Supplementary material

Lehecka, Tomas. 2021. Normative ratings for 111 Swedish nouns and corresponding picture stimuli. Nordic Journal of Linguistics XX(x), XXX–XXX.

The supplemented .csv file contains normative ratings collected from the Finland-Swedish subjects for the dataset described in the paper. The file also contains all the ratings to which the Finland-Swedish dataset is compared to in the paper. Below, a short description of each column is provided.

Item_nr_in_Multipic:

The ID number of the picture stimulus in question in the MultiPic database (Duñabeitia et al. 2017).

English_name:

The English word for the item in question, which corresponds to the modal answer to the stimulus in the English dataset of the MultiPic database (Duñabeitia et al. 2017).

Picture_nr_in_Greyscale_PNG:

The file name of the picture stimulus in Duñabeitia et al. (2017). This picture file was used to collect the norms for name agreement and category typicality.

Name_agreement_Modal_answer:

The modal answer to the picture stimulus (Survey 1).

Gender_of_Modal_answer_in_Swedish:

The grammatical gender of the modal answer (previous column) in Swedish.

Nr_modal_answers:

The number of subjects who gave the modal answer (Survey 1).

Pct_modal_answers:

The percentage of subjects who gave the modal answer (Survey 1).

Nr_different_answers:

The number of different answers that the subjects gave to the stimulus (Survey 1).

Nr_idiosyncratic_answers:

The number of answers that were given by no more than one subject (Survey 1).

Nr_invalid_answers:

The number of answers that were invalid (e.g. answers that were not words or that were clearly written as jokes) (Survey 1).

Shannon_H_Index:

The *H* information statistic (Shannon 1948) which quantifies the level of name agreement (Survey 1).

Semantic_category_Swedish:

The Swedish word for the semantic category that the item represents. This word was used in Survey 2 to collect the category typicality ratings.

Semantic_category_English:

The English translation of the category name (previous column).

Category_typicality_Mean:

The mean category typicality rating of the picture stimulus (Survey 2).

Category_typicality_SD:

The standard deviation of the category typicality ratings (Survey 2).

Stimulus_word_for_AoA&Subj_freq_ratings:

The stimulus word used to collect the AoA and subjective frequency ratings (Surveys 3 & 4). This word corresponds to the modal answer in Survey 1.

AoA_Mean:

The mean AoA rating of the stimulus word (Survey 3).

AoA_SD:

The standard deviation of the AoA ratings (Survey 3).

Subj_freq_Mean:

The mean subjective frequency rating of the stimulus word (Survey 4).

Subj_freq_SD:

The standard deviation of the subjective frequency ratings (Survey 4).

H_INDEX_[ ]_DUNABEITIA_ET_AL_2017:

The values of the *H* index for the picture stimulus in several languages, taken from the Duñabeitia et al. (2017) dataset. These values are used for the comparisons in Table 3 in the paper.

AOA_KUPERMAN_ET_AL_2012:

The AoA rating for the corresponding stimulus word in English (see the column English_name) in Kuperman, Stadthagen-Gonzalez & Brysbaert (2012). This rating is used for comparisons in Table 4.

AOA_BRYSBAERT_BIEMILLER_2017:

The AoA rating for the corresponding stimulus word in English in Brysbaert & Biemiller (2017). This rating is used for comparisons in Table 4.

AOA_LUNIEWSKA_ET_AL_2016_SWEDISH:

The AoA rating for the stimulus word in Łuniewska et al. (2016), collected in Sweden. This rating is used for comparisons in Table 4.

AOA_LIND_ET_AL_2015:

The AoA rating for the correponding stimulus word in Norwegian in Lind et al. (2015). This rating is used for comparisons in Table 4.

SUBJ_FREQ_ENCOUNTERED_BALOTA_ET_AL_2001:

The subjective frequency rating of the corresponding stimulus word in English in Balota et al. (2017). The rating represents the value collected using the verb *encounter* (as opposed to e.g. *hear* or *say*). This rating is used for comparisons in Table 5.

SUBJ_FREQ_SOARES_ET_AL_2017:

The subjective frequency rating of the corresponding stimulus word in Portuguese in Soares et al. (2017). This rating is used for comparisons in Table 5.

French_name:

The corresponding French word for the stimulus word in question. This word was used to find the corresponding subjective frequency ratings in the French datasets.

SUBJ_FREQ_FERRAND_ET_AL_2008:

The subjective frequency rating of the corresponding stimulus word in French in Ferrand et al. (2008). This rating is used for comparisons in Table 5.

SUBJ_FREQ_DESROCHERS_THOMPSON_2009:

The subjective frequency rating of the corresponding stimulus word in French in Desrochers & Thompson (2009). This rating is used for comparisons in Table 5.

FREQ_PER_1M_LOG10_[ ]:

The objective frequency of the stimulus word in several Swedish corpora in Språkbanken (http://spraakbanken.gu.se). The values represent the $\log_{10}$ of the relative frequency of the word per 1 Million tokens.

FREQ_PER_1M_LOG10_OPENSUB:

The objective frequency of the stimulus word extracted from the OpenSubtitles corpus (Lison & Tiedemann 2016).