Mapping meaning across cultures: a lexicographic resource for translators of Sanskrit Buddhist texts into English.¹

Ligeia Lugli

Mangalam Research Center for Buddhist Languages
University of California, Berkeley
ligeialugli@berkeley.edu

Abstract
This paper presents The Buddhist Translators Workbench (BTW), a digital lexicographic tool that I am currently developing in Berkeley, California. While I originally devised the proposed workbench for translators of Buddhist Sanskrit texts, this paper proffers that it can profitably be applied to any language, and could prove especially helpful to tackle translation from culturally distant languages, whose lexicon rarely matches “equivalents” in the target language.

BTW complements the lists of cognitive equivalents typical of bilingual lexicography with a system of semantic mapping based on the conceptual structure of the Historical Thesaurus of English. Visual charts allow users to compare the semantic spectrum of a word in the source language with that of its possible equivalents in the target language, as well as to contrast the semantic spectra of near-synonyms in the source language. Semantic mapping helps translators intuitively grasp the denotative range of words in the source language, and avoids superimposing on them any divergent meanings that their cognitive equivalents in target language may possess. This is particularly useful in the case of religious terminology, as cognitive equivalents often introduce the semantic baggage of the target culture’s dominant religious and philosophical traditions.

After presenting the conceptual model and practical implementation of BTW, this paper will raise the question of whether, and to what extent, the semantic categories of the Historical Thesaurus, which are based on the semantic configuration of the English language, can be used as tertium comparationis for comparing words from different languages and cultures.

Keywords: bilingual lexicography; digital lexicography; semantic mapping; Historical Thesaurus of English; cross-lingual comparison.

1. The Buddhist Translators Workbench

This paper presents the Buddhist Translators Workbench (BTW), a digital lexicographic tool that is currently being developed at the Mangalam Research Centre for Buddhist Languages in Berkeley (California). This resource was originally conceived for translators of Buddhist Sanskrit texts, but the lexicographic solutions that it introduces can profitably be applied to any language. In particular, they could prove

¹ Paper presented at the 9th International conference of AsiaLex on 26/6/2015; Words dictionaries and corpora: proceedings of the 9th International Conference of Asialex, HongKong: PolyU.
especially helpful to tackle translation from culturally distant languages, whose lexicon rarely matches “equivalents” in the target language.

BTW is presently available online as a proof-of-concept version consisting of a small repository of interlocking lexical entries (https://btw.mangalamresearch.org/). This repository is now being progressively expanded with new entries and will soon be supplemented with a suite of data-visualisation tools for intra- and inter-lingual comparison.

BTW aims to help translators craft suitable English renditions for individual Buddhist Sanskrit words; it does not tackle either phraseology or lexical items that have no specifically Buddhist applications. Its intended user-base includes the increasing number of non-native speakers of English who, for academic or religious reasons, decide to translate classical Buddhist texts into English to cater for an international readership. Like the present author, these users face the unusual challenge of translating from an ancient and culturally distant language that they do not speak (Sanskrit) into a contemporary language that they may speak fluently, but not perfectly (English). The ordeal of these translators poses an equally great lexicographic challenge for BTW, as it needs simultaneously to address the problems inherent in interpreting ancient texts, decoding a foreign language and encoding into a foreign language. To this end, BTW combines well established lexicographic practices with some innovative solutions.

2. Microstructure: overcoming the limits of equivalence

Following a widely established practice in Sanskrit bilingual lexicography, BTW structures its sense-discrimination section around cognitive equivalents. Each lemma is analysed into multiple senses, and each sense is accompanied by its English cognitive equivalent—or by a group of partial equivalents that serve to evoke its overall meaning.\footnote{In this paper I refer to the types of equivalence detailed in Adamska-Salaciak 2011, 5ff.}
For example BTW divides the Sanskrit word *prasāda* into four senses corresponding respectively to the English (1) serenity, (2) faith, confidence, trust, (3) devotion, and (4) favour. Explicating word-senses by means of equivalents has the advantage of providing users with a quick overview of the semantic spectrum of the Sanskrit lemma. This practice, however, has notorious drawbacks, including (1) inaccuracy in rendering lexical and terminological gaps, (2) blurring the difference between near-synonyms in the source language, and (3) difficulties in representing divergent polysemy (anisomorphism) between languages. To overcome these problems BTW adopts a number of lexicographic solutions of varying degrees of sophistication.

