

MULTICULTURAL THESAURUS CONSTRUCTION: HOMEMAKERS IN FINNISH AND BRITISH DISCOURSES

Susanna KERÄNEN, skeranen@abo.fi

Åbo akademi. Informationsförvaltning / University of Åbo. (Turku, Finland)

Abstract

This paper reports a part of an ongoing PhD study on problems related to multilingual social science thesaurus construction in the general framework of information science. The language and culture pair studied is Finnish – British English. The emphasis is on human effort and on Finnish language and practices. Several discourses and different kinds of material are studied: thesauri, database indexing, dictionaries, surveys and social scientists. The main analysis methods used are discourse analysis and co-word analysis. In the theoretical framework, the emphasis is on communicative equivalence theories. The thematic case studied is «family roles» and in the paper an example «homemaker» is reported. The results show how co-word analysis can be used to study contextual equivalence and how important it is to aim at predictability in the information-seeking situation when constructing multilingual thesauri. Multilingual thesauri should be also multicultural and conform to the practices of users. Within social sciences, the connotative level of words can cause major obstacles in information seeking.

Keywords

thesaurus construction, multilingual thesauri, discourse analysis, linguistic equivalence, translation, co-word analysis, ARTO, CSA, COPAC

1 BACKGROUND

1.1 The project

This paper reports a part of an ongoing PhD study Content Management – Concept and Indexing Term Equivalence in a Multilingual Thesaurus. The focus of the study is the translatability of British-English social science indexing terms into Finnish language and culture on a concept, term and indexing term level to support multicultural thesaurus construction. The study is a part of a bigger project *Cultural and linguistic differences in digital storage and retrieval of information*,¹ financed by the Academy of Finland.

1. <<http://www.abo.fi/instut/diginfo/index.html>>

The study is a qualitative case study, and both quantitative and qualitative methods are used. The main data collection method is focused interview and the main analysis methods are discourse analysis and bibliometric co-word analysis. The perspectives are linguistic and sociological – a combination through which a broader understanding of the phenomena is being aimed at in the general frame of information science. The aim is to identify different discourses and vocabularies existing in a particular information domain and to see how they are considered in information storage and in multilingual and multicultural thesaurus construction.

This paper is structured as follows. Chapter 1 gives a background and introduces the theoretical framework and perspective of the study. In chapter 2 is presented research questions together with the methods and material used. In chapter 3 the case *homemaker* and the discourses are discussed. Conclusions are presented in chapter 4.

1.2 Central concepts

A *definition* is a linguistic description of the concept. It outlines the extension of the concept and its objects (NYKÄNEN 1999). «A linguistic definition is a scientific hypothesis about the concept encoded in a given word» (ROBINSON 1950, p. 41). Like other scientific hypothesis, it cannot be proved to be right, but it can be tested and proved wrong – in which case it is discarded, or revised and tested again (WIERZBICKA 1996, p. 239).

In thesauri, the scope notes, non-preferred terms, broader, narrower and related terms manifest how a concept is understood in that specific context. A definition (or an understanding of the scope of the concept) is also needed when a descriptor is translated.

Each information search in a database covers at least five different languages: the authors, the indexers, the synthetic structure, the users and the search strategy (BUCKLAND 1999), which all represent a type of *discourse*. An indexer's or a specialist's ways to express their ideas and thoughts on a certain social environment differ from each other and in indexing this can cause problems. For example, when speaking about lone mothers, a politician prefers an eloquent term (*re-miss and mother*), a journalist an eye-catching one (*single moms*) and an indexer a term accordingly to the thesaurus practices (*mothers – divorced*).

When analysing different discourses it is essential to be aware of the context where the discourses take place. In this study, *cultural context* is understood according to Nida (1975, p. 229), as «the part of the context which includes both the total culture within which a communication takes place and the specific nonlinguistic circumstances of the communication». At a general level culture can be defined to refer to the customs, beliefs, and ways of life of a group of people. Culture has several subcultures like those of indexers and social scientists, which are subsets of the main cultural group.

1.3 Theoretical framework

When considering multicultural thesaurus construction, the heart of the theoretical context lays in indexing, in information seeking, in relation between language and culture, and in equivalence.

Iivonen (1989) has showed how consistency is likely to be greater in the conceptual analysis step and less in the term selection step (*ibid*). When discussing the consistency of indexing it is also important to keep in mind that the aim of the documentary languages, e.g. thesauri, is to increase *predictability*, not consistency. Even bad indexing may be highly consistent (e.g. *ibid*, p. 12, FUGMANN 1994, p. 420-421).

ISO 5964–1985 states, that the makers of a multilingual thesaurus face two additional kinds of problem which do not occur in monolingual work:

- a) management problems, which all call for decisions on, for example, the relative status of languages (i.e. the destination of a given language as the exchange language, the dominant language, or a secondary language), or the choice of a starting point for the work (for example translation of an existing thesaurus as opposed to *ab initio* construction);
- b) language problems which call for the choice of an appropriate procedure when a term in one language expresses a concept which cannot be represented by an exactly equivalent term in one or more of the other languages. (*Ibid*, p. 5)

According to Hudon (1997) the problems traditionally associated with multilingual thesauri are:

1. that of stretching a language to make it fit a foreign conceptual structure to the point where it becomes barely recognizable to its own speakers;
2. that of transferring a whole conceptual structure from one culture to another whether it is appropriate or not;
3. that of translating literally terms from the source language into meaningless expressions in the target language, etc. (*Ibid*)

One way to create a multilingual thesaurus is translating an existing monolingual thesaurus. Hudon (1997) points out that a monolingual thesaurus is always culturally biased, and a straight translation might lead to a form of «cultural imperialism».

Also Doerr (2001) states that even though semantic heterogeneity of terminological resources has frequently been referred as a problem, a systematic analysis of its intellectual basis and structure has not been carried out. According to him, translated thesauri are thesauri, «—where each concept is optimally interpreted in words or multiple languages, to allow speakers of those languages to understand better and use concepts of this thesaurus more effectively». He also stresses that such translations are in general not established indexing terms of the target language (*Ibid*).

The art of translation already meant for Horatius and Cicero a careful interpretation of the source text and composition of source text according to the principle *non verbum de verbo, sed sensum exprime de sensu* (not word for word, but thought for thought). This means that a translator is always responsible towards the target audience (BASSNET 1995, p. 62).

