

## Can prediction-based distributional semantic models predict typicality?

Tom Heyman<sup>1</sup> & Geert Heyman<sup>1</sup>

<sup>1</sup> University of Leuven

### Author Note

The manuscript was written in R (R Core Team, 2016) using the packages `papaja` (Aust & Barth, 2017) and `rmarkdown` (Allaire et al., 2016). One can find the data, the `.Rmd` file, including the analysis code, and the pre-print on the project's Open Science Framework page (<https://osf.io/nkfjy/>). This pre-print was created on 14/01/2018.

Correspondence concerning this article should be addressed to Tom Heyman, Department of Experimental Psychology, University of Leuven, Tiensestraat 102, 3000 Leuven, Belgium. E-mail: [tom.heyman@kuleuven.be](mailto:tom.heyman@kuleuven.be)

## Abstract

Recent advances in the field of computational linguistics have led to the development of various prediction-based models of semantics. These models seek to infer word representations from large text collections by predicting target words from neighboring words (or vice versa). The resulting representations are vectors in a continuous space, collectively called word embeddings. Although psychological plausibility was not a primary concern for the developers of predictive models, it has been the topic of several recent studies in the field of psycholinguistics. That is, word embeddings have been linked to similarity ratings, word associations, semantic priming, word recognition latencies, etcetera. Here, we build on this work by investigating category structure. Throughout seven experiments, we sought to predict human typicality judgments from two languages, Dutch and English, using different semantic spaces. More specifically, we extracted a number of predictor variables, and evaluated how well they could capture the typicality gradient of common categories (e.g., *birds*, *fruit*, *vehicles*,...). Overall, the performance of predictive models was rather modest, and did not compare favorably to that of an older count-based model. These results are somewhat disappointing given the enthusiasm surrounding predictive models. Possible explanations and future directions are discussed.

*Keywords:* distribributional semantics; typicality; predictive models

Word count: 8575

Can prediction-based distributional semantic models predict typicality?

Many models of semantic representation rely on the idea that words appearing in similar contexts have related meanings (Harris, 1954). In so-called distributional semantic models, words occupy a particular position in an  $N$ -dimensional space depending on their (co-)occurrence in certain contexts. By applying specific algorithms to, typically, large scale text corpora, one can derive a geometrical representation of the lexicon. Consequently, words with a related meaning, say *cat* and *dog*, tend to be located in each other's vicinity, as for instance measured by the cosine similarity between the respective vectors.

Even though the fundamental premise underlying all distributional semantic models is the same, there is considerable variability in terms of how context is defined, a relatively small window of five to 10 words or an entire text region such as a book chapter or webpage, and how information is acquired, Rescorla-Wagner or Hebbian-type learning mechanisms, for example (see Lenci, 2018; Yee, Jones, & McRae, 2018, for reviews). The present study specifically focused on the recently developed predictive models (e.g., Mikolov, Sutskever, Chen, Corrado, & Dean, 2013). They are called as such because these models learn to *predict* a target word based on context words (i.e., continuous bag-of-words architecture, henceforth CBOW) or vice versa (i.e., skip-gram architecture) by optimizing vector weights. Psychological plausibility was not necessarily a primary concern in the development of these models, yet they mimic human performance fairly well, at least in comparison to other types of distributional semantic models (Baroni, Dinu, & Kruszewski, 2014).

Because of their initial success, predictive models have been the topic of much research, primarily within computational linguistics, but recently also in cognitive psychology and psycholinguistics (Hollis, 2017; Hollis & Westbury, 2016; Mandera, Keuleers, & Brysbaert, 2017; Pereira, Gershman, Ritter, & Botvinick, 2016). These studies have related word embeddings to similarity ratings, word associations, semantic priming, word recognition latencies, and ratings of valence, dominance, arousal, and concreteness, thus providing converging evidence regarding the validity of such models in terms of resembling the human

mental lexicon. One aspect that has received little attention in this respect, is category structure. Seminal work by Rosch (1975) illustrated that categories’ internal structure is graded such that some exemplars are better, more typical members of a category (e.g., *robin* for the category *birds* or *chair* for the category *furniture*), than others (e.g., *chicken* for the category *birds* or *rug* for the category *furniture*). In general, people have similar intuitions about the continuous nature of most categories, as evidenced by their performance on a typicality rating task (e.g., “on a scale from 1 to 7, how typical is a *robin* for the category *birds*?”). That is, participants’ typicality ratings have been shown to be very consistent, which is suggestive of a shared category organization.

The present paper examines whether predictive models capture this underlying category structure. In a first experiment, we computed the similarity between a number of category exemplar vectors (e.g., *robin*, *crow*, *eagle*, *duck*, *chicken*, etc.) and the corresponding prototype vector (i.e., *bird* in this example). The resulting similarities were then correlated with typicality ratings obtained from Dutch and English norming studies (De Deyne et al., 2008; Morrow & Duffy, 2005). One would expect proximity to the prototype to be a good indicator of an exemplar’s perceived typicality. In fact, previous studies have provided evidence for this claim using non-predictive distributional semantic models (also called count models), more precisely, Latent Semantic Analysis (LSA) and the Bound Encoding of the Aggregate Language Environment (BEAGLE) model (Connell & Ramscar, 2001; Jones & Mewhort, 2007). Given that predictive models are generally considered to be superior in terms of mirroring human behavior (e.g., Baroni et al., 2014), one would expect sizeable correlations between typicality and model-based similarity to the prototype.<sup>1</sup>

That being said, a recent study by Vulic, Gerz, Kiela, Hill, and Korhonen (2017) showed that various (predictive) distributed models performed poorly in terms of mimicking

---

<sup>1</sup>Qualifying correlations as “sizeable” is admittedly rather vague. As such, the present study should be considered exploratory in that we aim to gauge the performance level of predictive models in the context of predicting typicality ratings.

graded lexical entailment ratings (e.g., “on a scale from 1 to 7, to what degree is *robin* a type of *bird*?”). Applying both unsupervised (e.g., cosine similarity between the word pairs’ vectors) and supervised techniques (i.e., learning to predict the lexical entailment ratings from the word embeddings using various regression models), did not result in a satisfying account of human performance. However, the data collection procedure Vulic et al. employed, differs considerably from the usual typicality rating task. For one, many categories appeared just a handful of times in the entire set, making it difficult to examine the representation of those categories. More importantly though, several word pairs did not follow the normal “X is a (potential) hyponym of Y” format (i.e., the stimulus set also featured antonyms like *girl-boy*, co-hyponyms like *green-blue*, meronyms like *eye-face*, hypernyms like *mammal-elephant*, and synonyms like *carpet-rug*), or were even completely unrelated (e.g., *foot-plant*). The inclusion of such items might have restricted the range of the ratings for the actual hyponyms in the experiment. Consequently, the ratings might only capture *within*-category structure to a limited extent. In addition, the present study considered a number of other predictor variables, inspired by research on concepts and categories in the domain of cognitive psychology, that might provide a better fit of human typicality/graded lexical entailment ratings.

## Experiment 1a

### Method

**Typicality ratings.** As mentioned in the introduction, we took typicality ratings from the norms of De Deyne et al. (2008), and Morrow and Duffy (2005). The former study was conducted in Dutch and involved 16 categories: *amphibians*, *birds*, *clothes*, *fish*, *fruit*, *insects*, *kitchen utensils*, *mammals*, *musical instruments*, *professions*, *reptiles*, *sports*, *tools*, *vegetables*, *vehicles*, and *weapons*. However, we omitted the category *amphibians* because it only contained five exemplars. Furthermore, we excluded multi-word category exemplars (e.g., *red cabbage*), as the composition of several terms is a non-trivial matter in

distributional semantics (Lenci, 2018). Exemplars with punctuation marks (e.g., *t-shirt*, *maïs* in Dutch) were also discarded because annotational variability could result in unreliable word embeddings. Ultimately, all categories did contain at least 20 exemplars.

Morrow and Duffy (2005) were interested in the influence of age on concept representations, and, as a consequence, they gathered typicality ratings from older and younger adults separately. Though the correlation between both sets of estimates was generally very high, we decided to keep this distinction in all our analyses. Morrow and Duffy (2005) examined the following 14 categories: *animals*, *birds*, *cutlery*, *clothes*, *flowers*, *fruit*, *furniture*, *insects*, *jewellery*, *musical instruments*, *tools*, *trees*, *vegetables*, and *vehicles*. Using the same procedure as described above, we omitted the categories *cutlery*, *jewellery*, and *trees* as these did not contain at least 20 single-word, non-punctuated exemplars (i.e., three, 17, and 11 exemplars, respectively).

**Word embeddings.** For the present study, we used off-the-shelf distributed semantic representations (from Mandera et al., 2017; Mikolov et al., 2013; Tulken, Emmery, & Daelemans, 2016), meaning that the respective model parameters were not tuned to obtain optimal performance in terms of predicting typicality ratings. Two semantic spaces, one per architecture, CBOW or skip-gram, were selected for each language. The Dutch word embeddings were taken from Tulken et al. (skip-gram with negative sampling, 320 dimensions, Wikipedia dump as training corpus), and Mandera et al. (CBOW with negative sampling, 200 dimensions, concatenation of subtitle and SONAR-500 corpora as training data). The English embeddings came from Mikolov et al. (skip-gram with negative sampling, 300 dimensions, Google News as training corpus), and Mandera et al. (CBOW with negative sampling, 300 dimensions, concatenation of subtitle and UKWAC corpora as training data).

From these semantic spaces, we extracted the exemplar vectors corresponding to the stimuli selected from De Deyne et al. (2008), and Morrow and Duffy (2005). In order to derive a prototype for each category, we followed two approaches. One involved taking the average of the respective exemplar vectors, in the other, we used the category label as a

proxy for its prototype (see Jones & Mewhort, 2007 for a similar procedure). Note that with regard to the latter approach, we slightly changed the category labels to better resemble the respective prototype (i.e., *animals* became *animal*, *birds* became *bird*,...). We did, however, keep the original category label when suchlike singularizing would create multi-word prototypes (i.e., *clothes* would become *piece of clothing*). Because the label *musical instrument(s)* comprises two words in any case, we only employed the averaging-across-exemplar-vectors method to that category. This was not an issue in Dutch because *musical instrument(s)* is one word (and so is *kitchen utensils*). Thus, except for *musical instruments* in English, there were two prototype vectors per category, one obtained by averaging across exemplars and one based on the category label itself. In a next step, we calculated the cosine similarity between every exemplar vector and the corresponding prototype vectors (see Tables 1 - 4 for summary statistics). The resulting similarities were then correlated with the typicality ratings for each category and prototype derivation method separately.<sup>2</sup> Note that in all experiments, we used the Pearson correlation coefficient.