2.1 Contextual glosses and citations: dealing lexical and terminological gaps

BTW specialises in Sanskrit Buddhist words. Most of them are culture-specific and lack an equivalent in English. BTW negotiates this difficulty simply by adding explanatory equivalents or contextual glosses (Adamska-Salaciak 2011, 5–6 Duval 2008, 274-275).

One of the senses of *prasāda*, for example, refers to a specific manifestation of devotion that is not lexicalized in English. It denotes a sudden and powerful surge of reverential devotion towards religious masters accompanied by a sense of awe and typically inducing the devotee to make huge donations or personal sacrifices. A contextual gloss containing minimal encyclopaedic information regarding the religious behaviour described in Buddhist devotional literature can successfully convey the specificity of this word-sense.

The same applies to terminological gaps. One of the senses of the Sanskrit word *śraddhā*, for instance, refers to a metaphysical entity described in classical Buddhist philosophical treatises as the innate ability to develop faith. Since the classical Buddhist metaphysical discourse has never taken place in English and there is no standardised
rendition for its terminology in contemporary Western scholarship, English lacks a word to render this technical sense of śraddhā (cf. Cabré 2010, 360). Again, combining the general English equivalent ‘faith’ with an encyclopaedic gloss solves the problem. To help translators identify technical uses of Sanskrit words, BTW includes information about the main collocation in which technical senses are found.3

In addition to equivalents and glosses, for each sense of a lemma BTW provides comprehensive citations that illustrate its uses in context. Selected citations are translated into English for ease of perusal. These translations can serve as examples of functional and translational equivalents, however, more work needs to be done in this direction once extensive parallel corpora for Sanskrit Buddhist literature become available.

### 2.2 The contrastive section: comparing near-synonyms

To grasp differences between near-synonyms constitutes one of the most challenging tasks for translators of ancient texts. Equivalents in the target language rarely succeed in pinning down these differences. To facilitate the comparison between near-synonyms, BTW offers an interactive contrastive section that encourages users to explore the network of semantic and etymological relations in which the lemma participates (Figure 1). For each sense of a lemma, BTW lists all the antonyms, cognates, and ‘conceptual proximates’4 with which the lemma is contrasted in our corpus.5

---

3 Presently, ‘collocation’ in BTW refers mainly to formulaic expressions clearly recognisable by lexicographers, computational statistical analysis of collocational patterns is impossible at this stage due to the limited size of BTW’s corpus; see Hanks 2012, 403-404

4 BTW uses the locution ‘conceptual proximate’ instead of the common label ‘near-synonym’. The category of ‘conceptual proximate’ allows BTW to group together words that share a similar meaning, irrespective of the part of speech to which they belong—a solution attuned to the requirements of Sanskrit compounding syntax.

5 For an overview of the debate over the status of synonymy as a relation between word-senses see Adamska-Salaciak 2013, 331–333.
The words listed in the contrastive section are colour-coded to signal their semantic or etymological relation to the lemma, and are hyperlinked to the relevant lexicographic entries. Thus, users can easily look up related words and compare their meanings and uses in various contexts.

BTW’s contrastive section constitutes a great improvement on the bilingual Sanskrit lexicography currently available. However, systematic semantic comparison
can only be achieved by measuring semantic variation against a fixed standard, a *tertium comparationis*.

### 2.3 Semantic fields for intra- and inter-lingual comparison

Comparative linguists have long despaired of ever finding the perfect semantic *tertium comparationis*, and such an ideal tool may well never be devised, or agreed upon (Altenberg and Granger 2002, 28-29). Fortunately, for the purposes of practical lexicography merely heuristic solutions may suffice. BTW finds one such solution in the conceptual structure of the *Historical Thesaurus of English* (HTE). I will discuss the problems of using such a resource for cross-lingual comparison later in this paper; for now, I would like briefly to outline why and how the HTE could prove useful for semantic mapping.

