these as e-tools, but if some publishing company should wish to make printed versions of one or more of these tools, this will also be possible.

This article was first printed in: *Estudios de Lexicografía*. Revista Mensual del Grupo de Las Dos Vidas de Las Palabras. Número 4, Junio de 2015: 121-136.

These following 54 data fields are the ones that are planned to be incorporated in the database:

Field
1. Main scientific fields (roll down list with the main scientific fields)
2. Branches of sport (roll down list with the described branches of sport)
3. Lemma
4. Type of English (for the lemma)
5. Equivalent in Chinese translated from English
6. Type of Chinese (for the equivalent)
7. Meaning in English for laymen
8. Remark(s) to the meaning in English for laymen
9. Internet link(s) to the meaning in English
10. Meaning in Chinese for laymen
11. Remark(s) to the meaning in Chinese for laymen
12. Internet link(s) to the meaning for laymen in Chinese
13. Meaning in English for semi-experts/experts
14. Remarks to the meaning in English for semi-experts/experts
15. Internet link(s) to the meaning for semi-experts/experts in English
16. Meaning in Chinese for semi-experts/experts
17. Remarks to the meaning in Chinese for semi-experts/experts
18. Introduction to the scientific or sports field for laymen
19. Introduction to the scientific or sports field for semi-experts/experts
20. Link to a certain place in the layman introduction to the scientific or sport field
21. Link to a certain place in the semi-expert/expert introduction to the scientific or sport field
22. Scientific discussions for experts in English
23. Scientific discussions for experts in Chinese
24. Internet link(s) to scientific discussions (all languages)
25. Operative data for laymen written in English
26. Operative data for semi-experts/experts written in English
27. Operative data for laymen written in Chinese
28. Operative data for semi-experts/experts written in Chinese
29. Operative data for laymen as pictures or as video clips in English
30. Operative data for laymen as pictures or as video clips in Chinese
31. Operative data for semi-experts/experts as pictures or as video clips in English

This article was first printed in: *Estudios de Lexicografía*. Revista Mensual del Grupo de Las Dos Vidas de Las Palabras. Número 4, Junio de 2015: 121-136.

32. Operative data for semi-experts/experts as pictures or as video clips in Chinese
33. Internet link(s) to scientific discussions about operative actions in English
34. Internet link(s) to scientific discussions about operative actions in Chinese
35. Interpretative data for laymen and semi-experts/experts written in English
36. Interpretative data for laymen and semi-experts/experts written in Chinese
37. Interpretative data for laymen and semi-experts/experts as pictures or as video clips in English
38. Interpretative data for laymen and semi-experts/experts as pictures or as video clips in Chinese
39. Internet link(s) to scientific discussions about interpretative actions in English
40. Internet link(s) to scientific discussions about interpretative actions in Chinese
41. Grammar (for the lemma)
42. Grammar (for the equivalent)
43. Synonym(s) for the lemma
44. Remark(s) to synonym(s) for the lemma
45. Antonym(s) for the lemma
46. Synonym(s) for the equivalent
47. Remark(s) to synonym(s) for the equivalent
48. Antonym(s) for the equivalent
49. Remark(s) to antonym(s) for the equivalent
50. English collocations with the lemma
51. Translation of the English collocations to Chinese
52. English examples with the lemma
53. Translation of the English examples to Chinese
54. Memo field

Field (1) will contain a description of the main scientific fields. At the current time, we have the following fields, but our experts on human movement will probably add more:

1. Biomechanics
2. Physiology
3. Anatomy
4. ...

Field (2) is a field to which we will assign each specific branch of sport. The following list is neither systematized nor complete, but it gives an impression of the amount of different branches of sport that are to be included in the information tools:

This article was first printed in: *Estudios de Lexicografía*. Revista Mensual del Grupo de Las Dos Vidas de Las Palabras. Número 4, Junio de 2015: 121-136.

1. Badminton
2. Cricket
3. Football
4. Rugby
5. Running
6. ....

Field (4) is a field with which different types of English are made available. If a user wants to search for information connected to a specific variant of English, e.g. Australian English, he can click the relevant button:

1. British English
2. American English
3. Australian English
4. Canadian English
5. South African English
6. ....

Field (6) is a field with which different types of Chinese are made available. If a user wants to search for information connected to a specific variant of Chinese, e.g. Mandarin, he can click the relevant button:

1. Mandarin
2. Cantonese
3. ...

It is essential that all lexicographers working with the specific branches of sport in the database are experts in these. The same is, of course, the case for fields (18) and (19), which are fields that contain a relatively short introduction to each branch of sport or scientific field (badminton, rugby, biomechanics, physiology, etc.). Fields (20) and (21) make it possible to link from articles to specific sections or paragraphs in these introductions, which is helpful for readers interested in learning more about the specific topic or term. To read more about this kind of systematic description of a certain scientific field, see Bergenholtz & Nielsen (2006). Each systematic description should be considered a short handbook on a specific field. For semi-experts in molecular biology, such a description in a printed dictionary was 38 pages long (Bergenholtz et al., 1998); for laymen in a music dictionary, it was 35 pages long (Bergenholtz, 1995). However, in our potential project, it is important to have two different handbooks: one for laymen and one for semi-experts. The one for semi-experts is to some extent not comprehensible for laymen, and the one for laymen is not useful for semi-experts.

This article was first printed in: *Estudios de Lexicografía*. Revista Mensual del Grupo de Las Dos Vidas de Las Palabras. Número 4, Junio de 2015: 121-136.

Fields (22) and (23) should not contain items for dictionary articles. Instead, they should contain papers about scientific problems in one of the single specialized fields, e.g. biomechanics or rugby. Here, we would have a normal peer review process like in a scientific journal, but an important part of the papers could be short discussion papers, e.g. a reaction to one of the other papers.

Fields (8), (11), (14) and (17) will contain remarks on the meaning that are relevant for laymen (8 and 11) as well as for semi-experts/experts (14 and 17). These remarks could for example provide recommendations for a consistent use of the single terms. Sometimes, experts and some terminologists call this an attempt to normalize the definition, i.e. a prescriptive method related to all uses of a certain term in texts for laymen, semi-experts and experts. We do not apply this method nor the strict descriptive method, but instead the proscriptive method with which we recommend a certain term or a certain definition, though we also present the non-recommended terms as lemmas, in these cases telling the dictionary user that this term is not recommended and also providing arguments for this solution. For more information about the use of the proscriptive method in lexicography, see Bergenholtz & Gouws (2010).

Fields (25) and (26) will contain operative data for laymen and semi-experts respectively. At first, we only anticipated a field for operative data for semi-experts, but in the current version of the suggested database, we have included one for laymen as well. The reasoning behind this decision is that the more expert knowledge a user has, the more detailed a description for acting he needs. For example, if a person wants to learn how to shoot a basketball, a beginner (or layman) would need a simple description of the basic technique behind a shot, whereas an advanced player (or a semi-expert) already knows the basic technique and would want to learn about other techniques. In addition to this, the semi-expert is familiar with basketball jargon in the same way as in the meaning explanation of *gene* above that a semi-expert is expected to understand what DNA and RNA are. Thus, the operative data in terms of written texts should be provided in two different ways depending on whether the user is a layman or a semi-expert.

Field (35) will contain interpretative data for laymen and semi-experts in English, and in field (37) this is supplemented with pictures and video clips in English. Contrary to the operative data, there is no distinction between interpretative data for laymen and semi-experts. For example, if a basketball referee indicates a specific type of foul, the meaning of this sign will be the same whether the person who wants to understand this signal is a layman or a semi-expert. The description should be short and easy to access as the user simply needs to understand the signal in the given situation.

3. Specific e-tools extracted from the proposed database

From the proposed database based on the choice of a given information tool, e.g. reception of English texts about rugby, and based on predetermined automatic searches in the database, we can present the data in such a way that is most suitable for a certain type of user in a certain type of user

This article was first printed in: *Estudios de Lexicografía*. Revista Mensual del Grupo de Las Dos Vidas de Las Palabras. Número 4, Junio de 2015: 121-136.

situation. The same has been done in the case of the database for the Danish Internet Dictionaries (see Bergenholtz, 2015, in this journal) and the database for the Danish, English and Spanish Accounting Dictionaries (see Bergenholtz, 2012).

Our suggestion would be to produce 17 information tools for the three mentioned basic user types:

A. laymen: (1) a dictionary for reception in English, (2) one for text production in English, (3) one for translation English-Chinese, (4) one for knowledge explained in English, (5) one for knowledge explained in Chinese, (6) one for acting explained in English, (7) one for acting explained in Chinese, (8) one for interpreting signs explained in English, (9) one for interpreting signs explained in Chinese

B. semi-experts: (10) a dictionary for knowledge explained in English, (11) one for knowledge explained in Chinese, (12) one for acting explained in English, (13) one for acting explained in Chinese, (14) one for interpreting signs explained in English, (15) one for interpreting signs explained in Chinese

C. experts: (16) a homepage for scientific papers and discussions in English, (17) a homepage for scientific papers and discussions in Chinese

As the list of data fields above shows, the interpretative data is the same for laymen and semi-experts, which is why information tool (8) is the same as (14) and information tool (9) is the same as (15). The reason for the redundancy is to provide both user types with each their information tool to be used in operative situations. Thus, even though they contain the same data, they will be presented as different information tools to the users.