Although there now seems to be a general consensus that there is no sense in demanding «mirror-translation», even today translating is not seen as an easy act. A translator is commonly seen as «a prisoner» of his/her own culture. Nida and Reyburn (NIDA 1981) have found that a translator usually understands the message in accordance with his/her own cultural-linguistic context. Usually a translator is aware of that and tries to solve the problem by using foreign terms. Regardless of that, a concept in the source language is not always semantically equivalent to the same concept in the target language. Generally descriptive phrases are therefore better (more equivalent) than foreign terms in translating a message into another culture (Ibid, p. 21-25). This can also be seen in thesaurus construction standards and guidelines where loan words are not recommended (in e.g. ISO 5964-1985; ISO 1986; SFS 5471-1988).

The early demands made upon translations do not conform to the nature of language, and that is also the main trend also seen in modern translation theories, terminology guides, as well as in thesaurus standards. The division of equivalence types in standards and guides and in translation theories the proportion of equivalence to the function of translation bases on this understanding. This study concentrates on the examination of the function of the translation, and of the equivalence type driven from focus. Therefore the divisions of translation equivalents, which base on the formal equivalence, fall outside the scope of the study. The functional, communicative translation theories are 1) Dynamic Equivalence, 2) Functional Equivalence and 3) SKOPOS theory. In the study, their influence can be seen in the perspective, hypotheses and in the research questions.

The theory of *dynamic equivalence* was developed by Nida in the 1960s and it is the first communicative translation theory. The «dynamic» refers to an idea, that a translation should cause the same *reaction* in the target audience as the source text does in the original context. Nida has further defined translation as «reproducing in the receptor language the closest natural equivalent of the message of the source language, first in terms of meaning and second in terms of style». (VEHMAS-LEHTO 1999, p. 54)

Functional equivalence theory is based on an idea that the function of the source text is the same as or similar to the function of the target text. The function of the translation adapts to the source text function (VEHMAS-LEHTO 1999, p. 70).

In *Skopos theory*, the *translation action* is stressed –translating is doing something. Vermeer (1989, p. 176) describes a translation action as a particular sort of behaviour:

for an act of behaviour to be called an action, the person performing it must (potentially) be able to explain *why* he acts as he does although he

could have acted otherwise. Furthermore, genuine reasons for actions can always be formulated in terms of aims or statements of goals (as an action «with a good reason»). (Ibid)

In Skopos theory, it is more important to fulfil the function of translation than translate in a certain style. However, the function of the target text is not necessarily the same as the source text has in its original context (REISS 1986, p. 54-59).

Vehmas-Lehto (1999, p. 12) has given a careful definition for translating: when translating one expresses something with the means of target language that has been expressed earlier with the means of source language. The definition displays that the *content* of a translation is the most important thing, the *meaning* –not *language* (Ibid).

The social sciences are connected not only to the development of science but also to the development of their surrounding culture and society. In a social sciences thesaurus, this phenomenon is seen more clearly than in, for example, thesauri of technology or medicine. Language is not static (see e.g. AITCHISON 1991; TSK, 1989; VARANTOLA 1990; WIERZBICKA 1997; LEHTONEN 2000) and therefore the language and documentation of social sciences is tied to culture and time.

According to Hiidenmaa (2000, p. 50-51) the idea of *language* merely as a tool of communication is narrow and first and foremost misleading and language should be considered more broadly: it constructs the world and guides values and choices. Suojanen (1993, p. 23) states that the world's *information culture* has commonly approved of «a neo-language» or «an euphemistic language» in public and official discourse –technology, politics, economics, and religion– especially when the context of a phenomenon feels heavy, scary, threatening. With this euphemistic language we try to guide the thinking of familiar things from a new perspective by the choice of words or a new compound term. Especially the media talks e.g. about «the poor» with words like «low-paid» (Ibid).

2 THE STUDY

The topic of the study deals with translation problems and indexing practices in creating multilingual and multicultural thesauri. The focus of this study is the translatability of British-English social science indexing terms into Finnish language and culture on a concept, term and indexing term level.

The samples studied are theoretical (not random) and therefore cases assumed to give answers to the questions of this study. The thesaurus *ELSST* (*European Language Social Science Thesaurus*) was used as a starting point and material (About ELSST see UKDA 2001; FORSMAN 2002; JÄÄSKELÄINEN 2003).

2.1 Research questions

The general research questions are:

1. How is equivalence understood in thesaurus construction guidelines and standards versus in modern translation studies versus in practice?
2. What is the pragmatic indexing term equivalence?

A theme, *family roles*, was selected as a case for the empirical part of the study. It was represented in the working version of ELSST as:

Family Roles

NT Breadwinners

Heads of Household

Homemakers

SN: A person, either male or female, whose role is to look after the home and family instead of earning money from employment.

NT Housewives (UKDA 2003)

The emphasis is on human effort and on Finnish language and culture. A background question asked is what is the sociological context of the studied concepts in the source and in the target culture. Statistical data about family role issues (division of labour, incomes, day care practices, attitudes, etc. in Finland and in UK) provide the answer to that question.

The general research questions are operationalised into several *sub-questions*:

- How are the studied concepts understood? To which extent are the differences due to institutional versus cultural differences? What is the semantic invariant?
- What are the studied terms about according to indexing and thesauri? How are the studied concepts used in indexing and why?
- How is equivalence understood? What do thesaurus constructors, indexers and social scientists aim at in their translations? Do the potential thesaurus users share same vision as thesaurus constructors?

2.2 Methods

The methods used to gather material are focused interview, word associations, simulated indexing task situation and datasets retrieved in databases, thesauri and dictionaries. The analysis methods are discourse analysis, semantic component analysis and co-word analysis.

Co-word analysis is very similar to co-citation analysis. Co-word analysis deals with co-occurrences of terms in documents, while co-citation

analysis deals with shared citations. It is thus about the relatedness of terms rather than documents (See CALLON 1991; PERSSON 1991; VON UNGERN-STERNBERG 1994; KÄRKI 1996; HORTON 1998; SCHNEIDER 2004).