The HTE arranges the English lexicon in over 200,000 semantic categories of increasing granularity. At the top-most level of the taxonomy one finds extremely general fields, such as *MENTAL CAPACITY*; this field subsumes categories such as *TOUGHT*, *KNOWLEDGE* and *BELIEF*, and inside these fields one finds very specific categories such as *SUPPOSITION* or *PRESUMPTION* (for a description of the HTE categorisation and its rationale see Kay 2011 and 2012). Each semantic category is associated with a unique identifier number which encodes its position in the taxonomy. Using these unique identifiers, BTW’s lexicographers assign each occurrence of a Sanskrit lemma to one or more of the HTE’s semantic categories. BTW can then generate lists of all the semantic fields attached to a lemma or to a word-sense in the entirety of its corpus, as well as in

---

6 Besides systems relying on extensive parallel corpora (which in the case of Buddhist Sanskrit literature are not yet sufficiently developed for lexicographic application), perhaps the strongest contender in the quest for a cross-linguistic *tertium comparationis* is the Natural Semantic Metalanguage (see e.g. Goddard 2001, 58). All theoretical controversy aside, this system appears to be impractical for lexicography and not easily amenable to data-visualisation, although more efforts could be made in this direction (cf. e.g. Atkins 2008, 277 n. 11 and Hanks 2013, 322–323 ).
specific texts and genres.

This facilitates both intra- and inter-lingual semantic comparison. By comparing the semantic categories assigned to near-synonyms, users can systematically contrast their meanings (Figure 2). Moreover, since the HTE uses the same categories to index English words, translators can compare the semantic categories assigned to a Sanskrit word in BTW with those indexed to its English “equivalents” in the HTE.

Manually comparing lists of semantic fields, however, is time consuming. BTW is now working towards the implementation a set of tools that will make semantic comparison much more quick and effective.

Fig. 2 Semantic fields for intra-lingual comparison in BTW

3. Dynamic features

3.1 Semantic mapping

Using the HTE’s categories, BTW will soon offer a suite of data-visualisation tools
capable of representing semantic data in a number of configurations. For instance, BTW will allow users to map the full denotational range of a Sanskrit word on the conceptual structure of the HTE. This representation has the advantage portraying the semantic spectrum of Sanskrit words without recourse to English equivalents. It avoids fragmenting the semantic continuum into a polysemy which is, at least in part, induced by the vagaries of English lexicalization. The visual representation of the semantic spectrum of prasāda (figure 3), for example, opens the possibility of conceiving the meaning of this word as a constellation of manifestations of faith that English does not lexicalize; that is, a mixture of trusting serenity, doctrinal assurance and reverential devotion. Unconstrained by the interpretive solutions imposed by English equivalents, translators will thus be able to reconceptualise the Sanskrit semantic landscape and explore new translational possibilities.
Fig. 3 HTE distribution of the semantic range of prasāda

Similar mapping systems can be employed to compare the semantic spectrum of near-synonyms. BTW will generate charts capable of contrasting the meaning of Sanskrit words as it is attested in its corpus (figure 4) or in specific texts (figure 5). Semantic mapping can also prove useful in inter-lingual comparison. Figure 6 shows the potential of contrastive charts as a tools to represent anisomorphism between Sanskrit lemmata and their possible English renditions. By visually comparing prasāda and English ‘confidence’, for example, translators will be alerted that the English word introduces an idea of courage and self-esteem that is extraneous to the Sanskrit word.
Fig. 4 Semantic comparison of prasāda and śraddhā
Fig. 5 Vikalpa-samjña contrast in the Bodhisattvabhūmi
### 3.2 Onomasiological searches and Semantic Tracker

The benefits of adopting the HTE’s conceptual structure for bilingual lexicography extend beyond data-visualization. BTW will soon be able to query the HTE database through an API and perform onomasiological searches tailored to the needs of translators. For example, users will be able to search for English words that match multiple semantic categories and approximate the polysemy of Sanskrit words.