## Results and discussion

The results are summarized in Figure 1 (Dutch), Figure 2 (English, younger adults), and Figure 3 (English, older adults). All figures consist of two panels depicting the results for the prototype-as-average method (top panel) and the prototype-as-category-label method (bottom panel). The x-axis always indicates the correlation between typicality and cosine similarity based on the CBOW model, whereas the y-axis does the same for the skip-gram model. As an illustration, consider the category *birds* in Dutch. The cosine similarity between the prototype-as-average and the exemplar vectors showed a correlation with typicality of .29 and .20, for the CBOW and skip-gram model, respectively. The corresponding results for the prototype-as-category-label method are .39 and .42.

All figures also contain grey dotted lines indicating the average correlation across

---

<sup>2</sup>In the analysis code, the term category vector refers to the prototype-as-category-label approach, whereas the prototype-average method is simply labelled prototype.

categories. As can be seen, these average correlations ranged from .08 to .36, and were in general rather low. Some categories unexpectedly demonstrated very small and even numerically negative correlations between prototype similarity and typicality, most notably *flowers* in English and *mammals* in Dutch. Figures 4 and 5 show the cases with the worst overall performance. The plots suggest that the low correlations are not the result of a few outliers, a non-linear pattern, or a restriction in range (see also Tables 1 - 4 and the tables in the Appendix). The conclusion based on the present results is rather sobering given the enthusiasm surrounding these models. Though the typicality gradient of some categories can be mimicked with reasonable accuracy (e.g., *vehicles* in Dutch, *furniture* in English), most correlations ranged from .10 to .40, meaning that prototype similarity explained only 1% to 16% of the variability in typicality ratings.

Note that these correlations depend on the reliability of the typicality ratings, and that of the similarity-to-the-prototype metrics. Typicality ratings are usually very reliable as indicated by the split-half reliability estimates reported by De Deyne et al. (2008), which ranged from .87 to .98 (see Table 5). For the English ratings, we took the correlation between the younger and older adults' ratings as an index of their reliability (see Table 6 of Morrow & Duffy, 2005). These were a bit smaller for some categories, ranging from .70 to .93, but still acceptable for the present purposes (see Table 6). To estimate the reliability of the prototype similarity measures, we correlated for each category the cosine similarities between exemplar and prototype vectors obtained from one semantic space (i.e., CBOW) with those based on the other semantic space (i.e., skip-gram). This way, one can get an idea of how stable such similarity-to-the-prototype metrics are across different semantic spaces. Indeed, one would expect the similarity between, say *lettuce* and the prototypical vegetable, to be consistent across training corpus and model architecture. The resulting correlations are also presented in Tables 5 (Dutch) and 6 (English). Averaged across categories, we obtained correlations of .59 (prototype-as-average) and .53 (prototype-as-category-label) for Dutch, and .64 (prototype-as-average) and .59 (prototype-as-category-label) for English. Within the



context of reliability estimates, these figures are fairly low, especially when we look at the level of individual categories (e.g., *kitchen utensils* in Dutch and *birds* in English).

In that sense, it is not surprising that the relation between prototype similarity and typicality is rather weak. After all, the correlation between two variables gets attenuated if the reliability of one or both variables is questionable. More precisely, following Spearman’s well-known formula (Spearman, 1904), the true correlation between two variables gets deflated by the square root of the product of their reliabilities. So, as an example, consider the typicality ratings for Dutch, with an average reliability estimate of .95, and the prototype-as-average approach, with an average reliability estimate of .59. In this case, the true correlation between prototype similarity and typicality, whatever it may be, would shrink by a factor  $\sqrt{.95 \times .59} = .75$ . Thus, a true correlation of, say, .80 would be reduced to  $.75 \times .80 = .60$ . Note that this figure is just a general approximation; in practice it will vary across categories. Still, the observed correlations in the present experiment do not come close to this mark. Importantly, one can, in principle, predict such typicality ratings to a much better extent, as illustrated by the study of Voorspoels, Vanpaemel, and Storms (2008). They used pairwise similarity judgments as input to construct a semantic space of which prototype-based, optimal predictions of typicality were derived (among other things). Depending on the semantic space’s dimensionality, Voorspoels et al. obtained an average correlation with typicality of up to .73 for that prototype predictor.

In sum, the performance of the currently tested predictive models was relatively underwhelming. In this regard, it is noteworthy that predictive models were also not able to approximate selectional preferences very well (Baroni et al., 2014). In the latter task, people have to judge the fit of a noun as an argument to a specific verb (e.g., *people eat* would get higher overall preference ratings than *clothes eat*). Following a similar approach as used in the present study, Baroni and colleagues averaged across the 20 most strongly associated nouns to obtain the prototypical argument of a particular verb. The cosine similarity between this prototype and various potential argument vectors was then correlated with

human preference ratings. The results showed that other distributional semantic models performed on par or better than predictive models, which led Baroni et al. to suspect that “this averaging approach... might be problematic for prediction-trained vectors” (p. 242).

It begs the question why averaging would be so problematic, though. Other mathematical operations on such vectors have proven to be successful, for instance in the context of solving analogies (e.g., the nearest neighbor of *king* - *man* + *woman* is typically *queen*). Furthermore, prototypes are usually operationalized as an average within the psychological literature (e.g., Voorspoels et al., 2008), so the approach by Baroni et al. is the most straightforward and parsimonious. With regard to the current experiment, one might object that using an unweighted average could have disproportionately shifted the prototype towards atypical exemplars. Put differently, *chicken* influenced the *bird* prototype to the same extent as *robin* did. This could have added noise to the resulting prototype similarity estimates, which, in turn, would have attenuated the correlation with the typicality ratings. The prototype-as-category-label approach does not suffer from this (potential) issue, yet its performance seems comparable to the unweighted averaging method. Hence, it does not appear to be the case that averaging, whether weighted or not, is the prime reason for the relatively poor performance of predictive models.

Thus far, we have exclusively considered singular nouns, except for some stimuli that are typically encountered in multiples like shoes or cymbals. However, it is plausible that the semantic vectors corresponding to the plural forms better capture the true meaning of a concept.<sup>3</sup> Indeed, one aspect that might have contributed to the relatively poor overall performance is the presence of polysemous words. If one has to evaluate, for instance, how typical an *orange* is for the category *fruit*, there is little to no confusion about the intended meaning of *orange*. A term’s position in a semantic space based on text corpora, though, gets affected by its various senses. Thus, the color interpretation of the word *orange* substantially impacts its location in the semantic space, thereby rendering the derived

---

<sup>3</sup>We want to thank Marc Brysbaert for bringing this to our attention.

typicality predictions unstable (Pierrejean & Tanguy, 2018), or even uninterpretable. When it concerns rather typical items such as *orange* for *fruit* or *saw* for *tools*, it could attenuate the overall correlation between the model predictions and human ratings. Therefore, we followed up on the previous experiment by exclusively using the semantic vectors of the plural forms, and examined whether this would improve the typicality predictions.

## Experiment 1b

### Method

The method was completely analogous to that of Experiment 1a. The only exception was that we now extracted the exemplar vectors corresponding to the plural forms of the stimuli taken from the typicality norms of De Deyne et al. (2008), and Morrow and Duffy (2005). Note that some stimuli were already plural (e.g., *shoes*), or had no distinct plural form. The latter was for instance the case for all exemplars of the category *sports* in Dutch. Next, we calculated a new prototype by averaging across the semantic vectors of the plural nouns. As for the prototype-as-category-label approach, we used the same category labels as in Experiment 1a.

### Results and discussion

The results are summarized in Figure 6 (Dutch), Figure 7 (English, younger adults), and Figure 8 (English, older adults). As can be seen, the use of semantic vectors corresponding to the stimuli’s plural form affected overall performance to a limited extent. When considering the CBOW models, the prediction of typicality, averaged across categories, seemed slightly worse in comparison to Experiment 1a. The skip-gram models on the other hand, performed somewhat better by relying on the vectors of the plural forms, except for the prototype-as-average approach in Dutch where the average correlation dropped from .23 to .18.

Taken together, even though plural forms can, in some cases, disambiguate singular

nouns (e.g., *orange* becomes *oranges*), we did not observe a general advantage of using their semantic vectors to predict human typicality ratings. This could be due to the fact that the stimuli in the typicality rating norms were (mostly) singular nouns, thus creating a mismatch with the semantic vectors. Also, the word embeddings of some plural forms might be more volatile relative to the singular counterparts, for instance, in cases where their occurrence frequency or contextual diversity is much lower. Indeed, the embeddings of low frequency words tend to be less stable in that their closest neighbors are more variable upon retraining the model (Pierrejean & Tanguy, 2018; Wendlandt, Kummerfeld, & Mihalcea, 2018).

To evaluate the reliability of the plural nouns’ semantic vectors, we correlated the similarity-to-the-prototype metrics across the two semantic spaces, as we did in Experiment 1a. The average correlations, obtained by collapsing over all categories, were .51 (prototype-as-average) and .44 (prototype-as-category-label) for Dutch, and .67 (prototype-as-average) and .59 (prototype-as-category-label) for English. When compared to the corresponding results of Experiment 1a, there is once again no clear-cut advantage of switching to plurals. The reliability estimates for Dutch are even a bit worse, perhaps because some plural forms are actually polysemous words with a dominant, alternative sense (e.g., *fluiten* means *to whistle* and *flutes*, or *vliegen* means *to fly* and *flies*).