The results of the co-word analysis can be shown as two-dimensional geographical *maps*. In the maps, the relative circle sizes correlates with the number of occurrences and width of lines correlates with the co-occurrence, i.e. the bigger the circles the greater the number of descriptors that occurred in the material, and the thicker the lines the greater the interaction between the linked descriptors. In co-word analysis the toolbox used is Bibexcel (BIBEXCEL 2001).

The *word association method* was also used earlier in the context of thesaurus construction. According to Nielsen (2002) word association method may result in a usable and workable thesaurus, performing as well as a thesaurus designed by traditional thesaurus construction methods. The advantage of word association method is that it may *enrich* thesaurus construction by providing current, relevant, and domain-specific information (Ibid).

Semantic component analysis is commonly used by translators and known also as denotation analysis. In component analysis the denotative meaning of the word is divided into smaller components, semantic characteristics. A component analysis is helpful especially in polysemy cases and in co-ordinated and related cases (See HÄKKINEN 1994, p. 173-175; INGO 1990; VEHMAS-LEHTO 1999; NIDA 1969; TSK 1989).

2.3 Material

The ten thesauri used to examine and compare the representation of the concept family roles are (in parenthesis possible full name and language versions available):

1. monocultural

Finnish

— Library of Parliament Subject Headings² (Finnish, Swedish and English)

— YSA³ (General Finnish Thesaurus, Finnish and Swedish)

British

— HASSET⁴ (Humanities and Social Science Electronic Thesaurus, English)

— SOSIG⁵ (Social Science Information Gateway – General Social Science Thesaurus, English)

2. <<http://www.eduskunta.fi/kirjasto/EKS/>>

3. <<http://vesa.lib.helsinki.fi/ysa/index.html>>

4. <<http://www.data-archive.ac.uk/search/hassetSearch.asp>>

5. <<http://sosig.esrc.bris.ac.uk/roads/cgi-bin/thesaurus.pl>>

American

— ERIC⁶ (The Educational Resources Information Center Thesaurus, English)

2. multicultural/international

— UNESCO Thesaurus⁷ (English, French, Spanish)

— Eurovoc⁸ (16 official languages – Spanish, Czech, Danish, German, Greek, English, French, Italian, Lithuanian, Hungarian, Dutch, Portuguese, Slovak, Slovene, Finnish and Swedish + Albania, Croatia, Latvia, Poland, Romania and Russia)

— ELSST⁹ (European Language Social Science Thesaurus, will be in English, German, French, Spanish, Finnish, Greek, Norwegian and in Danish)

— Sociological Thesaurus¹⁰ (Thesaurus of Sociological Indexing Terms, English)

— ICPSR¹¹ (The Inter-university Consortium for Political and Social Research Subject Thesaurus, English)

The databases used are international CSA (Cambridge Sociological Abstracts), British COPAC and Finnish ARTO. The studied databases are not identical –the contents, depth and ways of indexing differ substantially.

The search in CSA was limited to contain only British-English articles. Thus the topics represent the British (British i.e. here in British journals –all the articles are therefore not necessarily produced by British authors) discussion and the indexing terms (descriptors) are produced by the CSA indexers (in 2000 totally seven) with the help of a certain thesaurus, Sociological Thesaurus, and the identifiers by the authors of the articles.

COPAC is a British union online catalogue of members of the Consortium of University Research Libraries. Today, 26 institutions are contributing records to COPAC (UNIVERSITY OF MANCHESTER 2004). The indexers do not conform any established and common to all indexing guidelines nor controlled vocabularies.

ARTO is a reference database of Finnish articles (an index of articles from approx. 1000 Finnish journals). Altogether about 40 institutions have been indexing the ARTO database (HELSINKI 2005). A common indexing tool used is YSA.

In this paper, the emphasis on interview material is limited mainly to social scientists. Six British and six Finnish social scientists were interviewed (focused interview and word associations). The informants were known to be experts on matters related to family sociology and they were mostly on PhD level. Their research topics were e.g. how experts talk

6. <<http://searcheric.org/>>

7. <<http://lib-thesaurus.un.org/LIB/DHLUNBISThesaurus.nsf>>

8. <<http://europa.eu.int/celex/eurovoc/index.htm>>

9. <<http://www.fsd.uta.fi/tietoarkistolehti/english/12/elsst.html>>

10. <<http://www.csa.com/csa/HelpV6/suppl/sociothes.shtml>>

11. <<http://www.icpsr.umich.edu/IAED/THESAURUS/subject.html>>

narrower term but not any masculine related equivalent (ELSST, HASSET, SOSIG) or was used for housewives and had *mothers*, *married women* or *wives* as related terms (UNESCO and Sociological Thesaurus).

Illustrated below is the lexical network of homemakers in the six thesauri where the term occurred. It consists of 34 different terms, and the most popular co-occurring terms were housewives (with 5 hits), economic activity (3), family roles (3) and domestic responsibilities (3).

In the studied thesauri, *homemakers* is represented as a family role. Three significant clusters were found: issues related to 1) housewives, women, men, sex and responsibilities, to 2) economic activity (incl. labour and employment issues) and to 3) housework and family management. It is thus strongly connected to economic activity and to housework. It is represented as a near-equivalent to housewives and it refers to women.

What were the near-equivalents like (within the context of the studied theme, i.e. family role) in the other thesauri studied?

In *Eurovoc* a near-equivalent *housewife* was found (Finnish equivalent *kotirouva*). In ICPSR the nearest equivalents found were *husbands* and *wives* (BT *spouses* -> BT *family relationships*). *Library of Parliament Subject Headings* had no near-equivalents to refer to a person, but instead terms like *housework/domestic work* (*kotityö*), *parents* (*vanhemmat*), *family* (*perhe*).

YSA does not have equivalents in English. The nearest equivalents found were *koti-isät* (stay-at-home fathers = a father who stays at home to take care of children and the home instead of earning money from employment) and *kotiäidit* (stay-at-home mothers = a mother, who stays at home to take care of children and the home instead of earning money from employment).

Although *homemaker* can be considered as a gender-neutral term it associated as a female related term in the thesauri studied. From the gender-neutral discourse perspective the most modern thesauri studied were in this case ICPRS, *Library of Parliament Subject Headings* and YSA, all of which did not have an exact equivalent for homemaker.