A preliminary test of this concept has yielded promising results. The test focuses on a Sanskrit pun that is particularly difficult to translate, and compares the semantic spectrum of the words involved in the Sanskrit paronomasia with that of the word used in a felicitous English translation of the pun. Let us analyse this example in detail. The Sanskrit pun revolves around the assonance between the adjective *kalpika*, which means proper, and the verbs *kalpayati* and *vikalpayati*, which, broadly speaking, mean...
respectively ‘to conceptualise’ (especially wrongly) and ‘to discriminate’. I mapped the semantic spectra of the three Sanskrit words by assigning *kalpika* to the categories pertaining to propriety, good manners and decorum, and the verbs *kalpayati* and *vikalpayati* to the semantic fields of conceptualisation, fantasy and categorisation (figure 7). The translation of the pun, brilliantly crafted by Prof. Luis Gómez, hinges upon the English word ‘form’. As figure 8 shows, the word ‘form’ matches the semantic categories shared by *kalpika*, *kalpayati* and *vikalpayati* and thus succeeds in conveying the Sanskrit pun. This preliminary test confirms that the successful English rendition could be retrieved by searching the HTE for a word matching the specified multiple categories. Further testing of BTW’s onomasiological search is currently under way and a gamma of query options optimised for translation will be developed in the near future.
Vimalakirti 2006 p. 69

mā bhadanta śāriputro evaṃ vocaḥ | tat kasmād dhetoh | etāni hi puspāṇi kalpikāni | kiṃ kāraṇam | tathā hy etāni puspāni na kalpayantī na vikalpayantī | sthavirah punah śāriputrah kalpayati vikalpayati ca | ye bhadanta śāriputro svākhāyate dharmavinaye pravrajyāṃ kalpayantī vikalpayantī ca, te 'kalpikāh | sthāviras tu kalpayati vikalpayati ca | ye punar na kalpayantī na vikalpayantī, te kalpikāh | paśya bhadanta śāriputro eṣāṃ mahāsatvānām kāye puspāṇi na śiśyantī | yathāpi nāma sarvakaṇḍa pānā māhāsatvāṁ
dharmavinaye pravrajyām kalpikāni

Fig. 7 Semantic mapping of a Sanskrit pun
Finally, the workbench aspect of BTW will include a Semantic Tracker application. This feature is designed to allow translators to assign semantic fields to citations taken from whichever texts they are working on. Users will be able to compare meanings and track semantic variation of any desired lexical item across a corpus of their choice (virtually, in any language). Within the BTW environment, translators will thus be able to collect citations, assign semantic fields, generate infographics and share semantic data with their colleagues to negotiate agreement on interpretation and rendition.

4. Using the HTE for cross-lingual comparison: problems and (preliminary) solutions

In the preceding sections I have presented some potential applications of the conceptual structure of the HTE for semantic comparison in bilingual lexicography. It is now time
to address the theoretical problems that such practice raises.

The conceptual structure of the HTE is not language neutral. Its semantic categories are not intended to reflect universal concepts; on the contrary, they are explicitly derived from the English lexicon and are meant to represent the conceptual configuration of English (Kay 2011, 265–266). On the one hand, this may not be totally undesirable for a semantic mapping system aimed at translators. After all, translation consists in the transposition of the conceptual universe of the source language into that of the target language. Representing the semantic spectrum of Sanskrit words on the conceptual map of English could facilitate this task. On the other hand, however, interpreting Sanskrit words through the lens of English semantic categorisations may result in anglocentric distortion of their meaning.

This is a serious risk and requires careful consideration. BTW is still testing its model of semantic mapping and any definitive conclusions regarding the value of the HTE’s conceptual structure for translation-oriented bilingual lexicography would be premature. All I can offer at this stage is a description of the practical solutions that I have devised to address the difficulties that I have encountered so far, and a preliminary evaluation of them based on limited case studies.

The cross-linguistic application of the HTE’s taxonomy raises two main problems: (1) asymmetry in the representation of culture-specific categories, and (2) difference in the lexicalization and conceptualisation of reality in English and Sanskrit.