In sum, there is ample room for improvement in terms of predicting typicality ratings based on predictive distributional semantic models, regardless of whether one relies on the embeddings of (mostly) singular nouns or their plural counterparts. In an attempt to provide a better account of the typicality gradient in common categories, we conducted three follow-up experiments, each examining different potential predictor variables extracted from the semantic spaces, inspired by existing literature regarding concepts and categories.

## Experiment 2a

According to Rosch and Mervis (1975), the most typical members of a category do not only possess features characteristic of that category, but they also have the least features in

common with members of contrasting categories. Thus far, we have only considered the former aspect, yet a number of studies have shown contrast effects in accounting for typicality judgements (e.g., Ameel & Storms, 2006; Dry & Storms, 2010; Voorspoels, Storms, & Vanpaemel, 2012). For instance, Ameel and Storms (2006) derived a semantic space by applying multidimensional scaling to a pairwise similarity matrix, after which they calculated a centroid for each category by averaging across the resulting exemplar vectors. The similarity between these exemplar and corresponding prototype vectors showed substantial correlations with typicality ratings, but performance improved when the centroid was moved in a direction opposite to where (members of) contrast categories were situated (Ameel & Storms, 2006 operationalized (dis)similarity as Euclidean distance). So in the present experiment, we examined whether dissimilarity to a contrast category would improve upon the typicality predictions of Experiment 1a.

## Method

In order to identify contrast categories, we relied on a study by Verheyen and Stukken (2010) in which native Dutch speakers were asked to generate a contrasting category for one abstract and one concrete category, following the procedure introduced by Malt and Johnson (1992). For each category that was included in the norms of De Deyne et al. (2008), we then selected the most frequent response (see Table 7 for an overview). As comparable English data were not readily available for all categories in the norms of Morrow and Duffy (2005), we limited our analyses to Dutch.

Analogous to the prototype-as-category-label approach of Experiment 1a, we calculated the cosine similarity between each exemplar vector and that of the corresponding contrast category label (e.g., *banana* and *vegetable*). Unfortunately, there was no semantic vector corresponding to the contrast category of *kitchen utensils* (i.e., *poetsgerief* meaning *cleaning equipment*) for the skip-gram model. As Experiment 1b did not show an advantage of using plural forms for Dutch, if anything there might have been a slight disadvantage, we

stuck with the original, mostly singular stimuli. The semantic spaces were also those of Experiment 1a (and 1b). Note that an equivalent similarity measure for the prototype-as-average approach was not calculated, because it required exemplars of contrast categories that were not part of De Deyne et al.’s (2008) dataset (e.g., *hobbies*).

Next, the cosine similarity between exemplars and the contrast category label was correlated with the typicality ratings of Experiment 1a, *controlling* for the cosine similarity between the exemplars and the label of the category it actually belongs to (e.g., *banana* and *fruit*). These partial correlations thus express the unique contribution of (dis)similarity to contrast categories in explaining typicality ratings.

## Results and discussion

The results are summarized in Figure 9, which has the same structure as Figures 1 - 3, but with just one panel. The axes now represent partial correlations, with the similarity between the exemplars and the category label as the sole covariate. As expected from Rosch and Mervis’ (1975) account, the correlation coefficients were generally negative. Exemplars that were the most similar to the contrast category tended to be less typical when controlling for similarity to the actual category. However, the average across categories was rather close to zero, and only a few categories exhibited a substantial contrast effect. The most notable one, was the category *fish*, where similarity to the contrast category *mammal* was a strong predictor of typicality. This had to do with the fact that a number of exemplars are really mammals, even though they are frequently mentioned when people have to generate exemplars of the category *fish* (e.g., *dolphin*, *whale*, *orca*). Those items received fairly low typicality ratings (see Table A3 of the Appendix), but they nevertheless resemble the prototypical fish fairly well (i.e., they live in the ocean, they have fins, etcetera), which arguably contributed to the weak correlations observed in Experiment 1a. By taking their similarity to the contrast category *mammals* into account (i.e., they require air to breathe, they are child-bearing, etcetera), the prediction of the typicality ratings thus improves

markedly. That being said, most categories showed no or a very weak contrast effect, so there is still no satisfying explanation for the relatively poor overall performance observed in Experiment 1a.

### Experiment 2b

Thus far, we have only entertained the idea of an abstract category representation, whether it is by averaging across exemplars or by considering the label itself. That is to say, fine-grained, exemplar-level information gets filtered out this way. However, according to the instantiation principle (Heit & Barsalou, 1996), a category’s representation includes detailed information about its members. Consequently, the somewhat disappointing results of Experiments 1a and 1b could potentially be attributed to the use of inaccurate category architectures. Indeed, a study by Storms, De Boeck, and Ruts (2000) demonstrated the value of a predictor inspired by the instantiation principle in explaining typicality ratings (as well as three other dependent variables: response times from a speeded categorization task, category-naming frequencies, and exemplar-generation frequencies). They showed that pairwise similarity ratings between exemplars and the most frequently generated exemplar(s) of a category (e.g., *apple* for the category *fruit*) correlated significantly with the corresponding typicality ratings. In the current experiment, we followed a similar approach, but limited the instantiation of a category to a single exemplar, as assumed by Heit and Barsalou (1996). That is, we selected the most generated exemplar for each category, and examined whether the cosine similarity between its representation and that of all category exemplars could indeed predict typicality ratings.

### Method

To select the best exemplar for each category (i.e., the member that is the most likely to instantiate the category), we relied on data from an exemplar generation task, as did Storms et al. (2000). In this task, participants are presented with a category label, and have to name or write down as many exemplars as possible given a certain time window. In

addition to typicality ratings, the norms of De Deyne et al. (2008) also contain such exemplar generation data for the same categories. From those, we selected the most frequent response for each category (see Table 8 for an overview).

Next, we calculated the cosine similarity between each exemplar vector and that of the corresponding best exemplar (e.g., *banana* and *apple*, *lemon* and *apple*, and so on). Note that the list of category exemplars always included the best exemplar. Thus, in those cases (e.g., *apple* and *apple*), the cosine similarity is equal to 1. The similarity measures were derived from the same semantic spaces used in the previous experiments. The resulting cosine similarities were subsequently correlated with the typicality ratings described in the Method section of Experiment 1a. As in Experiment 2a, we limited our analyses to Dutch.

## Results and discussion

The results are summarized in Figure 10, which again has the same structure as Figures 1 - 3, but with just one panel. Overall, the instantiation-based predictor seemed to provide a decent account of typicality ratings. Interestingly, it succeeded in predicting typicality ratings for some categories (to a certain extent), where the other approaches discussed above failed. However, one might argue that the improved performance is in part artificially induced. The most generated exemplar is usually considered a very typical member of that category. As the cosine similarity for that item is equal to 1 by definition, it becomes an outlier that generally inflates the overall correlation. Indeed, if we were to exclude those exemplars from the analyses, the average correlation across categories would drop from .37 (CBOW) and .33 (skip-gram) to, respectively, .31 and .24.

With the latter caveat in mind, one can conclude that the predictive models tested here only capture category structure to a limited extent. More specifically, they performed quite well in some cases, but failed in other domains. This becomes more clear when considering the results of a multiple linear regression analysis on the typicality ratings for each category, using all four predictors (i.e., similarity to the prototype, both average-based



and label-based, similarity to the label of the contrast category, and similarity to the best exemplar)<sup>4</sup>. The goodness-of-fit measures (see Table 9), especially  $R^2_{adj}$ , which takes the number of independent variables into account, indicated that the internal organization of some categories was reasonably well approximated (e.g., *vehicles*, *fish*, *mammals*). Yet, this contrasted with the poor fit in other categories like *reptiles*, *clothes*, and *insects*.

A potential explanation for the mixed performance of predictive models is that many dimensions of the semantic space contain no information regarding category structure. As all dimensions were weighted equally, the (possibly) irrelevant ones may overwhelm the valid information conveyed by a limited number of critical dimensions. In a next experiment, we set out to test this hypothesis.

## Experiment 2c

The idea that specific dimensions convey certain information has already received some empirical support. More concretely, Hollis and Westbury (2016) linked specific principal components extracted from skip-gram-based word embeddings to certain variables like valence, arousal, and dominance. Analogously, it is possible that typicality is linked to a particular dimension or component as well. Along the same lines, it is also conceivable that the component or dimension comprising information regarding category structure differs across categories. In fact, a recent study by Jang and Myaeng (2017) provided evidence for this notion in the context of a non-predictive distributional semantic model. In particular, they sought to identify so-called significant properties for each category. That is, they collapsed across respective exemplar vectors as in the prototype-as-average approach (see above), and then selected the dimensions for which the mean score crossed a certain threshold. The value for each exemplar on those dimensions was subsequently correlated with the corresponding graded lexical entailment judgments of Vulic et al. (2017). Out of the five

---

<sup>4</sup>Here and throughout the paper, we fitted main-effects-only models featuring all predictors under consideration. Potential interactions were not taken into account as the number of observations per category was rather limited.

categories Jang and Myaeng (2017) tested (i.e., *animals*, *birds*, *food*, *fruit*, and *instruments*), the dimension with the highest average correlated strongly with lexical entailment (i.e., Pearson’s  $r$  of .55, .78, .26, .74, and .93, respectively). Hence, we examined whether we can obtain similarly encouraging results for the current predictive models and typicality data.

## Method

We took a slightly different approach as Jang and Myaeng (2017), in that we did not define an arbitrary threshold to select significant properties. Instead, we just used the dimension with the highest average value, which yielded, without exception, the best performance according to Jang and Myaeng’s results. For each category, we took the prototype-as-average vector from Experiment 1a, and selected the dimension with the top score. In contrast to Experiments 2a and 2b, we now again considered both the Dutch and English datasets. Also, semantic spaces and typicality ratings were those described under Experiment 1a.

Table 10 and 11 show the thus selected dimension for each category. As in Jang and Myaeng (2017), we see, especially for Dutch, that these dimensions tend to be shared by taxonomically related categories. In a final step, we correlated the exemplars’ score on that particular dimension with their typicality ratings.