Although the studied term did not have exact equivalents in the Finnish thesauri studied, it is not alien –although still foreign. Especially in the YSA, the subject field of a caregiver parent is covered, but from another perspective and thus with other kind of terms. The emphasis is more on parenthood and not in economic roles as in English thesauri studied.

— The overall discourse of thesauri shows that thesauri conform to the traditional language usage and practices. Homemaker was a foreign concept in the studied Finnish thesaurus context.

How was the term used in indexing in the dataset isolated? The descriptor/identifier-search resulted in 24 records, published during the years 1986-2001. Records having *homemakers* only in the identifier field, given by the author(s) of the indexed article, were one, and the other 23 had *homemakers* as one of the descriptors, given by an indexer. Eleven (11) documents had both identifiers and descriptors.

The descriptor-sample (articles indexed with the descriptor *homemaker*) includes 24 records and the descriptors included to the co-word analysis had two or more occurrences. The picture below illustrates the co-occurrences and the network of the most used descriptors:

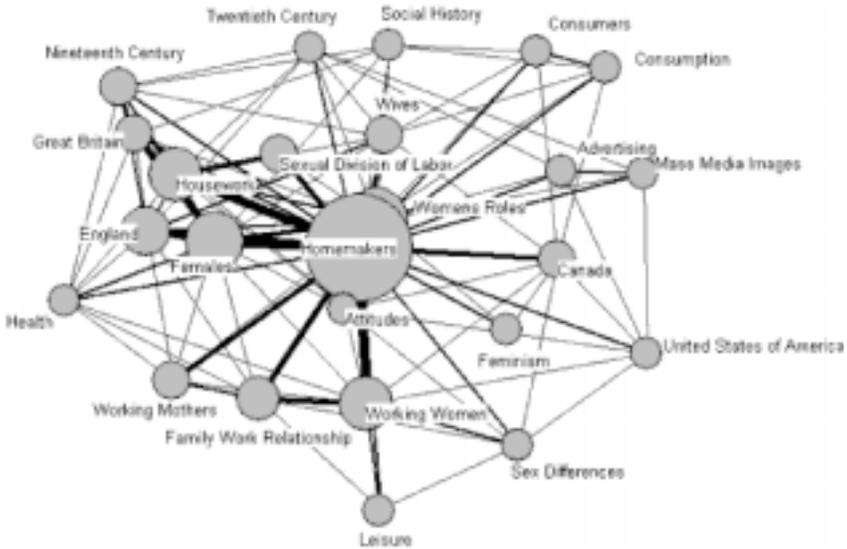


FIGURE 3. Homemakers as a descriptor in CSA

The major topic is women's roles and participation in working life and the most co-used descriptors are *women's roles*, *females*, *working women*, *housework*, *family work relationship*. The major cluster consists of *females*, *housework*, *England* and *women's roles*. It is used also in coherence with the thesaurus used in indexing and as it is commonly understood both in the other studied thesauri and in the studied dictionaries.

How British is the material retrieved? CSA is an international database. The search was limited to articles published in Britain. Since this study concentrates on Finnish and British cultures it was still interesting to know how British the authors actually were. The 45 records retrieved were analysed according to their affiliation-field in order to find out, what countries the writers of the articles represent. Two-thirds of the authors of the retrieved records were European. The biggest groups were United Kingdom (25/45) and United States of America (5/45). In CSA it was clearly more a British concept than the other terms studied.

Was homemaker used by Finns in work originally published in Finnish or in Finland? – Only once and in a different sense than it is commonly understood in CSA's English indexing (Municipal Homemaker).

- The overall discourse of CSA shows that it also conforms to the traditional language usage and practices and thesaurus discourse. *Homemaker* was understood similarly by the indexers and the authors (who used the concept as a title, abstract and /or an identifier word) as a female and a family role.

3.2.2 COPAC

In the British COPAC database, *homemakers* occurred 154 times as a subject term, that is as a descriptor (1869-2003) and 76 times as a title word (author word, 1905-2003). It was used ambiguously both by the indexers and the authors. Still the indexers were using the term in a more constant way than the authors did in titles.

How was *homemaker* used as subject term in the 1990s? It was used less ambiguously and mostly (in nine cases of twelve) to refer to stay-at-home mothers and housewives. In two cases it referred to a home help and in one case to a proper name (television program).

But how did *homemaker* occur in titles? During the 1990-2002 time period it occurred four times: It referred three times to home interior decorator and once to a stay-at-home mother. In 1980-1989, it occurred more often, that is 11 times: It was used four time to refer to a visiting nurse, three times to a stay-at-home mother, once to home help /housekeeper, once to a bride (a future homemaker) and was used twice in a compound, displaced homemakers (~divorced or widowed mothers who used to take care of the household and children and were economically dependent on their spouses).

The picture below illustrates how *homemaker* was used as a subject term. The sample covers the twelve most recent records, i.e. from 1991 to 2002:

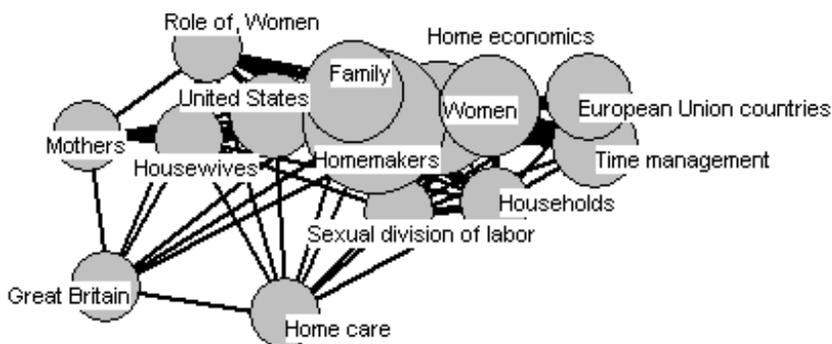


FIGURE 4. Homemakers as a subject term in COPAC

The sample had 76 individual indexing terms with 109 occurrences altogether (17 terms occurred more than once). The main topics are rather similar to the CSA database. Even here, homemakers are discussed together with women, household, family and women's roles. In COPAC, the household aspect is more strong, but the labour force participation aspect is a weaker one than in CSA.