4.1 Culture-specific categories

The HTE includes categories devised to accommodate culture-specific concepts. These are not suitable for mapping the semantic universe of other cultures. For example, the field PHILOSOPHY (category 02.01.15 of the HTE) includes only categories relevant

---

7 For a discussion of the culture-specificity of thesauri’s notional structures see Fischer 2004.
to the English philosophical discourse, and these cannot safely be applied to the ancient South Asian context. Some categories, such as PRAGMATISM (02.01.15.24), or EXISTENTIALISM (02.01.15.20), simply have no counterpart in Sanskrit Buddhist texts and should obviously be ignored in cross-linguistic comparison. Others, such as SCHOLASTICISM (02.01.15.03) or METAPHYSICS (02.01.15.05), could theoretically be applied to Buddhist Sanskrit terminology. However, doing so would suggest a rather strong parallelism between Western and South Asian philosophical systems, and whether such parallelism obtains is a question better left to comparative philosophers. Hence, BTW ignores all the sub-categories that the HTE includes in the field PHILOSOPHY. Conversely, the HTE lacks specifically Buddhist philosophical categories, such as Yogācāra or Abhidharma. It is technically possible to create these categories and align them with the HTE taxonomy, but this would transcend the scope of our project. As a rule, when the HTE lacks categories suitable for representing culture-specific meanings of Sanskrit words, BTW lexicographers choose superordinate categories and assign the Sanskrit words to rather vague concepts. So, BTW indexes all instances of Buddhist philosophical terminology to the superordinate category PHILOSOPHY. As a result, BTW’s semantic mapping achieves different levels of granularity when charting the semantic spectrum of English words, whose meaning is precisely categorised in the HTE, and when charting Sanskrit lemmas, whose meaning is assigned to rather general categories in the HTE. We thus sacrifice precision in order to avoid misrepresenting cultural diversity.

4.2 Differences in conceptual representation

Culture-specific and technical terminology is not the only problem BTW faces. Several every-day phenomena are conceptualised and lexicalized differently in English and Sanskrit. The vocabulary of religious devotion, for example, is richer and more nuanced
in Sanskrit than in English. This divergence in the lexicon does not necessarily imply a cultural difference in the spiritual life of Brits and ancient Indians, but still involves problems for cross-linguistic semantic mapping. The HTE subsumes devotion under **ASPECTS OF FAITH** (03.08.01) → **PIETY** (03.08.01.23), and does not contain any distinction within **PIETY** that could capture the semantic nuances of Sanskrit words such as *prasāda* or *bhakti*. Broadly speaking, both words can mean ‘devotion’, but the former contains an element of awe and reverence, while the latter emphasises love for the object of worship. BTW bypasses the taxonomical lacuna in the HTE by assigning multiple categories to these word-senses. *Prasāda* in the sense of ‘devotion’ links to both **PIETY** (03.08.01.23) and **REVERENCE** (02.02.09.02.01); while *bhakti* links to **PIETY** (03.08.01.23) and **LOVE** (02.04.13).

It is important to note that the HTE’s categories assigned to a lemma need not contain an equivalent of the lemma. *Bhakti*, for example, is described by the intersection of the two fields **LOVE** and **PIETY**, but the HTE’s category **LOVE** does not contain any term that could, on its own, translate *bhakti*. Rather, it includes lexical items that could participate in a phraseological rendition of the Sanskrit (e.g. ‘loving devotion’). The semantic categories of the HTE, thus, can serve as semantic coordinates to chart meanings that are not lexicalised in English and elicit creative renditions in the translator’s mind.8

### 4.3 A preliminary test

Differences in the lexicalization of devotion in English and Sanskrit, one may argue, are relatively minor. They do not reflect a cultural fracture in the conceptualisation of this notion. Would simple recourse to multiple semantic categories taken from the HTE

---

8 The possibility of using semantic irrespective of the lexical items constitutes a great advantage of using a Thesaurus with an explicit conceptual structure, like the HTE, over employing a WordNet-type resource (See Sylvester 2004, cf. Teuber 2002, 197ff).
prove successful in representing two radically different conceptualisations of a phenomenon? At this stage of development, BTW’s database is too small to answer this compelling question. Hence, I suggest we run a preliminary test based on a case study external to BTW.

