A comparative study of free associations to loanword/native synonyms and variant pairs in Polish

Joanna Rączaszek-Leonardi, Faculty of Psychology, University of Warsaw

in collaboration with

Mirosław Bańko, Faculty of Polish Studies, University of Warsaw
Adianna Kucińska, Faculty of Polish Studies, University of Warsaw

Objective

The goal of this study was to observe differences in free associations evoked by pairs of synonyms, one of which is of foreign origin (and still recognizable as foreign), the other one native, as well as pairs of variant forms of a loanword, one of which has the original (or nearly original) spelling, the other one being assimilated graphically in the recipient language. The study was undertaken for two reasons: first, to identify possible semantic dimensions along which pairs of loan/native words and unassimilated/assimilated words may differ, and second, to check if the distinctions between the members of the pairs observed in linguistic analyses coincide with those identified in free associations.

Methods

Material

To study free associations, we chose 28 synonym pairs and 14 variant pairs, most of them from a list of 50 word pairs which had been previously compiled for the purpose of linguistic analysis. The criteria used in the selection were as follows: 1) both members of each pair were one-word expressions, 2) neither of them was an obsolete or very rare word in Polish, 3) neither of them was polysemous in an obvious way and 4) in the case of synonyms – their meaning was evaluated by the experimenters as being particularly close. Inadvertently, one pair was selected which fails to meet criterion 3, namely *doktor – lekarz*.

Next, this set of 28 synonymic and 14 variant pairs was divided in two subsets (with 14 synonymic and 7 variant pairs in each), roughly balanced for word classes. Finally, each of the subsets was further subdivided in such a way as to contain only one member of each pair and care was taken to balance the resulting subsets in terms of loan/native words and unassimilated/assimilated variants. Thus 4 lists of words were created, each 21 words long, each containing 14 members of synonymic pairs and 7 members of variant pairs, balanced for being loan words (or unassimilated variants).

The lists then served as the basis for association questionnaires. These took the form of 4 booklets in which each word from the given list was printed on a separate page, with three dotted lines on the right side of the word where participants were to write down three associations. Beneath each word, on the bottom of each page, there was also a scale on which participants were asked to mark how familiar they were with the word. One sample page of such a booklet is presented in Appendix 1. The order of the words within each booklet was maintained constant across participants.

Participants

The 88 participants were mostly university students studying at various departments who volunteered to take part. They were asked to participate by the members of research team in connection with their regular university courses. The gender ratio of women to men was 2:1.

Each participant was asked to fill out one, and only one, of the four questionnaires. Each of the four questionnaires was filled out by 22 participants, yielding a total of 88 questionnaires. As the task was to produce three associations for every word given in the questionnaire, we obtained 66 associations for each word.

Procedure

Questionnaires were given to participants together with the following instructions:

Masz przed sobą broszurę zawierającą listę słów. Każdy arkusz dotyczy innego słowa i postaraj się potraktować każdy z nich jako oddzielną całość. Uzupełniaj jeden arkusz na raz.	In this booklet you will find a list of words. Each page is devoted to one particular word and we'd like you to work through them one page at a time.
Jak uzupełniać arkusz	**How to fill in the questionnaire**
Przeczytaj słowo znajdujące się na górze strony, **pomyśl o nim przez chwilę**, a następnie:	Read the word on top of the page, **take a moment to think about it** and then:
(1) Wpisz obok słowa swoje trzy pierwsze skojarzenia.	**(1) Fill the given space next to the word with your first three associations for the word given**.
Nie przejmuj się tym, jaką postać mają twoje skojarzenia – wpisz to, co przyjdzie ci do głowy (słowa, określenia, zwroty itd.).	Don't worry about the form of the association – write whatever comes to mind (words, attributes, collocations, etc.)
Ważne: nawet, jeśli nie znałeś/znałaś słowa wcześniej, postaraj się wpisać, z czym ci się ono kojarzy (jego brzmienie, wygląd, jakiej dziedziny, twoim zdaniem, może dotyczyć etc.).	Important: whether you know the meaning of the word or not, try to associate it with something (maybe you can associate the sound of it or a domain to which it seems to pertain, etc.).
(2) Określ, w jakim stopniu jest ci znane każde słowo, zakreślając kółkiem odpowiedni punkt na podanej osi.	**(2)** Mark how well you believe you know the word by circling the appropriate answer on the scale below it.