- The overall discourse of COPAC showed clearly the ambiguity of homemaker within English language and culture. The discourse of COPAC also revealed a lack of common guidelines and vocabulary tools for indexing –there was great heterogeneity in near-equivalents, writing forms and especially in the depth of indexing.

3.2.3 ARTO

In the Finnish database, a search (13.10.04) homemaker? OR «home maker?» in any field of the reference returned only five references, published in 1984-1997. The search string appeared in titles in all the records retrieved. Three of them were written in English and two in Finnish with an abstract in English.

In the oldest cases, the term *homemaker* occurred in a compound, *communal homemakers/municipal homemakers*, and referred to an occupational title (in Finnish *kunnallinen kodinhoitaja*). In three more recent cases, the term was used to refer to *housewives/homemothers* as persons who decide about nutrition issues in families. The concept studied had no equivalents in the descriptor field.

No references were retrieved through the use of the Finnish ELSST equivalent «*kotitaloustyötä tekev?*» (descriptive phrase, ~«persons doing household work», 13.10.04). It is thus not used as a descriptor or in titles.

The Finnish near-equivalents related to *stay-at-home mothers* and *stay-at-home fathers* returned documents predominantly about childcare and parental leave.

- The studied concept homemaker was thus foreign and not commonly used in ARTO. The caregiver and homemaker role within a family is understood differently from the other samples studied.

3.3 Dictionaries

The definitions of *homemaker* given in online dictionaries are rather similar, e.g.:

- one who manages a household especially as a wife and mother (MERIAM-WEBSTER 2002).
- A housewife, esp. one in charge of the domestic arrangements (as opp. to a paid housekeeper); also, one who manages a household. So home-making n. and a. (OED ONLINE 2002).

The characteristic common to all of the studied definitions was «manages a household». The second most common characteristic was «being unpaid»: taking care of household is usually not a paid duty or work. The third most common characteristic was «gender»: four out of the six expressed in some way homemaker to be a married woman. Only one definition was totally neutral in the gender issue while the other ones referred to a female in its example of use. The fourth most common characteristic was «taking care of children».

When using dictionary translations from one language to another the results are often (in the case of family role terms) misleading –as in the case of *homemaker*. In the Finnish-English-Finnish dictionary (MOT Englanti 4.0 suomi-englanti), *homemaker* is translated as 1. *perheenemäntä* (~*homemaker/housewife*) and *kotiäiti* (*stay-at-home mother*). Its near-equivalent, *housewife*, is translated as *kotirouva*, *perheenemäntä*, *kotiäiti*. According to these equivalents the concepts, the terms represented are greatly the same, but contrary to other studied discourses, *housewife* would be broader in its extension:

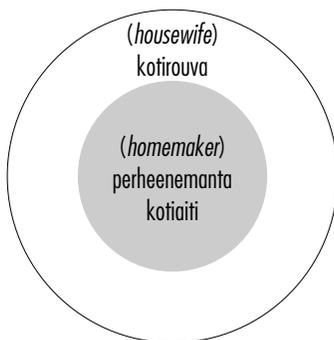


FIGURE 5. Homemakers vs. housewives in dictionary translations

- The overall discourse of dictionaries shows that *homemaker* is commonly understood as a woman, who takes care of the home and children instead of earning money from employment. In most of the studied dictionaries and thesauri, *homemaker* is represented as an unambiguous term, although it evidently is not. The discourse of the studied dictionaries represents a kind of consensus and does not reflect the different practices in real language usage.

3.4 Language usage guides and surveys

In the British guides for non-discriminatory language use for academic staff, *homemaker* is considered as a modern and gender-neutral equivalent for the term *housewife* (see e.g. UCL 2000; BRUNEL 1999; SCHOOL OF PSYCHOLOGY 2002; UNIVERSITY OF SOUTHAMPTON 1997).

Gender has traditionally been associated with words for particular roles, for example «foreman» and «housewife». Wherever possible it is best to

use gender-neutral terms such as «supervisor», «carer», «homemaker» or another suitable alternative. (BRUNEL 1999)

In the large international surveys studied, such as World Value Surveys and Eurobarometers, the term *homemaker* is not used. Instead, in the English original versions the term *housewife* is used, which is translated into Finnish as *kotiäiti* (*stay-at-home mother*).

En: Being a *housewife* is just as fulfilling as working for pay (WVS 2000(a), V116).
 Fi: *Kotiäitinä* [stay-at-home mother] oleminen on aivan yhtä antoisaa kuin ansiotyössä käyminen (WVS 2000(b), Q46D).
 En: If no paid employment:
 — *Housewife* not otherwise employed (WVS 2000(a), V229)
 Fi: — *Kotiäiti* [stay-at-home mother], ei muuta työtä (WVS 2000(b), Q96)

— The overall discourse of large international surveys shows that *homemaker* is not regarded as a usable gender-neutral term. Terms referring to mothers who stay at home taking care of children are used instead. Fathers are included in «others».

3.5 Interviews

3.5.1 British social scientists

Six British social scientist representing four different universities and five different social science disciplines were interviewed. Four out of six interviewees were used to performing their own searches in databases and all six were used to giving keywords to their articles. None of them used a thesaurus for either indexing or information seeking and they were not very familiar with thesauri designed for indexing and information retrieval.

How did British social scientists consider the concept and term *homemaker*? Firstly, they were asked to provide about five word associations (response words) for the stimulus word *homemaker* and then define the concept and evaluate its usability as a descriptor.

The following picture illustrates which associations the word *homemaker* returned (see figure 6).

Three clear clusters were found: issues related to 1) usability of word, 2) activities and 3) gender. It was commonly associated as an unusable, ambiguous and artificial word. As an activity it was usually associated positively, so the negativeness occurred mostly on a term, and not a concept, level. In practice, *homemaker* was considered to refer to a person who is female and looks after home and children. *Homemaker* was thus commonly understood to be effectively a synonym for *housewife*, but *housewife* produced even more negative associations. Neither one was deemed usable as a descriptor, and the interviewees preferred to avoid using them.