A very fine specimen for this purpose can be found in Goddard and Wierzbicka’s discussion of visual semantics in Walpiri (Goddard and Wierzbicka 2014, 85–101). I do not know Walpiri and I have no means to evaluate the accuracy of Goddard and Wierzbicka’s lexical analysis. I wish to focus on this study because it contains an attentive critique of anglocentric distortion in cross-linguistic comparison, and, for this very reason, constitutes the ideal test bench for the application of the HTE’s taxonomy to bilingual semantic mapping.

Goddard and Wierzbicka observe that while English relies heavily on the concept of ‘colour’ for describing the visual world, languages such as Walpiri do not. Hence, interpreting Walpiri words such as pirarr-pirarrarrapa, liirlpari, ratarata or kuruwarri-kuruwarri through the English concept colour is misleading. According to Goddard and Wierzbicka, these words indicate non-chromatic forms of visual conspicuousness. For example, they explain that while pirarr-prararrpa is glossed in the Walpiri Dictionary as referring to “bright colour or light colour (white, yellow, orange, red, silver) as opposed to dark colours (black, blue, green, purple), in fact, it seems clear that the meaning of pirarr-pirarrarpa does not refer to ‘colour’ at all: what unites the range including elements like ‘yellow’, ‘orange’, ‘silver’, ‘light’ and ‘shiny’ is not colour but high visibility (against a different background).” (Goddard and Wierzbicka 2014, 88).

The HTE’s conceptual structure allows for representing the meaning of pirarr-pirarrarpa in line with Goddard and Wierzbicka’s interpretation, without taking recourse in the concept of colour. This could be achieved by assigning pirarr-
pirararrpa to the general field pertaining to sight/sight and vision (01.09.08) and its subordinate visible/visibility (01.09.08.11), and then by choosing the desired level of specificity within the taxonomical chains:

Clearly visible (01.09.08.11|13) → sharp/distinct (01.09.08.11|13.04) → by contrast (01.09.08.11|13.04.01) + conspicuousness (01.09.08.11|07.02) → state of being clearly visible (01.09.08.11|07.02.01) → person/thing: stand out (01.09.08.11|07.02.01).

Fig. 9 Conceptual chains for pirarr-pirararrpa based on the taxonomy of the HTE

The HTE’s conceptual structures appears to be also capable of representing culturally encoded meanings that are foreign to the English conceptual universe. Goddard and Wierzbicka note that, in English, visual patterns are interpreted as variegation or alternation between different colours, whereas in Walpí this is not the case. Words like jiirlpari-jiirlpari and Kuruwarri-kuruwarri refer to patterns on animals, with reference not to colouring, but to visual contrast and markings. The
former word, if I understand Goddard and Wierzbicka correctly, can refer to monochrome patterning where different textures create visual contrast; while the latter denotes markings that evoke the idea of ceremonial designs:

The meaning of the reduplicated form kuruwarri-kuruwarri is not about ‘stripes’ and other form of variegation (patches, blotches, spots, etc.), but rather about visual patterns which look like markings made somewhere by someone to convey some meaning (as in some of the senses of kuruwarri itself). (Goddard and Wierzbicka 2014, 92).

While English speakers may conceptualise visual patterns differently, it is possible to map the meaning of jiirlpari-jiirlpari and Kuruwarri-kuruwarri on the HTE conceptual structure with some degree of accuracy.