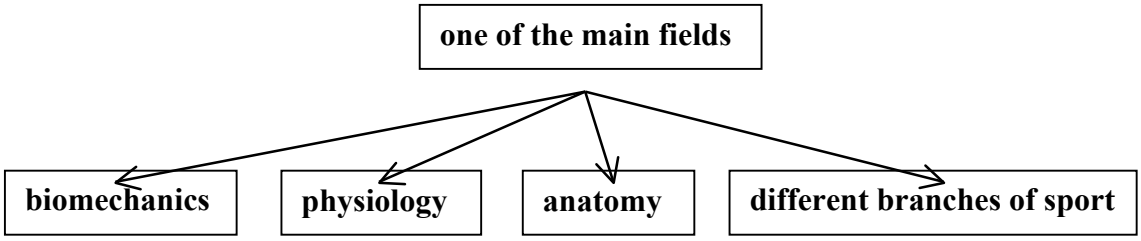
In principle, we could have as many monolingual or bilingual information tools as the number of divisions into scientific areas and branches of sport, not only the 17 information tools mentioned above, for example an English reception dictionary for laymen about biomechanics, an English reception dictionary for semi-experts about biomechanics, a Chinese information lexicon for laymen about rugby, a Chinese lexicon for semi-experts about rugby etc.

The way that these many information tools should be structured would be similar to the structuring of different electronic folders that can be opened and can contain a number of other folders or documents on a computer; thus, the user will have to make a selection of "folders" before he gets the information tool that can help him solve his problem:

As a first step, the user has to decide whether he wants to search through the whole database (all information tools) or only a certain part of the database (one of the main fields):

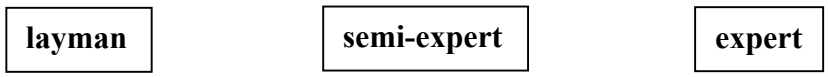
This article was first printed in: *Estudios de Lexicografía*. Revista Mensual del Grupo de Las Dos Vidas de Las Palabras. Número 4, Junio de 2015: 121-136.

If clicking **one of the main fields**, the user can choose from:

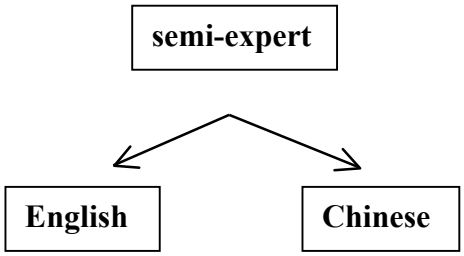


In the case of **different branches of sport,** the user will be provided with a long list containing each branch of sport. This list will be sorted alphabetically, but it will also be possible to makes searches in it so the user does not have to leaf through the hundreds of different sports.

Having made this choice, the user will have to choose a suitable user profile:

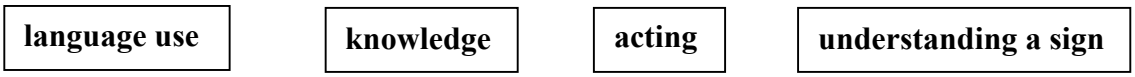


If the user chooses **semi-expert**, he identifies himself as a semi-expert, but this user profile will also be chosen by many experts as an expert may also lack knowledge in a certain subfield or wants to make sure that his understanding is similar to the data in one of the existing information tools. In addition to choosing a user type, the user will also have to choose whether the data should be provided in English or in Chinese:

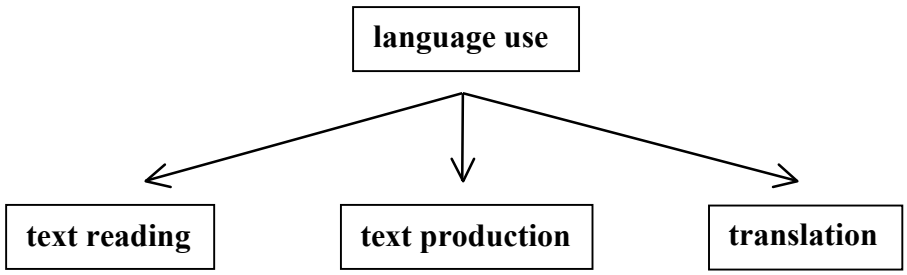


After having decided on the language, the user will have to decide in which type of situation he finds himself:

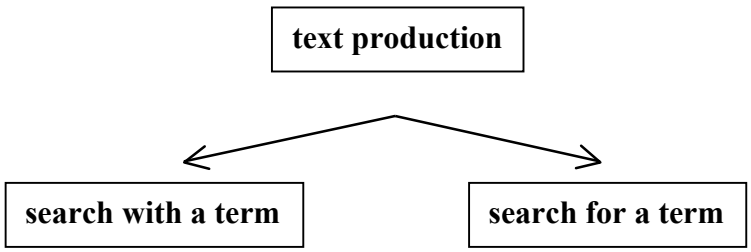
This article was first printed in: *Estudios de Lexicografía*. Revista Mensual del Grupo de Las Dos Vidas de Las Palabras. Número 4, Junio de 2015: 121-136.



These four labels correspond to the following expressions that we use in theoretical texts: communicative, cognitive, operative and interpretative. In the case of **language use**, the user is given three more specific situations from which he has to choose one:



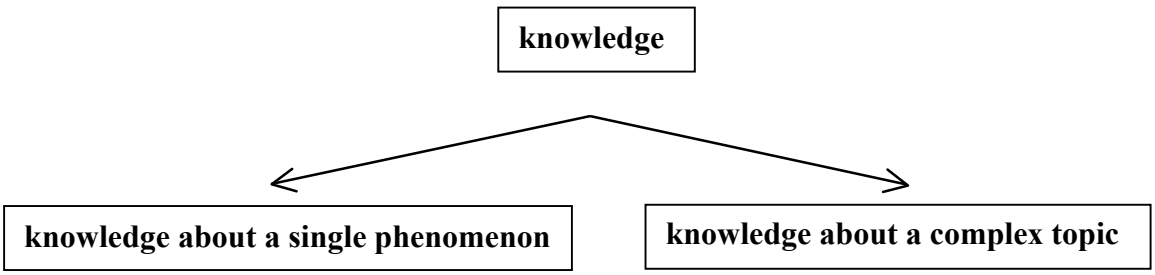
For each of these options, there will be a search field in which the user can write a search term or expression. In the case of **text reading**, the answer will be very brief, normally only in the form of a meaning explanation for semi-experts in English (had the user chosen layman as his user profile, the definition provided would, of course, have been for laymen). In the case of **text production**, there will be two options as in some situations (1) the user will already know the word but is not sure about how to use it, while in other situations (2) the user is not sure about what word to use and therefore searches for suggestions based on a Boolean search involving a combination of words that describe the meaning of the needed word. In the case of **translation**, the user can write one or more words, which will be translated into Chinese.



Of course, the user will obtain knowledge when solving a communicative problem, but this is not the primary purpose of the information search. The primary goal is to get help to solve a problem

This article was first printed in: *Estudios de Lexicografía*. Revista Mensual del Grupo de Las Dos Vidas de Las Palabras. Número 4, Junio de 2015: 121-136.

that has arisen when reading or writing a specific text. When a solution has been reached, the user can insert this solution into the specific text and then either remember or forget this piece of information. This is different from a cognitive situation (the **knowledge** button) in which case the user is looking for knowledge that may have emerged when working on a text, from a discussion or simply from a sudden need for knowledge. When this need has been met, the knowledge gained can be used in other contexts, but this was not the primary reason for the information search. We distinguish between two different types of knowledge:



For both options, the user can type in a search item in the search field or he can choose between systematic overviews of each theme (a main field or a subfield) in the folder. Especially for **knowledge about a complex topic**, it is necessary to make a search with Boolean operators available to the user.

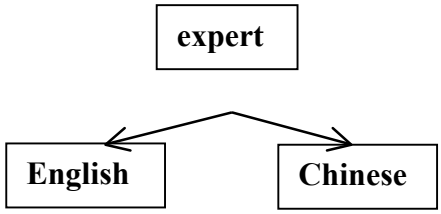
The situation called **acting** is when a person seeks information on how to act in relation to sports or other kinds of physical activity, e.g. how to measure one's pulse or how to do a certain exercise. In this situation, he partially wants to know "why", but first and foremost "how". In some cases, this kind of data can be provided in the form of a text, but it will almost always be connected to an illustration or a video clip in which the act is carried out and each step in the process is explained. If he uses the button **acting**, a search field will appear in which he can type in a single search term or make a Boolean search and thus type in a number of keywords, which will result in a small number of suggestions from which he can choose.