Participants proceeded to fill in the questionnaires. It took them approximately 20 minutes to produce associations for each of the 21 words in each booklet. At the end of each booklet there were questions about the sex, age, education and native language of the participants.

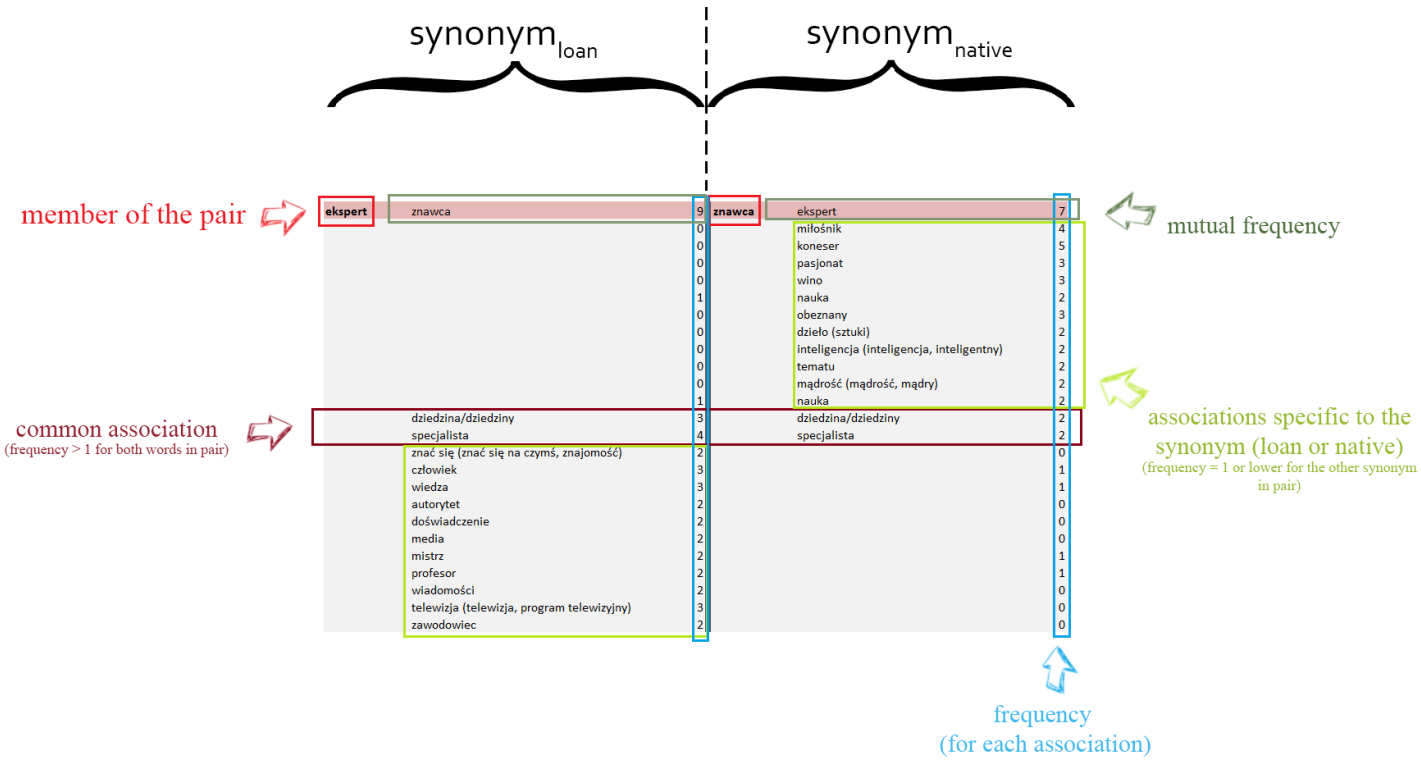
Data coding

Data were coded for item analysis: i.e., for each word, a list of 66 associations was compiled and frequency distributions were calculated on this basis. Any 1) unintelligible (unreadable) associations, 2) blank spaces left by participants or 3) instances of the same associations being repeated twice or three times were considered null responses and were coded as such; their number for each word was recorded as well.