## Results and discussion

The results are summarized in Figure 11 (Dutch) and 12 (English, with separate panels for younger and older adults). Overall, the correlations were very close to zero, especially for Dutch. In addition, there did not appear to be much consistency across semantic spaces. That is to say, providing a reasonable account of a certain category’s typicality ratings for one semantic space, did not indicate that the best-dimension approach performed well when using the other semantic space. In comparison, the approaches of previous experiments either (partly) succeeded or failed to predict typicality ratings for certain categories, regardless of which semantic space was considered.

These results stand in stark contrast to the ones observed by Jang and Myaeng (2017). While the discrepancy could in principle be attributed to the difference between predictive and non-predictive models, we believe that Jang and Myaeng’s correlations got inflated due to a combination of (some) relatively small category sizes (i.e., 16, 9, and 14 exemplars for *birds*, *fruit*, and *instruments*, respectively), and a number of idiosyncratic items. The category *birds* for instance, contained the clear non-members *hamburger*, *wing*, *feather*, and *kite*, which, for obvious reasons, received low lexical entailment ratings. So, if the best dimension allows to separate birds from non-birds, or even living from non-living things, it might correlate strongly with the lexical entailment ratings without necessarily conveying any information about within-category structure. A similar argument can be made for the category *instruments*. Most of its items were musical instruments, but a few (i.e., *compass*, *scalpel*, and *stethoscope*) did not fall in that category. As the latter items received the lowest entailment ratings, it suffices if the selected dimension somehow discriminates between musical and non-musical instruments to achieve a high correlation. Taken together, there is limited evidence that individual dimensions capture *within*-category structure, regardless of whether one relies on predictive or non-predictive distributed semantic models.

That being said, one might wonder how well non-predictive distributed semantic models would fair in capturing the typicality gradient of categories. Based on previous studies (Connell & Ramscar, 2001; Jones & Mewhort, 2007), one would expect that such models are able to predict typicality ratings. However, the question is how they would stack up against the, supposedly superior, predictive models. The goal of our next experiment is to shed some light on this issue.

### Experiment 3

In this experiment, we sought to directly compare the performance of a predictive model to that of a more traditional count model. To this end, we used the predict and count spaces from Mandera et al. (2017), as their models were trained on the exact same corpus,

thus facilitating comparability. From these spaces we derived all five predictor variables introduced in Experiment 1a, and Experiments 2a through 2c. So, for each semantic model, we obtained a set of five predictors, which were then entered into two separate multiple linear regression analyses. In order to assess their relative performance, we then contrasted the corresponding Akaike’s Information Criterion (henceforth, AIC) values.

## Method

The typicality ratings and procedure to derive the predictor variables were the same as before. Because the contrast category and instantiation predictors were not available for English, we again only considered the Dutch data. As for the semantic spaces, we opted for one predictive and one count model. The former was the CBOW model from previous experiments (Mandera et al., 2017), which achieved slightly superior performance compared to the skip-gram model. The count model was inspired by Lund and Burgess’ Hyperspace Analogue to Language (HAL; Lund & Burgess, 1996), as implemented by Mandera and colleagues. More specifically, they constructed a word by word co-occurrence matrix for the 300,000 most frequent words by sliding a window through the same corpus on which the predictive model was trained (i.e., a concatenation of subtitle and SONAR-500 corpora). Then, a positive pointwise mutual information weighting was applied to the resulting matrix (see Mandera et al., 2017 for more details).<sup>5</sup>

Based on those semantic vectors, we extracted scores for the following predictor variables: cosine similarity between the exemplar vectors and the average exemplar vector (as in Experiment 1a), cosine similarity between the exemplar vectors and the category label vector (as in Experiment 1a), cosine similarity between the exemplar vectors and the contrast category vector (as in Experiment 2a), cosine similarity between the exemplar vectors and

---

<sup>5</sup>Initially, we wanted to include an LSA-inspired model as well. Using the Gensim toolkit (Rehurek & Sojka, 2010), we obtained semantic vectors that appeared anomalous, though. Mandera et al. (2017) experienced similar issues, and therefore decided to focus only on the HAL-based model, which we did here as well.

the category’s best exemplar vector (as in Experiment 2b), and each exemplar’s value on the dimension with the highest average score across exemplars of a given category (as in Experiment 2c). These variables were then correlated with the typicality ratings, but only for the count model as the other correlations were already calculated in previous experiments (see Table 12 for an overview of the latter results). Note that for the contrast category variable, we again computed a partial correlation with typicality controlling for the cosine similarity between the exemplar vectors and the actual category label vector. However, there was one missing value for the category *kitchen utensils* as there was no semantic vector for the contrast category label (i.e., *poetsgerief* meaning *cleaning equipment*).

In addition to the correlational analyses, we also entered all five predictors simultaneously to a linear regression model for both the CBOW and HAL-based semantic spaces separately. The only exception was the HAL-based regression model for the category *kitchen utensils*, which contained four instead of five predictor variables due to the missing contrast category. Furthermore, we limited these analyses to the exemplars that were represented in both spaces, thus allowing for a straightforward evaluation. In order to assess which of the two (non-nested) models provided the best fit of human typicality ratings, we then compared their respective AIC values for each category.

## Results and discussion

The (partial) correlations for the HAL-type model are displayed in Table 13. As can be seen, the predictions for the prototype-as-average and prototype-as-category-label approaches are on average quite comparable to those of the CBOW model. If anything, the correlations were generally slightly higher (i.e., prototype-as-average .32 for HAL versus .29 for CBOW; prototype-as-category-label .28 for HAL versus .25 for CBOW). Despite the similarity in terms of average performance, there are marked differences if one looks at the individual categories. For instance, the HAL-based variables approximated the typicality gradient of the category *fruit* fairly well, whereas the CBOW-based variables were largely

unsuccessful in this respect, and vice versa for categories like *kitchen utensils*.

The contrast category variable for the count model did not have much predictive value overall (average partial correlation = -.07), but there was considerable variability across categories. In particular, the categories *musical instruments* and *professions* unexpectedly yielded substantial *positive* correlation coefficients. This finding calls into question whether the current implementation of similarity to the contrast category was sensible. Recall that the average correlation for the predictive models was rather low as well (see Experiment 2a). Alternatively, the influence of contrast categories on typicality ratings for common, pre-existing categories might be minimal to begin with, thus explaining the low overall correlation. We will revisit this issue in the General discussion.

The variable inspired by the instantiation principle yielded a moderate average correlation with typicality (i.e., .31). This figure is quite similar, albeit numerically smaller than that based on the CBOW model (i.e., .37). Moreover, both CBOW and HAL-based instantiation metrics seemed to achieve comparable results at the level of the individual categories. That is, *professions* and *kitchen utensils* were, for instance, both located at the lower end of the spectrum with small correlations between typicality and the instantiation predictor. Categories like *vehicles* and *musical instruments*, on the other hand, were positioned towards the upper end of the scale, though the HAL-based correlations were more homogeneous.

Finally, the best dimension predictor also correlated with typicality to a modest extent, at least on average (i.e., .30). This finding is the most pronounced difference with the results obtained for the CBOW model, where we observed an average correlation of only .09. However, it is important to point out that the respective semantic spaces differ in their dimensionality (i.e., 200 for the CBOW model versus 300,000 for the HAL-based model), and the interpretability of those dimensions. Still, the correlation coefficients fall short of those obtained by Jang and Myaeng (2017), thus further calling into question whether the latter findings truly speak to within-category structure.

The AIC scores for all multiple linear regression models are shown in Table 14, along with  $R^2$  and  $R^2_{adj}$  values. Somewhat surprisingly, we found that the count model outperformed the predictive model (i.e., having lower AIC values) on 10 out of 15 categories. In addition, the HAL-inspired model had a slight edge in terms of the average performance as well. Further inspection of Table 14 also revealed that both models tended to predict category structure to a similar degree. They, for instance, captured the typicality gradient of categories like *musical instruments*, *vehicles*, and *mammals* quite well, but were largely unsuccessful in other domains, most notably *insects* and *reptiles*. In sum, the predict model we evaluated in terms of accounting for human typicality ratings did not compare favourably to a HAL-based count model. This finding thus nuances the often-heard claim that predictive models are generally superior (see Baroni et al., 2014).

Hence, both predictive and non-predictive distributed semantic models face the challenge of closing the gap in terms of approaching actual human typicality judgments. We mention a number of avenues worthy of further exploration in the General discussion, but a last study is devoted to one particularly promising aspect. That is, we examined to what extent predictions could be improved by including variables that capture the accessibility of the concepts under consideration. Indeed, previous research has suggested that other factors such as familiarity, word frequency, and age of acquisition influence, or at least correlate to some degree, with typicality ratings (i.a., Malt & Smith, 1982; Schröder, Gemballa, Ruppín, & Wartenburger, 2012). Should these variables yield substantial predictive value on top of the model-based predictors from Experiments 1a-2c, then one would have a (partial) explanation of the rather disappointing performance thus far, as well as inspiration for future improvement<sup>6</sup>.

---

<sup>6</sup>We want to thank Marc Brysbaert for this suggestion.

### Experiment 4

Building on the results of Experiment 3, we aimed to address two questions: 1) to what degree, if at all, does the addition of the variables familiarity, word frequency, and age of acquisition, improve upon the previously observed goodness-of-fit estimates, and 2) do the model-based predictors from Experiments 1a-2c still contribute in their own right, after taking the three accessibility covariates into account? To assess the relative performance of the respective regression models we contrasted the resulting AIC values with those obtained in Experiment 3 (i.e., Table 14).

### Method

The employed procedure and stimulus set are completely analogous to those of Experiment 3, except for the addition of familiarity, age of acquisition, and word frequency estimates, taken from the norms of De Deyne et al. (2008). The former two variables were based on subjective ratings, whereas the word frequencies were actually the log-transformed lemma counts from the CELEX database (Baayen, Piepenbrock, & van Rijn, 1993). These three predictors were then added to the linear regression models described under Experiment 3 and the resulting AIC values were contrasted. Subsequently, we conducted a regression analysis featuring only the accessibility covariates, and compared the AIC values with those from the eight-predictor analysis in order to determine the role of the model-based predictors. Note that the age of acquisition estimate for one exemplar of the category *kitchen utensils* was missing, which hinders a straightforward evaluation for that category.