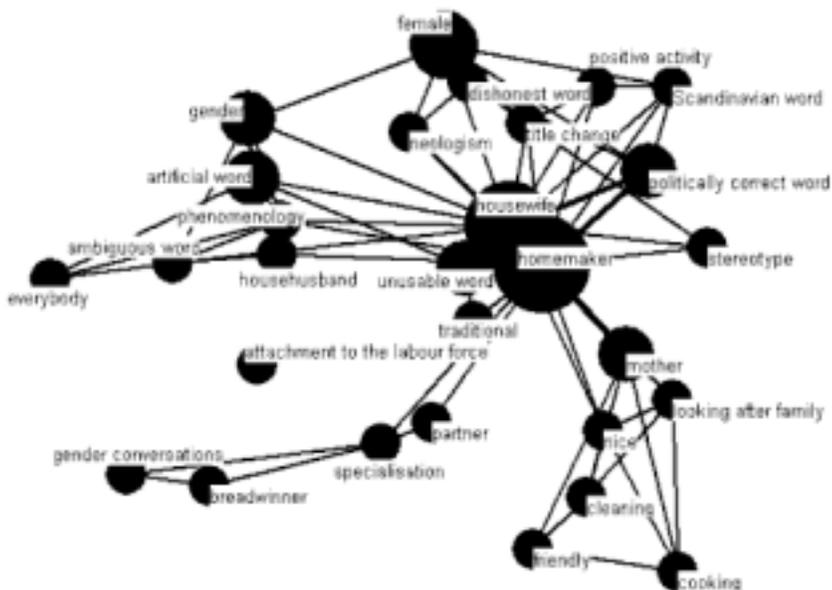


FIGURE 6. Word associations for homemaker by British social scientists

- The overall discourse of British social scientists shows that on the denotative level, homemaker is understood similarly as in dictionaries, thesauri and databases. But there is a significant difference on the connotative level, and due to that difference the discourses would not meet in (possible) information seeking situations.

3.5.2 Finnish social scientists

Six Finnish social scientist representing three different universities and four different social science disciplines were interviewed. Three out of six interviewees were used to performing their own searches in databases and all six were used to giving keywords to their articles. None of them were very familiar with thesauri designed for indexing and information retrieval.

How did the Finnish social scientists consider the concept and term *homemaker*? Firstly, they were asked to provide about five word associations (response words) for the English stimulus word *homemaker* and then define the concept, translate it into Finnish and evaluate its usability as a descriptor.

Two of the informants gave some response words in English and some in Finnish, one informant gave solely in English and three solely in Finnish.

The following picture illustrates the response words given for the stimulus word *homemaker* (Finnish words are translated into English by the author):

care of the family. The topics seen in co-word analysis of databases conform the way the term is represented in thesauri and dictionaries. The same relatedness between women, family roles and homemakers is also seen clearly in the indexing of the international and British database studied. But what is often missing is, for example, the other side of the phenomenon, namely relatedness to those who are taken care of, children, spouses, elderly people etc. In the Finnish ARTO database *homemaker* was a foreign concept –as it was also in the Finnish thesauri studied.

In dictionaries (which represent a common and «consensus-based language usage»), *homemaker* is usually represented as a woman and usually as a synonym for *housewife*. In thesauri, *homemaker* is often either a narrower or a broader term for a concept that refers to females. When translating into Finnish, aiming at gender neutrality and the «taking care of household» characteristic cause a major problem. Although Finland is considered to be one of the most egalitarian countries, and our language as gender-neutral, we do not have an equivalent to cover both genders. Moreover, the idea of having a wife or a husband at home for reasons other than parenthood, disability to work or unemployment is foreign.

According to the results obtained in the study, co-word analysis seems to be a valuable method especially when studying the contextual equivalence (question *how*). It illustrates a kind of semantic network of the studied concept in context. Other methods are needed, such as focused interview, in order to find out relevance equivalence (further question *why*). The biggest problem with co-word analysis is that it does not show you what is missing –or in other words, you need other methods, like the focused interview, to find the missing parts and help to interpret the results– to find the relevance equivalence.

Regarding discourse matters and problems in practice (thesaurus construction), problems appear on two levels: there are clear differences in Finnish versus British discourses (due to differences in societal practices) but also between social scientists versus indexers and thesaurus constructors (due to differences in practices at the content management and information seeking level).

The results of the study also show that indexers and thesaurus constructors consider a common vocabulary tool as a necessity in multicultural information seeking context, but they question the traditional forms and aims of multilingual thesauri. In the ideal case in thesaurus discourse, translation conforms to the thesaurus and user practice of the target environment. In the practical multicultural thesaurus context, the function could also be to clarify the meaning of a source language descriptor in its original cultural context and not to provide a usable equivalent in the target context. The equivalence aims are thus sometimes regarded as misleading, resulting in unusable descriptors.

Thesauri operate on the denotative level of the words, but if we want to improve predictability in the information seeking context, we should also pay more attention to the connotative level. For example, in the case of family roles it seems that ideological reasons and connotative mean-

ings form the biggest gap between specialists (as information seekers) and indexers. Connotative meanings can cause a major obstacle in the use of multicultural thesaurus within social sciences. The word association method reveals clearly the difference between the two levels.

It is therefore important to consider documentary language, i.e. also thesauri, from the perspective of natural language problems and to consider the pragmatics and norms of social science discourse. It is also important in translations to aim at predictability in information seeking context and to follow the norms of social science discourse. The true challenge is to construct multicultural thesauri for information seeking, not only multilingual ones for indexing. It is unrealistic to expect the users to study the discourse of thesauri before making their database searches. And when the thesaurus practice differs greatly from social scientists' ways of expressing the information need, there is evidently a gap in practice.

This study was supported by The Academy of Finland, The Finnish Cultural Foundation and NORSLIS (Nordic Research School in Library and Information Science).

The author warmly thanks the project leader and supervisor Sara von Ungern-Sternberg for her valuable comments during the project, the informants for their participation in the study, Ken Miller, Pam Miller, UK-DA (UK Data Archive) and FSD (Finnish Social Science Data Archive) for the research material and support.