The HTE gives us the choice to distinguish between patterns with reference to colour, or with reference to the animal world. To represent the meaning of jiirlpari-jiirlpari as explicated by Goddard and Wierzbicka, one should discard the field PERTAINING TO COLOUR (01.10.09) and its subcategories, such as VARIEGATION/VARIEGATED (01.10.09.08) and SPOTTING/BEING SPOTTED (01.10.09.08.06) or STRIPED (01.10.09.08.02), MARKED WITH RINGS OF COLOUR (01.10.09.08.04), etc. One should choose instead the field PERTAINING TO ANIMAL BODY (01.05.11) and its subordinate MARKINGS/COLOURINGS (01.05.11.01). Then one could select DAPPLED/SPOTTED (01.05.11.01|03) or STRIPED/LINED (01.05.11.01|04), and so on, as appropriate.
While there is reference to colouring in category 01.05.11.01, the fact that it is subsumed under PERTAINING TO ANIMAL BODY (01.05.11), rather than under PERTAINING TO COLOUR (01.10.09), approximates the Walpiri conceptualisation of patterning. Also, the emphasis can easily be shifted from the idea of colour to that of visual conspicuousness.
by adding reference to the fields of visual contrast (01.09.08.11|13.04.01), conspicuousness (01.09.08.11|07.02.01) and striking manifestness (03.09.02.01|01). Similarly, the culture-specific association of *Kuruwarri-kuruwarri* with meaningful markings and visual signs can be conveyed by adding the chain serving as sign (03.09.04) → marked (03.09.04.03).

Finally, it seems that the HTE’s categories can successfully handle subtle semantic nuances that might not be very salient in British culture, such as those obtaining between the words *liirlpari* and *ratarata*. According to Goddard and Wierzbicka this lexical pair is representative of the Walpiri “preoccupation with things “shining” somewhere in the speaker’s environment” (Goddard and Wierzbicka 2014, 89). *Liirlpari*, the two linguists explain, indicates an object shining in the distance which “visually ‘stands out’ in a particular time and place” (ibidem), while *Ratarata* “does not refer to visibility in the distance, but it too refers to standing out against the surroundings: ‘typically used of something white which stands out on a dark surface’.” (Goddard and Wierzbicka 2014, 90). Examples of the use of this word include “There is a lot of edible sap glistening right there”, and “we can see them (drops of water) glistening on the grass.” (ibidem).

In order to map *liirlpari* and *ratarata* in conformity with Goddard and Wierzbicka’s analysis, one could start by assigning both words to the categories clearly visible (01.09.08.11|13), be manifest (03.09.02) → strikingly: stand out (03.09.02|02), and reflective (01.10.08.05) → lustrous/shining with reflected light (01.10.08.05|04). Then one could differentiate between *liirlpari* and *ratarata* by linking the former to visible at a distance (01.09.08.11|04), and the latter to clearly visible (01.09.08.11|13) → sharp/distinct (01.09.08.11|13.04) → by contrast (01.09.08.11|13.04.01).
All in all, from this preliminary test on the Walpiri visual lexicon it appears that, while it surely does not provide the “maximally neutral epistemological perspective” that Goddard and Wierzbicka strive to achieve (Goddard and Wierzbicka 2014, 85), the HTE taxonomy is capable of representing radically different conceptual representations of phenomena. In other words, it can serve as a heuristic tool for cross-lingual semantic mapping. Needless to say, the application of the HTE conceptual structure in cross-linguistic context needs to be tested systematically on many more case studies, and even if it were to pass all tests, it will always suffer from the ultimate subjectivity of semantic categorisations. After all, it is up to the individual linguist to interpret the word and decide which categories better represent it—and there is no strictly scientific methodology for this task (cf. Kay 2004, 48 and Fischer 2004, 56). Yet, paired with traditional lexicographic information and (where available) data gathered from extensive parallel corpora, cross-lingual semantic mapping based on the HTE taxonomy
could prove a valuable addition to the standard features of digital bilingual dictionaries.

Acknowledgements
The development of BTW has been made possible by two NEH Digital Humanities Start-up grants, Level I and II, and by the continuous support of the Mangalam Research Center for Buddhist Languages. The software of BTW is the work of Dr. Louis-Dominique Dubeau, while the lexicographic template is mainly the output of Prof. Luis Gómez, who is co-directing the project with myself. I am indebted to Ruth-Ann Liu-John and Bennett Buchanan for the infographics included in this paper. Finally, I would like to express my sincere gratitude to Prof. Christian Kay and her team at the University of Glasgow for granting permission to use the Historical Thesaurus of English (HTE) within BTW. The integration of the HTE’s semantic categories in BTW is my own idea, and any flaws in their use for bilingual lexicography are my sole responsibility.

References


Kay, Christian. 2012. Developing the Historical Thesaurus of the OED. In *Current