### Results and discussion

The results of the eight-predictor analyses (i.e., five derived from the semantic models plus the three covariates) are displayed in Table 15. The addition of familiarity, word frequency, and age of acquisition yielded, on average, a sizeable increase in terms of the proportion of explained variance (i.e.,  $\Delta R^2 = .19$ ,  $\Delta R^2_{adj} = .15$ , for the CBOW model, and



$\Delta R^2 = .18$ ,  $\Delta R^2_{adj} = .14$ , for the HAL-based model). For certain categories, most notably *fruit* and *musical instruments*, it actually provided a very satisfying account of human typicality ratings. Furthermore, based on the AIC scores, the eight-predictor model was preferred over the model without the three accessibility covariates in 12 (CBOW) and 10 (HAL) out of the 15 categories.

The conclusion regarding the comparison between the CBOW and HAL-inspired model remained the same. The latter was favored in 10 out of 15 categories, and it also had a slight advantage when considering the average performance.

Finally, Table 16 shows the results for a regression model including only the three covariates. When contrasted with Table 14, it becomes clear that the variables derived from the semantic spaces contributed substantially to the prediction of typicality ratings. In particular, the proportion of explained variance, averaged across categories, nearly doubled after adding the five model-based predictors introduced in Experiments 1a-2c ( $\Delta R^2 = .28$ ,  $\Delta R^2_{adj} = .19$ , for the CBOW model, and  $\Delta R^2 = .29$ ,  $\Delta R^2_{adj} = .21$ , for the HAL-based model). Moreover, the regression model featuring those predictors in addition to the three covariates was preferred over the model with only the three covariates in 12 out of the 15 categories (for both CBOW and HAL).

Taken together, from Experiment 3, we learned that a) the five predictors seeking to capture category structure based on the semantic spaces were able to predict human typicality ratings to some degree, and b) if anything, the CBOW model was outperformed by the HAL-based model in this respect. The current experiment confirmed and extended those findings. However, the present results also qualified the conclusion that the so-called predictive distributed semantic models did not live up to the lofty expectations. That is to say, their performance could potentially be ameliorated if one would succeed in extracting information such as familiarity, word frequency, and age of acquisition from the semantic spaces. In fact, the study by Hollis and Westbury (2016) implies that this is actually possible to some extent, at least for word frequency and age of acquisition. So future work

could explore this promising avenue, yet, given our results, it would probably still fail to mimic human performance for certain categories, most notably *insects* and *kitchen utensils* (see Table 15).

### General discussion

This study examined how well predictive distributed semantic models fare in recovering the typicality gradient of common categories (e.g., *fruit*, *birds*, *vehicles*,...). Previous work suggested that these kind of models are especially well-equipped to mimic human behavior (Baroni et al., 2014), perhaps because the underlying learning mechanisms are psychologically plausible (Mandera et al., 2017). Given that past validation efforts predominantly focused on predicting similarity-driven variables such as relatedness ratings, semantic priming effects, and word associations, it was unclear whether predictive models also capture the structure of broader categories.

To address that question, we conducted a first experiment in which we derived prototypes from two model-based word embeddings, one with CBOW and one with skip-gram architecture, and this for two languages, English and Dutch. Prototypes were either obtained by averaging across exemplar vectors (e.g., *robin*, *dove*, *sparrow*,...; prototype-as-average) or by merely using the vector corresponding to the category label itself (e.g., *bird*; prototype-as-category-label). The cosine similarity between each exemplar vector and the two respective prototype vectors was then related to human typicality ratings. On the whole, the analyses revealed relatively small correlations. Looking at the average across categories, only one combination yielded an estimate that would cross Cohen’s moderate effect size threshold of .30.

In a first follow-up experiment, we examined whether the relatively modest performance was due to the focus on single nouns. In Experiment 1b, we used the semantic vectors corresponding to the plural forms instead, but this approach had little impact on the overall quality of the typicality predictions. Alternatively, one could potentially attribute the

low correlations to the nature of the predictor variables. Similarity to the prototype of the particular category might just be weakly related to typicality in general (though see e.g., Voorspoels et al., 2008). Experiments 2a-2c addressed that possibility by testing out different predictor variables: (dis)similarity to a contrast category (Experiment 2a), similarity to the best exemplar of the category in question (Experiment 2b), and the respective scores on the dimension with the highest average value across exemplars of a certain category (Experiment 2c, see Jang & Myaeng, 2017). Neither variable provided a satisfying account of the typicality ratings, although the exemplar-based predictor, inspired by the instantiation principle of Heit and Barsalou (1996), offered a slight improvement over the prototype approaches introduced in Experiments 1a and 1b. Nevertheless, as predictive models are often considered the proverbial new kid on the block, the achieved performance was a bit disappointing, especially in light of the relative success of other (older) distributional semantic models (Connell & Ramscar, 2001; Jones & Mewhort, 2007).

In order to directly compare such so-called count models with predictive models, we conducted an additional experiment. Somewhat surprisingly, Experiment 3 found that a HAL-based count model (Lund & Burgess, 1996), was generally superior to a CBOW model trained on the same corpus, which had yielded the best overall performance up to that point. That being said, the difference was rather small, and potentially coincidental. It does suggest that predictive distributed semantic models do not necessarily offer an advantage over count models in terms of mirroring human behavior.

Interestingly, our findings mesh well with those of Vulic et al. (2017) on graded lexical entailment, of which we had no knowledge when initiating the current project. Even though we took a rather different approach (i.e., other data collection procedure, stimuli, word embeddings, and independent variables), our conclusions are remarkably similar: there is a substantial gap between what state-of-the-art distributed semantic models predict about category structure, and actual human data.

To be clear, we do not imply that these models fail to recover categories' typicality

gradient. In most domains, they did succeed in explaining at least a small amount of variance, and in some cases, performance was actually quite good (see Table 14). It does beg the question of why predictive models were not more successful overall. In a final experiment, we sought to break some ground in this respect by examining whether familiarity, word frequency, and age of acquisition would contribute to the prediction of typicality ratings when controlling for the five variables introduced in Experiments 1a-2c. The three covariates indeed turned out to provide substantial, additional predictive value. Hence, one way to close the gap with actual human performance would be to somehow extract that information from the semantic spaces. Interestingly, Hollis and Westbury (2016) showed that certain principal components derived from a skip-gram representation correlated moderately with word frequency and age of acquisition, thus there seems to be some potential in that regard.

That being said, one would also need to carefully consider the semantic spaces and how they are constructed, the (psychological) theories on semantic memory, and the operationalization of the respective theoretical constructs, to provide a truly satisfying account of human typicality judgments. Isolating the root cause(s) of the poor performance for some categories might not be a straightforward endeavour, though. Based on the current results it would, for example, be tempting to dismiss the importance of contrast categories in shaping the structure of semantic memory. However, it could also be the case that the corresponding semantic vectors are not reliable, or that one should consider alternative implementations (e.g., relying on the contrast categories' best exemplar, or the average over many exemplars, perhaps using an a priori weighting scheme, . . . ).

Furthermore, if one would lay the blame (in part) on the semantic spaces, one still needs to determine whether the model's architecture itself is really the culprit. Alternatively, one might envision that training these models purely on text corpora does not suffice to fully capture category structure. Perceptual information arguably plays a big role in this respect, and some recent studies have begun to train models on both linguistic and visual input (e.g., Kiela & Bottou, 2014; Lazaridou, Pham, & Baroni, 2015). The current experiments could

thus serve as benchmarks against which future studies could be compared.

In sum, the present study showed that prediction-based distributional semantic models trained on text corpora can predict typicality ratings to some extent. However, contrary to results in other domains (Baroni et al., 2014), they do not appear to have an edge over count-based semantic models. It is possible that such predictive models' representations simply do not mimic human semantic memory, which is not inherently problematic because psychological plausibility was no principal goal in itself. A more nuanced (and plausible) explanation is that a) the semantic spaces are still not entirely on point, b) the information regarding category structure is encoded in a more opaque, yet to be discovered fashion, and c) typicality ratings are likely influenced by other factors like familiarity, word frequency, and/or age of acquisition, which could potentially be derived from the semantic spaces in their own right.

### **Author contributions**

TH and GH developed the study concept. TH performed the data analysis and interpretation. TH drafted the manuscript, with GH providing critical revisions. Both authors approved the final version for submission.

### **Acknowledgments**

TH is a postdoctoral fellow of the Research Foundation-Flanders (FWO-Vlaanderen). We want to thank Steven Verheyen and Wouter Voorspoels for providing valuable suggestions. Correspondence should be addressed to Tom Heyman, Department of Experimental Psychology, University of Leuven, Tiensestraat 102, 3000 Leuven, Belgium. E-mail: [tom.heyman@kuleuven.be](mailto:tom.heyman@kuleuven.be)

## References

- Allaire, J., Cheng, J., Xie, Y., McPherson, J., Chang, W., Allen, J., . . . Hyndman, R. (2016). *rmarkdown: Dynamic Documents for R*. Retrieved from <https://CRAN.R-project.org/package=rmarkdown>
- Ameel, E., & Storms, G. (2006). From prototypes to caricatures: Geometrical models for concept typicality. *Journal of Memory and Language*, 55(3), 402–421. doi:[10.1016/j.jml.2006.05.005](https://doi.org/10.1016/j.jml.2006.05.005)
- Aust, F., & Barth, M. (2017). *papaja: Create APA manuscripts with R Markdown*. Retrieved from <https://github.com/crsh/papaja>
- Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). The CELEX lexical database. [CD-ROM]; Philadelphia: University of Pennsylvania, Linguistic Data Consortium.
- Baroni, M., Dinu, G., & Kruszewski, G. (2014). Don't count, predict! A systematic comparison of context-counting vs. context-predicting semantic vectors. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics* (pp. 238–247).
- Connell, L., & Ramscar, M. (2001). Using distributional measures to model typicality in categorization. In *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (pp. 226–231). Mahwah, NJ: Erlbaum.
- De Deyne, S., Verheyen, S., Ameel, E., Vanpaemel, W., Dry, M. J., Voorspoels, W., & Storms, G. (2008). Exemplar by feature applicability matrices and other Dutch normative data for semantic concepts. *Behavior Research Methods*, 40(4), 1030–1048. doi:[10.3758/brm.40.4.1030](https://doi.org/10.3758/brm.40.4.1030)
- Dry, M. J., & Storms, G. (2010). Features of graded category structure: Generalizing the family resemblance and polymorphous concept models. *Acta Psychologica*, 133(3), 244–255. doi:[10.1016/j.actpsy.2009.12.005](https://doi.org/10.1016/j.actpsy.2009.12.005)
- Harris, Z. S. (1954). Distributional structure. *Word*, 10(2-3), 146–162.

doi:[10.1080/00437956.1954.11659520](https://doi.org/10.1080/00437956.1954.11659520)

Heit, E., & Barsalou, L. W. (1996). The instantiation principle in natural categories.