The same search options are given if the user chooses the button **understanding a sign** in which situation the user wants to know what a certain sign or gesture means and whether it involves a certain kind of response. One of the ways to carry out this search would be to use the sign as a search string or to do a Boolean search in which case more than one search term can be typed in the search field.

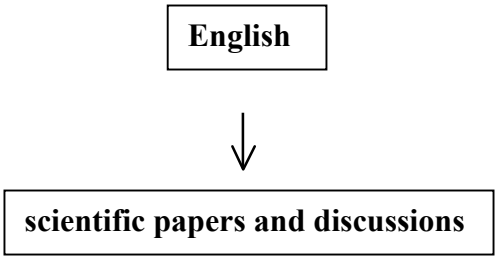
If the user instead of the button **semi-expert** chooses the button **layman**, this will result in the same options as described above, but in this case, of course, only data from the database fields targeted at laymen will be presented to the user.

If the user chooses the user type **expert**, he will be offered access to scientific papers and discussions. But first, the user will have to choose which language the data should be presented in:

This article was first printed in: *Estudios de Lexicografía*. Revista Mensual del Grupo de Las Dos Vidas de Las Palabras. Número 4, Junio de 2015: 121-136.



If he chooses the button **English**, he gets the following folder:



For each paper and discussion, there will be a maximum of 10 keywords that the user can search for in this folder either with a simple search or a Boolean search, or he can open the folder to see the list of titles of each contribution. The button **Chinese** will work in the same way.

As demonstrated, the user will need to make a number of choices in order to reach the information tool that he needs. However, we will provide a list of for example the last 8 information tools that the user has searched in and provide them in an order in which the one that he used most recently will be at the top. Each information tool will be given a number or a name so the user can easily find it again. Typically, a user will mainly be interested in only some of the available tools, e.g. ones about football and handball, and therefore it will be redundant for him every time to go through the number of selections that he had to make the first time he wanted to make a search.

The database described in this article is extremely comprehensive. We could begin by focusing on information tools only for semi-experts and only in English and then work with the database fields relevant for these tools. The result is that we could offer the users a high number of monolingual information tools. In principle, the database could be created in such a way that it encompasses more than two languages. This is what we did with the Accounting Dictionaries (Bergenholtz 2012). However, experience has shown that the programming of a database is very time consuming. Compared to the Accounting Dictionaries, the concept for our suggested database, which involves the division into three user types and four basic types of user situations, is so complicated that in the first phase of the project, the focus should be on a maximum number of two languages.

This article was first printed in: *Estudios de Lexicografía*. Revista Mensual del Grupo de Las Dos Vidas de Las Palabras. Número 4, Junio de 2015: 121-136.

4. References

BERGENHOLTZ, Henning (2012): «Concepts for monofunctional accounting dictionaries», *Terminology 18*, pp. 243-263.

BERGENHOLTZ, Henning (2015): «Historia y actualidad de los Diccionarios en Línea de Danés», (in this journal)

BERGENHOLTZ, Henning and Rufus GOUWS (2010): «A Functional Approach for the Choice between Descriptive, Prescriptive and Proscriptive Lexicography», *Lexikos 20*, pp. 26-51.

BERGENHOLTZ, Henning and Sandro NIELSEN (2006): «Subject-Field Components as Integrated Parts of LSP Dictionaries», *Terminology 12 (2)*, pp. 281-303.

BERGENHOLTZ, Henning and Sven TARP (2003): «Two opposing theories: On H.E. Wiegand's recent discovery of lexicographic function», *Hermes 31*, pp. 171-196.

BERGENHOLTZ, Henning and Uwe KAUFMANN (1997): «Terminography and lexicography. A critical survey of dictionaries from a single specialised field», *Hermes 18*, pp. 91-125.

BERGENHOLTZ, Henning and Uwe KAUFMANN in cooperation with Bjarne STUMMAN, Sven TARP, Laura DE LA ROSA MARABET, Nelson LA SERNA TORRES and Gladys LA SERNA MIRANDA (1998): *Encyclopedic Dictionary of Gene Technology. English–Spanish.* Toronto: Lugus.

BERGENHOLTZ, Inger (1995): *Politikens musikordbog.* Copenhagen: Politikens Forlag.

Oxford British & World English Dictionary. Oxford Dictionaries Online. http://oxforddictionaries.com/. Accessed 11 March 2015.

TARP, Sven (2008): «Revival of a Dusty Old Profession», *Hermes 41,* pp. 175-188.

This article was first printed in: *Estudios de Lexicografía*. Revista Mensual del Grupo de Las Dos Vidas de Las Palabras. Número 4, Junio de 2015: 121-136.