REFERENCES

- (AITCHISON 1991) AITCHISON, Jean. *Language Change: Progress or Decay?* 2nd ed. Fontana Paperbacks: Suffolk, 1991.
- (BASSNET 1995) BASSNET, Susan. *Teoksesta toiseen. Johdatus kirjallisuuden kääntämiseen*. Ed. Riitta Oittinen. Jyväskylä: Vastapaino. [Original work: *Translation Studies* 1980 (1991), Routledge, 1995].
- (BIBEXCEL 2001) BIBEXCEL [electronic resource]: *a toolbox for bibliometricians*. Version 2001-05-21. Developed by PERSSON, Olle, Inforsk, Umeå University, Sweden. <<http://www.umu.se/inforsk/Bibexcel/index.html>>. [Cited August 2004].
- (BRUNEL 1999) BRUNEL UNIVERSITY, Human Resources. *Non-Discriminatory Language Guide* [electronic resource]: *a guide for students and staff*. <<http://www.brunel.ac.uk/admin/humanresources/equalopp/langguid.html>>. [Cited August 2004].
- (BUCKLAND 1999) BUCKLAND, Michael. «Vocabulary As A Central Concept In Library And Information Science». Preprint of paper published as «Vocabulary as a Central Concept in Library and Information Science». In: International Conference on Conceptions of Library and Information Science (3rd: 1999: Dubrovnik, Croatia). *Digital Libraries: Interdisciplinary Concepts, Challenges, and Opportunities. Proceedings of the Third International Conference on Conceptions of Library and Information Science (CoLIS3, Dubrovnik, Croatia, 23-26 May 1999)*. Ed. by ARPANAC, T., et al. (Zagreb: Lokve), p. 3-12.
- (CALLON 1991) CALLON, M.; COURTIAL, J. P.; LAVILLE, F. «Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry.» *Scientometrics*, v. 22, n.1 (1991), p. 155-205.

- (DOERR 2001) DOERR, Martin. «Semantic Problems of Thesaurus Mapping» [electronic resource]. *Journal of Digital Information*, v. 1, n. 8 (2001). <<http://jodi.ecs.soton.ac.uk/Articles/v01/i08/Doerr/>>. [Cited 26 March 2001].
- (FORSMAN 2002) FORSMAN, Maria; KERÄNEN, Susanna. «From HASSET to MADIERA [electronic resource]: FSD develops multilingual content description». *FSD neWWWs*, n. 9 (March 2002). <<http://www.fsd.uta.fi/tietoarkistolehti/english/09/madiera.html>>. [Cited 27 February 2005].
- (FUGMANN 1994) FUGMANN, Robert. «Representational Predictability: Key to the Resolution of Several Pending Issues in Indexing and Information Supply». In: INTERNATIONAL ISKO CONFERENCE (3rd: 1994: Copenhagen, Denmark). *Knowledge organization and quality management proceedings of the Third International ISKO Conference, 20-24 June 1994, Copenhagen, Denmark*. H. Albrechtsen; S. Oernager (eds.). (Frankfurt; Main: INDEKS Verlag, 1994), p. 420-421.
- (HÄKKINEN 1994) HÄKKINEN, Kaisa. *Kielitieteen perusteet*. Tampere: Suomalaisen kirjallisuuden seura, 1994.
- (HELSINKI 2005) HELSINKI UNIVERSITY LIBRARY. «Linnea Database Services [electronic resource]: ARTO». <<http://www.lib.helsinki.fi/english/libraries/linnea/ARTO.htm>>. [Cited 4 February 2005].
- (HIIDENMAA 2000) HIIDENMAA, Pirjo. «Suomen kielen eloonjäämisoppi». In: *STUDIA GENERALIA. Suomi EU:n puheenjohtajana, suomi Euroopan kielenä*. HEINO, Päivi; WESTERMARCK, Harri (eds.). (Helsinki: Helsingin yliopiston vapaan sivistystyön toimikunta, Yliopistopaino, 2000), p. 45-56.
- (HORTON 1998) HORTON, Thomas; COULTER, Neal; GRANT, Emanuel. «Applying Content Analysis to Humanities Computing Research Literature» [electronic resource]. In: ALLC/ACH JOINT INTERNATIONAL CONFERENCE (1998: Debrecen, Hungary). *Joint International Conference ALLC/ACH'98, Association for Literary and Linguistic Computing, Association for Computers and the Humanities, 5-10 July 1998, Lajos Kossuth University Debrecen, Hungary*. <<http://lingua.arts.klte.hu/allcach98/abst/abs20.htm>>. [Cited 23 April 2001].
- (HUDON 1997) HUDON, Michele. «Multilingual thesaurus construction: integrating the views of different cultures in one gateway to knowledge and concepts». *Information Services & Use*, v. 17, n. 2/3 (1997), p. 111-123.
- (IIVONEN 1989) IIVONEN, Mirja. *Indeksointituloksen riippuvuus indeksointiympäristöstä*. [The impact of the indexing environment on the result of indexing]. University of Tampere, Department of Library and Information Science, 1989. (Studies No. 26/1989).
- (INGO 1990) INGO, Rune. *Lähtökielestä kohdekieleen. Johdatusta käänntieteeseen*. Porvoo, Helsinki, Juva: WSOY, 1990.
- (ISO 2788:1986) INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO International Standard 2788. *Documentation – Guidelines for the establishment and development of monolingual thesauri*. Geneva: International Organization for Standardization, 1986.
- (ISO 5964:1985) INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO International Standard 5964. *Documentation – Guidelines for the establishment and development of multilingual thesauri*. Geneva: International Organization for Standardization, 1985.
- (JÄÄSKELÄINEN 2003) JÄÄSKELÄINEN, Taina; FORSMAN, Maria. «Finnish Version of the Multilingual Thesaurus ELSSST» [electronic resource] *FSD neWWWs*, n. 12 (March 2003). <<http://www.fsd.uta.fi/tietoarkistolehti/english/12/elsst.html>>. [Cited 27 February 2005].
- (KÄRKI 1996) KÄRKI, Riitta; KORTELAINE, Terttu. *Johdatus bibliometriikkaan*. Tampere: Informaatiotutkimuksen yhdistys, 1996.

