

example, *large – big* in English has an average similarity score of 5.9, whereas Spanish, Kiswahili and Welsh speakers rate the closest concept pair in their native language to have a similarity score between 2.7 and 3.8. What is more, *woman – wife* receives an average similarity of 0.9 in English, 2.9 in Spanish, and greater than 4.0 in Kiswahili and Welsh. The examples from Spanish include *banco – asiento* (*bank – seat*) which receives an average similarity score 5.1, while in the other three languages the similarity score for this word pair does not exceed 0.1. At the same time, the average similarity score of *espantosamente – fantásticamente* (*amazingly – fantastically*) is much lower in Spanish (0.4) than in other languages (4.1 – 5.1). In Kiswahili, an example of a word pair with a higher similarity score than the rest would be *machweo – jioni* (*sunset – evening*), having an average score of 5.5, while the other languages receive 2.8 or less, and a notably lower similarity score is given to *wa ajabu - mkubwa sana* (*wonderful – terrific*), getting 0.9, while the other languages receive 5.3 or more. Welsh examples include *yn llwyr - yn gyfan gwbl* (*purely – completely*), which scores 5.4 among Welsh speakers but 2.3 or less in other languages, while *addo – tyngu* (*promise – swear*) is rated as 0 by all Welsh annotators, but in the other three languages 4.3 or more on average.

There can be several explanations for the differences in similarity scores across languages, including but not limited to cultural context, polysemy, metonymy, translation, regional and generational differences, and most commonly, the fact that words and meanings do not exactly map onto each other across languages. For example, it is likely that the other three languages do not have two separate words for describing the concepts in the concept pair: *big – large*, and the translators had to opt for similar lexical items that were more distant in meaning, explaining why in English the concept pair received a much higher average similarity score than in other languages. A similar issue related to the mapping problem across languages arose in the Welsh concept pair *yn llwyd – yn gyfan gwbl*, where Welsh speakers agreed that the two concepts are very similar. When asked, bilingual speakers considered the two Welsh concepts more similar than English equivalents *purely – completely*, potentially explaining why a higher average similarity score was reached in Welsh. The example of *woman – wife* can illustrate cultural differences or another translation-related issue where the word ‘wife’ did not exist in some languages (for example, Estonian), and therefore had to be described using other words, affecting the comparability of the similarity scores. This was also the case with the *football – soccer* concept pair. The pair *bank – seat* demonstrates the effect of the polysemy mismatch across languages: while ‘bank’ has two different meanings in English, neither of them is similar to the word ‘seat’, but in Spanish, ‘*banco*’ can mean ‘bank’, but it can also mean ‘bench’. Quite naturally, Spanish speakers gave the pair *banco – asiento* a higher similarity score than the speakers of languages where this polysemy did not occur.

An example of metonymy affecting the average similarity score can be seen in the Kiswahili version of the word pair: *sunset – evening* (*machweo – jioni*). The average similarity score for this pair is much higher in Kiswahili, likely because the word ‘sunset’ can act as a metonym of ‘evening’. The low similarity score of *wonderful – terrific* in Kiswahili (*wa ajabu - mkubwa sana*) can be explained by the fact that while ‘*mkubwa sana*’ can be used as ‘terrific’ in Kiswahili, it technically means ‘very big’, adding to the examples of translation- and mapping-related effects. The word pair *amazingly – fantastically* (*espantosamente – fantásticamente*) brings out another translation-related problem: the accuracy of the translation. While ‘*espantosamente*’ could arguably be translated to ‘amazingly’, more common meanings include: ‘frightfully’, ‘terrifyingly’, and ‘shockingly’, explaining why the average similarity score differs from the rest of the languages. Another problem was brought out by *addo – tyngu* (*promise – swear*) in Welsh,