Lastly, the rough data were aggregated into clusters. Each such cluster included derivatives of the same parent word (i.e. words having a common stem), word forms differing due to inflection and/or phrases including any of the above-mentioned words or word forms, whether idiomatic or not. The results of this step of analysis are given in Appendix 2.

Data analysis

For each pair we listed associations that were shared between the members of the pair and those which were idiosyncratic to one member of the pair. By definition, the former have a frequency of at least 2 for both members of the pair, while the latter, again by definition, have a frequency of at least 2 for at least one member of the pair. Associations of lower frequencies were not included, with the exception of mutual associations (occurring between the two pair members themselves), which were included irrespective of their frequency. Results are presented in Appendix 3 (which is based on Appendix 2, but uses more advanced clustering, so occasional discrepancies can be found between Appendix 2 and Appendix 3). The manner of presentation is explained in the diagram below.



Lastly, an index of semantic proximity was computed, as the proportion of the associations shared by the members of each pair (plus the number of mutual associations between the pair members) to the overall number of associations obtained for this pair (see Appendix 4).

Discussion

The distribution of associations demonstrates that in each word pair the perception of one element differs from the perception of the other. In some pairs the distinctions are very much akin to those observed in the previous linguistic analysis, cf. *dealer – diler*, *infekcja – zakażenie*, *komponent – składnik*, *strofa – zwrotka*, *toksyczny – trujący*. In other pairs, however, the distinctions identified in the associations are different from those arrived at in the linguistic research, cf. *chips – czips* or *helikopter – śmigłowiec*. Further studies are needed to test the relevance of particular dimensions observed in the linguistic and the psycholinguistic research. These will be conducted at a later stage of the current project.

Though the elements of the pairs inspected differ on the 'foreign – native' or 'unassimilated – assimilated' scales, it is rather unlikely that the difference in their perception is related directly to their origin. Loanwords are perceived differently from their native synonyms not because they have a foreign origin, but because they have some characteristics resulting from their origin, e.g. a different phonological structure and different (usually lower) frequency.

The indices of semantic proximity vary from 0.05 (for *fan – miłośnik*) to 0.74 (for *e-mail – mejl*) on a scale from 0 to 1 and have the average value of 0.49. One might argue that they do not reflect the real similarity of words within particular pairs and would be higher if calculated on data grouped previously according to some semantic criteria. Semantic grouping would indeed increase the number of associations shared by both members of particular pairs, but would have to rely on subjective criteria and could be questioned. It was decided, therefore, to base the indices on data grouped according to only formal criteria. Though not perfect, the indices calculated this way are still reliable to a certain extent, e.g. they are generally higher for variant pairs (the average value of 0.57) than for synonym pairs (the average value of 0.45), which accords with expectations. All in all, the proportions among the indices are more important than their absolute values. At a later stage of the project, the semantic proximity between the same words will be calculated in semantic spaces and the results will be compared to those obtained from the free association study.

It is worth noticing that the mutual attraction between members within a pair is not equal: the second element is more often found among the associations given for the first element than vice versa. The difference is not a small one: in total, 234 mutual associations were given for loanwords and only 110 for native words. As the mutual frequencies in variant pairs are generally low (most often zero), it is the synonym pairs that are responsible for this effect. The asymmetry in mutual associations may be caused by different textual frequencies of the synonym pair members or by different familiarity of these words, as declared by the subjects. If time allows, the potential relationship between mutual association frequencies, textual frequencies and the declared familiarity of words will be investigated later on in the project.

Published on 30/04/2014, updated on 29/06/2015