- (LEHTONEN 2000) LEHTONEN, Mikko. *Merkitysten maailma. Kulttuurisen tekstin-tutkimuksen lähtökohtia*. 3rd ed. Tampere: Osuuskunta Vastapaino, 2000.
- (MERRIAM-WEBSTER 2002) *MERRIAM-Webster's Speaking Dictionary* [electronic resource]. <<http://www.m-w.com/>>. [Cited 12 March 2002].
- (NIDA 1969) NIDA, Eugene A.; TABER, Charles R. *The theory and practice of translation*. Leiden: E.J. Brill, 1969.
- (NIDA 1975) NIDA, Eugene A. *Componential analysis of meaning: an introduction to semantic structures*. The Hague; Paris: Mouton, 1975.
- (NIDA 1981) NIDA, Eugene; REYBURN, William. *Meaning Across Cultures*. New York: Orbis Books, 1981. (American Society of Missiology Series, n. 4).
- (NIELSEN 2002) NIELSEN, Marianne Lykke. *The word association method: a gateway to work-task based retrieval*. Doctoral dissertation. Åbo: Åbo Akademi University Press, 2002.
- (NYKÄNEN 1999) NYKÄNEN, Olli; KALLIOKUUSI, Virpi. *Sanastotyön sanastoa* [electronic resource]. [*Terminology of terminology work*]. The Finnish Terminology Centre TSK, 1999. <<http://www.tsk.fi/san/sts-s.html>>. [Cited 11 September 2002].
- (OED ONLINE 2002) *OED Online*. 2nd ed. Oxford: Oxford University Press, 1989. <<http://www.oed.com/>>. [Cited 3 September 2002].
- (PERSSON 1991) PERSSON, Olle. *Forskning i bibliometrisk belysning*. Umeå: INUM, 1991.
- (REISS 1986) REISS, Katharina; VERMEER, Hans J. *Mitä kääntäminen on*. [Grundlegung einer allgemeinen Translationstheorie. Ed. ROINILA, Pauli] Helsinki: Gaudeamus, 1986.
- (ROBINSON 1950) ROBINSON, Richard. *Definition*. Oxford: Clarendon Press, 1950. [Cited here WIERZBICKA 1996].
- (SFS 5471-1988) *Guidelines for the establishment and maintenance of Finnish language thesauri. Suomenkielisen tesauuksen laatimis- ja ylläpito-ohjeet. Suomen Standardisoimisliitto*. The Finnish Standards Association, SFS, 1988.
- (SCHNEIDER 2004) SCHNEIDER, Jesper. *Verification of bibliometric methods' applicability for thesaurus construction*. Doctoral dissertation. Department of Information Studies, Royal School of Library and Information Science, Aalborg, 2004.
- (SCHOOL OF PSYCHOLOGY 2002) School of Psychology, Queen's University, Belfast. *Guidelines for the Use of Non-sexist Language* [electronic resource, reprinted from *The Psychologist*, February 1988.] (päivitetty 22.10.2002): <<http://www.psych.qub.ac.uk/ug/guide/nonsexist.html>>. [Cited November 2004].
- (SUOJANEN 1993) SUOJANEN, Päivikki. «Identiteetti, perinne, vallan viestit». In: SUOJANEN, Päivikki; SUOJANEN, Matti K. 1993, 16: *KULTTUURIN KALEIDOSKOOPISTA. Kirjoituksia kielestä ja kulttuurista*. Tampere.
- (TSK 1989) The Finnish Terminology Centre TSK. *Sanastotyön käsikirja. Soveltavan terminologian periaatteet ja työmenetelmät* (TSK 14, SFS-käsikirja 50) Suomen Standardisoimisliitto SFS, 1989.
- (UCL 2000) University College London. *Guide to non-discriminatory language*. [electronic resource] <http://www.ucl.ac.uk/hr/docs/non_discrim_language.php>. [Cited 4 September 2002].
- (UKDA 2001) The UK Data Archive, University of Essex. (29.1.2005). *User guide to ELSST Thesaurus* [electronic resource]. <www.limber.rl.ac.uk/Internal/Deliverables/D4_2_final_V2.doc>. [Cited 29 January 2005].
- (UKDA 2003) The UK Data Archive, University of Essex. «Unpublished working version of ELSST.»
- (UNIVERSITY OF MANCHESTER 2004) UNIVERSITY OF MANCHESTER. «About COPAC» [electronic resource]. <<http://copac.ac.uk/about/>>. [Cited 9 December 2004].

- (UNIVERSITY OF SOUTHAMPTON 1997) University of Southampton, Faculty of Social Sciences. *Research Students Handbook* [electronic resource]. <<http://www.soc-stats.soton.ac.uk/teach/research-handbook.pdf>>. [Cited 4 September 2002].
- (VARANTOLA 1990) VARANTOLA, Krista. *Tekniikan suomi yhdentyvässä Euroopassa: Sanastotyön merkitystä koskeva selvitys*. Helsinki: Tekniikan Sanastokeskus ry (TSK).
- (VEHMAS-LEHTO 1999) VEHMAS-LEHTO, Inkeri. *Kopiointia vai kommunikointia: Johdatus käännteoriaan*. Helsinki: Oy Finn Lectura Ab, 1999.
- (VERMEER 1989) VERMEER, Hans, J. «Skopos and commission in translational action». In: CHESTERMAN, Andrew (ed.). *Readings in Translation Theory*, p. 173-187. Loimaa: Oy Finn Lectura Ab, 1989.
- (VON UNGERN-STERNBERG 1994) VON UNGERN-STERNBERG, Sara. *Verktyg för planering av tvärvetenskaplig informationsförsörjning: En tillämpning på ämnesområdet bioteknik i Finland*. Doctoral dissertation. Åbo: Åbo Akademis förlag, 1994. (Diss: Åbo Akademi).
- (WIERZBICKA 1996) WIERZBICKA, Anna. *Semantics Primes and Universals*. New York: Oxford University, 1996.
- (WIERZBICKA 1997) WIERZBICKA, Anna. *Understanding Cultures Through Their Key Words: English, Russian, Polish, German, and Japanese*. New York: Oxford University Press, 1997.
- (WVS 2000a) The University of Michigan, Institute for Social Research. *World Values Survey* [electronic resource]: *Questionnaire (2000-2001 fieldwork)*. <<http://wvs.isr.umich.edu/ques4.shtml>>. [Cited 28 February 2005].
- (WVS 2000b) Suomen Gallup & Kirkon tutkimuskeskus. *World Values Survey 2000* [electronic resource]: *Suomen aineisto*. Tampere: [Finnish Social Science Data Archive], 2002. <http://www.fsd.uta.fi/aineistot/quF0154_fin.pdf>. [Cited 28 February 2005].