*Memory*, 4(4), 413–451. doi:[10.1080/096582196388915](https://doi.org/10.1080/096582196388915)

Hollis, G. (2017). Estimating the average need of semantic knowledge from distributional semantic models. *Memory & Cognition*, 45(8), 1350–1370.

doi:[10.3758/s13421-017-0732-1](https://doi.org/10.3758/s13421-017-0732-1)

Hollis, G., & Westbury, C. (2016). The principals of meaning: Extracting semantic dimensions from co-occurrence models of semantics. *Psychonomic Bulletin & Review*, 23(6), 1744–1756. doi:[10.3758/s13423-016-1053-2](https://doi.org/10.3758/s13423-016-1053-2)

Jang, K.-R., & Myaeng, S.-H. (2017). Elucidating conceptual properties from word embeddings. In *SENSE 2017* (pp. 91–95).

Jones, M. N., & Mewhort, D. J. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological Review*, 114(1), 1–37.

doi:[10.1037/0033-295x.114.1.1](https://doi.org/10.1037/0033-295x.114.1.1)

Kiela, D., & Bottou, L. (2014). Learning image embeddings using convolutional neural networks for improved multi-modal semantics. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing* (pp. 36–45).

Lazaridou, A., Pham, N. T., & Baroni, M. (2015). Combining language and vision with a multimodal skip-gram model. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 153–163).

Lenci, A. (2018). Distributional models of word meaning. *Annual Review of Linguistics*, 4(1). doi:[10.1146/annurev-linguistics-030514-125254](https://doi.org/10.1146/annurev-linguistics-030514-125254)

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers*, 28(2), 203–208. doi:[10.3758/bf03204766](https://doi.org/10.3758/bf03204766)

Malt, B. C., & Johnson, E. C. (1992). Do artifact concepts have cores? *Journal of Memory*

- and Language*, 31(2), 195–217. doi:[10.1016/0749-596x\(92\)90011-1](https://doi.org/10.1016/0749-596x(92)90011-1)
- Malt, B. C., & Smith, E. E. (1982). The role of familiarity in determining typicality. *Memory & Cognition*, 10(1), 69–75. doi:[10.3758/bf03197627](https://doi.org/10.3758/bf03197627)
- Mandera, P., Keuleers, E., & Brysbaert, M. (2017). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language*, 92, 57–78. doi:[10.1016/j.jml.2016.04.001](https://doi.org/10.1016/j.jml.2016.04.001)
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems* (pp. 3111–3119).
- Morrow, L. I., & Duffy, M. F. (2005). The representation of ontological category concepts as affected by healthy aging: Normative data and theoretical implications. *Behavior Research Methods*, 37(4), 608–625. doi:[10.3758/bf03192731](https://doi.org/10.3758/bf03192731)
- Pereira, F., Gershman, S., Ritter, S., & Botvinick, M. (2016). A comparative evaluation of off-the-shelf distributed semantic representations for modelling behavioural data. *Cognitive Neuropsychology*, 33(3-4), 175–190. doi:[10.1080/02643294.2016.1176907](https://doi.org/10.1080/02643294.2016.1176907)
- Pierrejean, B., & Tanguy, L. (2018). Predicting word embeddings variability. In *Proceedings of the Seventh Joint Conference on Lexical and Computational Semantics* (pp. 154–159).
- R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Rehurek, R., & Sojka, P. (2010). Software framework for topic modelling with large corpora. In *In Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks* (pp. 45–50). Valletta, Malta: ELRA.
- Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental*



- Psychology: General*, 104(3), 192–233. doi:[10.1037/0096-3445.104.3.192](https://doi.org/10.1037/0096-3445.104.3.192)
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7(4), 573–605. doi:[10.1016/0010-0285\(75\)90024-9](https://doi.org/10.1016/0010-0285(75)90024-9)
- Schröder, A., Gemballa, T., Ruppín, S., & Wartenburger, I. (2012). German norms for semantic typicality, age of acquisition, and concept familiarity. *Behavior Research Methods*, 44(2), 380–394. doi:[10.3758/s13428-011-0164-y](https://doi.org/10.3758/s13428-011-0164-y)
- Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1), 72–101. doi:[10.2307/1412159](https://doi.org/10.2307/1412159)
- Storms, G., De Boeck, P., & Ruts, W. (2000). Prototype and exemplar-based information in natural language categories. *Journal of Memory and Language*, 42(1), 51–73. doi:[10.1006/jmla.1999.2669](https://doi.org/10.1006/jmla.1999.2669)
- Tulkens, S., Emmery, C., & Daelemans, W. (2016). Evaluating unsupervised Dutch word embeddings as a linguistic resource. In *Proceedings of the 10th international conference on language resources and evaluation (LREC)*. European Language Resources Association (ELRA).
- Verheyen, S., & Stukken, L. (2010). [Contrast categories]. Unpublished raw data.
- Voorspoels, W., Storms, G., & Vanpaemel, W. (2012). Contrast effects in typicality judgements: A hierarchical Bayesian approach. *The Quarterly Journal of Experimental Psychology*, 65(9), 1721–1739. doi:[10.1080/17470218.2012.662237](https://doi.org/10.1080/17470218.2012.662237)
- Voorspoels, W., Vanpaemel, W., & Storms, G. (2008). Exemplars and prototypes in natural language concepts: A typicality-based evaluation. *Psychonomic Bulletin & Review*, 15(3), 630–637. doi:[10.3758/pbr.15.3.630](https://doi.org/10.3758/pbr.15.3.630)
- Vulic, I., Gerz, D., Kiela, D., Hill, F., & Korhonen, A. (2017). Hyperlex: A large-scale evaluation of graded lexical entailment. *Computational Linguistics*, 43(4), 781–835. doi:[10.1162/COLI\\_a\\_00301](https://doi.org/10.1162/COLI_a_00301)
- Wendlandt, L., Kummerfeld, J. K., & Mihalcea, R. (2018). Factors influencing the surprising instability of word embeddings. In *Proceedings of the 2018 Conference of the North*

*American Chapter of the Association for Computational Linguistics* (Vol. 1, pp. 2092–2102).

Yee, E., Jones, M. N., & McRae, K. (2018). The Stevens' handbook of experimental psychology and cognitive neuroscience. In J. T. Wixted & S. Thompson-Schill (Eds.), (4th Edition., Vol. 3: Language and thought). New York: Wiley.

Table 1

*Summary statistics of the cosine similarity between every exemplar vector and the corresponding prototype vector for the Dutch embeddings using the prototype-as-average approach*

Category	CBOW		skip-gram	
	$M$ ( $SD$ )	Range	$M$ ( $SD$ )	Range
birds	.56 (0.10)	[.40, .75]	.60 (0.09)	[.38, .77]
clothes	.62 (0.13)	[.28, .77]	.59 (0.17)	[.17, .83]
fish	.66 (0.08)	[.49, .75]	.62 (0.13)	[.35, .81]
fruit	.65 (0.11)	[.34, .80]	.71 (0.14)	[.16, .86]
insects	.53 (0.13)	[.25, .76]	.53 (0.20)	[.09, .76]
kitchen utensils	.55 (0.11)	[.33, .73]	.54 (0.15)	[.17, .72]
mammals	.60 (0.09)	[.35, .74]	.62 (0.12)	[.16, .78]
musical instruments	.69 (0.09)	[.45, .80]	.80 (0.11)	[.50, .93]
professions	.45 (0.08)	[.19, .58]	.56 (0.11)	[.31, .70]
reptiles	.57 (0.12)	[.35, .76]	.65 (0.18)	[.27, .82]
sports	.54 (0.13)	[.24, .69]	.65 (0.18)	[.23, .89]
tools	.52 (0.17)	[.03, .76]	.57 (0.15)	[.10, .76]
vegetables	.68 (0.13)	[.34, .82]	.74 (0.15)	[.20, .88]
vehicles	.52 (0.10)	[.23, .75]	.56 (0.13)	[.26, .80]
weapons	.55 (0.13)	[.09, .72]	.60 (0.08)	[.45, .74]

Table 2

*Summary statistics of the cosine similarity between every exemplar vector and the corresponding prototype vector for the Dutch embeddings using the prototype-as-category-label approach*

Category	CBOW		skip-gram	
	$M$ ( $SD$ )	Range	$M$ ( $SD$ )	Range
birds	.39 (0.11)	[.20, .55]	.44 (0.12)	[.15, .61]
clothes	.44 (0.11)	[.17, .63]	.38 (0.16)	[-.04, .62]
fish	.52 (0.09)	[.29, .63]	.43 (0.14)	[.14, .65]
fruit	.33 (0.07)	[.15, .47]	.51 (0.15)	[.06, .75]
insects	.39 (0.13)	[.10, .57]	.10 (0.07)	[-.02, .23]
kitchen utensils	.26 (0.10)	[.08, .47]	.32 (0.14)	[.04, .55]
mammals	.34 (0.11)	[.14, .54]	.28 (0.12)	[-.02, .70]
musical instruments	.54 (0.08)	[.33, .65]	.56 (0.05)	[.43, .67]
professions	.17 (0.10)	[.01, .42]	.20 (0.12)	[.05, .63]
reptiles	.46 (0.12)	[.23, .62]	.36 (0.12)	[.10, .57]
sports	.38 (0.12)	[.14, .55]	.39 (0.14)	[.13, .60]
tools	.33 (0.13)	[.01, .56]	.38 (0.14)	[-.04, .59]
vegetables	.43 (0.10)	[.15, .56]	.60 (0.16)	[.14, .76]
vehicles	.31 (0.11)	[.05, .56]	.39 (0.14)	[.08, .63]
weapons	.41 (0.17)	[.13, .83]	.28 (0.11)	[.13, .61]

Table 3

*Summary statistics of the cosine similarity between every exemplar vector and the corresponding prototype vector for the English embeddings using the prototype-as-average approach*

Category	CBOW		skip-gram	
	$M$ ( $SD$ )	Range	$M$ ( $SD$ )	Range
animals	.49 (0.10)	[.16, .72]	.58 (0.11)	[.06, .78]
birds	.51 (0.11)	[.16, .70]	.59 (0.14)	[.08, .78]
clothes	.55 (0.13)	[.15, .77]	.58 (0.14)	[.12, .76]
flowers	.58 (0.11)	[.36, .81]	.68 (0.11)	[.20, .82]
fruit	.55 (0.13)	[.23, .75]	.65 (0.12)	[.06, .82]
furniture	.50 (0.13)	[.19, .75]	.52 (0.17)	[.13, .82]
insects	.51 (0.10)	[.22, .65]	.59 (0.13)	[.23, .80]
musical instruments	.60 (0.13)	[.14, .75]	.67 (0.14)	[.22, .86]
tools	.47 (0.12)	[.15, .68]	.48 (0.15)	[.09, .73]
vegetables	.58 (0.11)	[.36, .76]	.69 (0.11)	[.40, .83]
vehicles	.47 (0.08)	[.21, .65]	.52 (0.13)	[.13, .74]

Table 4

*Summary statistics of the cosine similarity between every exemplar vector and the corresponding prototype vector for the English embeddings using the prototype-as-category-label approach*

Category	CBOW		skip-gram	
	$M$ ( $SD$ )	Range	$M$ ( $SD$ )	Range
animals	.31 (0.09)	[.06, .49]	.38 (0.10)	[.04, .64]
birds	.37 (0.09)	[.04, .59]	.48 (0.13)	[.01, .68]
clothes	.37 (0.12)	[.10, .66]	.36 (0.11)	[.06, .65]
flowers	.41 (0.10)	[.19, .60]	.52 (0.10)	[.13, .69]
fruit	.34 (0.11)	[.06, .52]	.47 (0.11)	[.08, .66]
furniture	.27 (0.10)	[.06, .47]	.29 (0.12)	[.06, .50]
insects	.36 (0.11)	[.08, .54]	.49 (0.16)	[.09, .77]
musical instruments	-	-	-	-
tools	.23 (0.09)	[.00, .45]	.17 (0.07)	[-.01, .30]
vegetables	.29 (0.08)	[.14, .44]	.47 (0.10)	[.24, .63]
vehicles	.26 (0.11)	[.08, .65]	.33 (0.15)	[.02, .78]

Table 5

*Reliability estimates for the similarity-to-the-prototype metrics and the typicality ratings for Dutch*

Category	Prototype-as-average	Prototype-as-category-label	Typicality
birds	.72	.61	.97
clothes	.66	.59	.98
fish	.52	.54	.96
fruit	.70	.55	.96
insects	.53	.43	.87
kitchen utensils	.30	.31	.91
mammals	.42	.73	.94
musical instruments	.62	.23	.94
professions	.81	.64	.92
reptiles	.47	.70	.90
sports	.70	.73	.98
tools	.69	.43	.95
vegetables	.55	.52	.94
vehicles	.69	.59	.98
weapons	.54	.40	.98

Table 6

*Reliability estimates for the similarity-to-the-prototype metrics and the typicality ratings for English*

Category	Prototype-as-average	Prototype-as-category-label	Typicality
animals	.64	.69	.93
birds	.38	.14	.87
clothes	.63	.67	.83
flowers	.44	.30	.70
fruit	.49	.57	.87
furniture	.74	.81	.76
insects	.71	.77	.88
musical instruments	.86	-	.88
tools	.71	.54	.79
vegetables	.73	.58	.89
vehicles	.72	.80	.90



Table 7

*Most frequently generated contrast category for each category in the norms of De Deyne et al. (2008)*

Category	Contrast category
birds	mammals
clothes	accessories
fish	mammals
fruit	vegetables
insects	birds
kitchen utensils	cleaning equipment
mammals	amfibians
musical instruments	radios
professions	hobbies
reptiles	amfibians
sports	hobbies
tools	vehicles
vegetables	fruit
vehicles	tools
weapons	tools

Table 8

*Most frequently generated exemplar for each category in the norms of De Deyne et al. (2008)*

Category	Best exemplar
birds	sparrow
clothes	pants
fish	salmon
fruit	apple
insects	fly
kitchen utensils	knife
mammals	dog
musical instruments	guitar
professions	psychologist
reptiles	snake
sports	soccer
tools	hammer
vegetables	lettuce
vehicles	car
weapons	rifle

Table 9

*Goodness-of-fit measures for the multiple linear regression model with all four predictors from Experiments 1a, 2a, and 2b, for each category and architecture separately*

Category	CBOW		skip-gram	
	$R^2$	$R^2_{adj}$	$R^2$	$R^2_{adj}$
birds	.27	.15	.51	.43
clothes	.16	.01	.15	-.01
fish	.64	.56	.59	.50
fruit	.24	.12	.35	.25
insects	.16	.00	.18	.03
kitchen utensils	.34	.25	.02	-.09
mammals	.60	.54	.77	.74
musical instruments	.56	.48	.34	.22
professions	.54	.47	.43	.34
reptiles	.15	-.08	.11	-.13
sports	.40	.30	.39	.29
tools	.25	.12	.09	-.08
vegetables	.28	.15	.18	.03
vehicles	.63	.57	.64	.58
weapons	.47	.33	.46	.32
$M$	.38	.26	.35	.23

Table 10

*The dimension with the highest average value across all exemplars of each separate category for Dutch*

Category	CBOW	skip-gram
birds	D27	D272
clothes	D138	D106
fish	D27	D272
fruit	D103	D272
insects	D27	D272
kitchen utensils	D114	D271
mammals	D27	D270
musical instruments	D171	D52
professions	D160	D72
reptiles	D27	D272
sports	D171	D257
tools	D171	D271
vegetables	D103	D272
vehicles	D114	D270
weapons	D171	D271

Table 11

*The dimension with the highest average value across all exemplars of each separate category for English*

Category	CBOW	skip-gram
animals	D199	D30
birds	D244	D54
clothes	D100	D108
flowers	D268	D291
fruit	D268	D89
furniture	D293	D72
insects	D281	D48
musical instruments	D177	D147
tools	D225	D44
vegetables	D216	D291
vehicles	D146	D224

Table 12

*Correlation between typicality and the various predictors derived from the CBOW model*

Category	Prototype-as-average	Prototype-as-category-label	Contrast category <sup>a</sup>	Instantiation	Best dimension
birds	.29	.39	-.23	.43	.07
clothes	.34	.33	-.16	.39	.13
fish	.02	.18	-.75	.50	-.22
fruit	.03	.16	-.32	.39	.34
insects	.02	-.02	.39	.25	-.05
kitchen utensils	.57	.39	.03	.17	.46
mammals	.01	-.46	-.20	.44	-.21
musical instruments	.34	.07	-.03	.49	.28
professions	.68	.22	.12	.07	.13
reptiles	.32	.28	-.14	.20	.33
sports	.46	.47	-.42	.29	-.20
tools	.44	.38	-.10	.46	.30
vegetables	.08	.29	-.06	.31	-.32
vehicles	.61	.62	-.46	.73	.20
weapons	.17	.52	-.51	.50	.07
<i>M</i>	.29	.25	-.19	.37	.09

<sup>a</sup> The values in this column are partial correlations (see main text for details).

Table 13

*Correlation between typicality and the various predictors derived from the HAL-type model*

Category	Prototype-as-average	Prototype-as-category-label	Contrast category <sup>a</sup>	Instantiation	Best dimension
birds	-.35	-.03	.03	.30	.26
clothes	.37	.40	.09	.37	.55
fish	.29	.33	-.69	.36	.57
fruit	.53	.46	-.31	.39	.53
insects	.05	-.04	-.06	.28	.08
kitchen utensils	.12	.20	-	.07	.11
mammals	.79	-.23	-.25	.43	.15
musical instruments	.67	.20	.74	.45	.59
professions	.15	.16	.45	-.08	.43
reptiles	.35	.25	-.01	.21	-.11
sports	.66	.62	-.01	.36	.24
tools	.26	.41	-.17	.36	.38
vegetables	.32	.48	-.25	.32	.20
vehicles	.64	.66	-.25	.46	.61
weapons	-.08	.32	-.25	.35	-.09
<i>M</i>	.32	.28	-.07	.31	.30

<sup>a</sup> The values in this column are partial correlations (see main text for details).

Table 14

*Goodness-of-fit measures for the multiple linear regression model with all five predictors, for each category and architecture separately*

Category	CBOW			HAL		
	AIC	$R^2$	$R^2_{adj}$	AIC	$R^2$	$R^2_{adj}$
birds	140.70	.28	.12	131.63	.47	.35
clothes	139.31	.17	-.03	131.49	.38	.23
fish	101.74	.67	.57	106.54	.59	.47
fruit	136.42	.32	.18	131.17	.43	.31
insects	119.77	.18	-.03	118.88	.21	.01
kitchen utensils	135.20	.39	.27	147.46	.06	-.08
mammals	102.42	.60	.52	78.84	.82	.78
musical instruments	87.54	.64	.56	86.19	.66	.58
professions	97.31	.57	.48	110.72	.32	.18
reptiles	107.06	.19	-.10	107.89	.16	-.15
sports	144.58	.40	.27	140.52	.48	.37
tools	136.19	.20	.01	131.31	.33	.18
vegetables	127.15	.39	.25	125.22	.44	.30
vehicles	152.27	.63	.55	150.99	.64	.57
weapons	114.45	.48	.29	115.91	.44	.23
$M$	122.81	.41	.26	120.98	.43	.29



Table 15

*Goodness-of-fit measures for the multiple linear regression model with all five predictors as well as the covariates word frequency, age of acquisition, and familiarity, for each category and architecture separately*

Category	CBOW			HAL		
	AIC	$R^2$	$R^2_{adj}$	AIC	$R^2$	$R^2_{adj}$
birds	139.79	.43	.20	132.10	.56	.39
clothes	138.49	.35	.07	133.44	.46	.23
fish	105.64	.69	.52	105.46	.70	.52
fruit	89.32	.88	.84	69.88	.94	.92
insects	121.72	.30	-.03	122.30	.28	-.06
kitchen utensils	133.91	.46	.27	146.86	.14	-.11
mammals	95.93	.74	.64	70.46	.89	.84
musical instruments	60.18	.90	.85	55.41	.91	.88
professions	98.99	.63	.48	105.27	.54	.36
reptiles	98.53	.61	.32	96.61	.64	.39
sports	141.22	.57	.40	143.20	.54	.35
tools	133.84	.41	.15	126.46	.55	.36
vegetables	126.50	.53	.32	130.43	.45	.21
vehicles	148.65	.73	.63	145.01	.76	.67
weapons	106.50	.74	.55	106.26	.74	.55
$M$	115.95	.60	.41	112.61	.61	.43

Table 16

*Goodness-of-fit measures for the multiple linear regression model with only the covariates word frequency, age of acquisition, and familiarity, for each category separately*

Category	AIC	$R^2$	$R^2_{adj}$
birds	143.59	.08	-.03
clothes	131.79	.27	.17
fish	117.07	.22	.10
fruit	92.23	.82	.80
insects	117.97	.11	-.01
kitchen utensils	141.82	.05	-.05
mammals	103.00	.54	.48
musical instruments	74.67	.74	.71
professions	107.41	.31	.23
reptiles	97.98	.37	.25
sports	147.10	.25	.16
tools	136.78	.05	-.07
vegetables	132.30	.15	.04
vehicles	149.92	.61	.56
weapons	121.04	.11	-.06
$M$	120.98	.31	.22

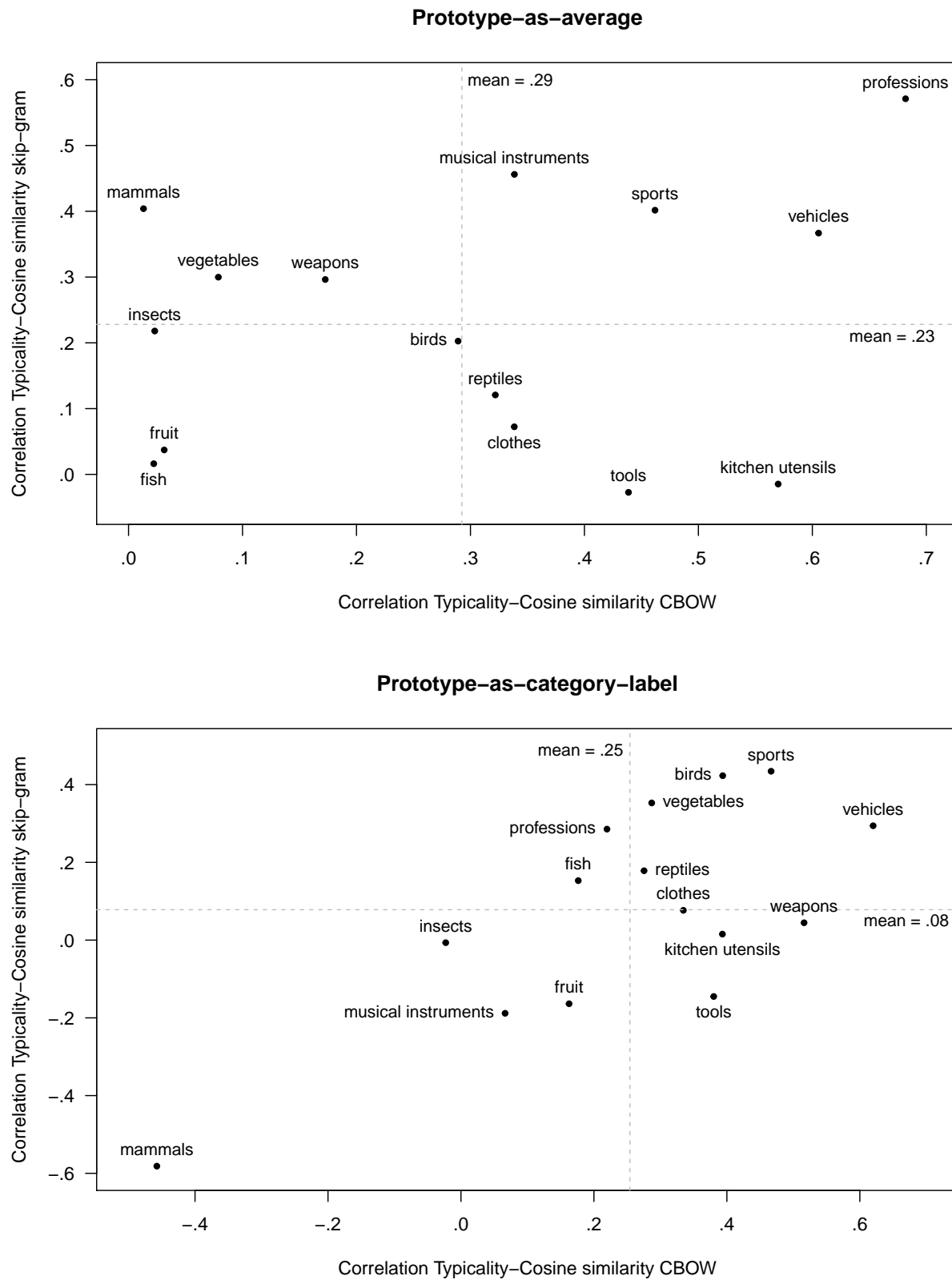


Figure 1. Summary of the results of Experiment 1a for Dutch, see main text for details

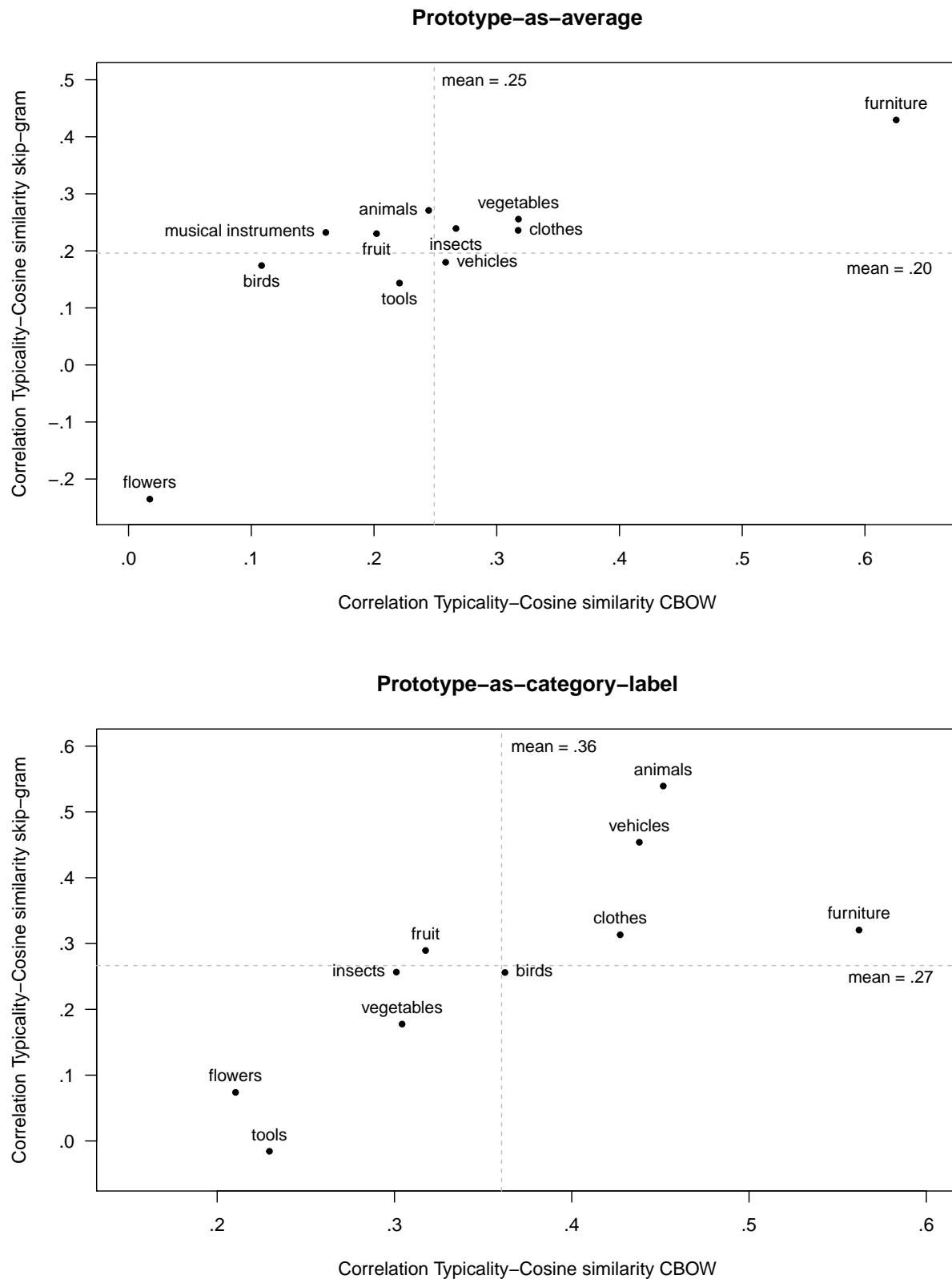


Figure 2. Summary of the results of Experiment 1a for English younger adults, see main text for details

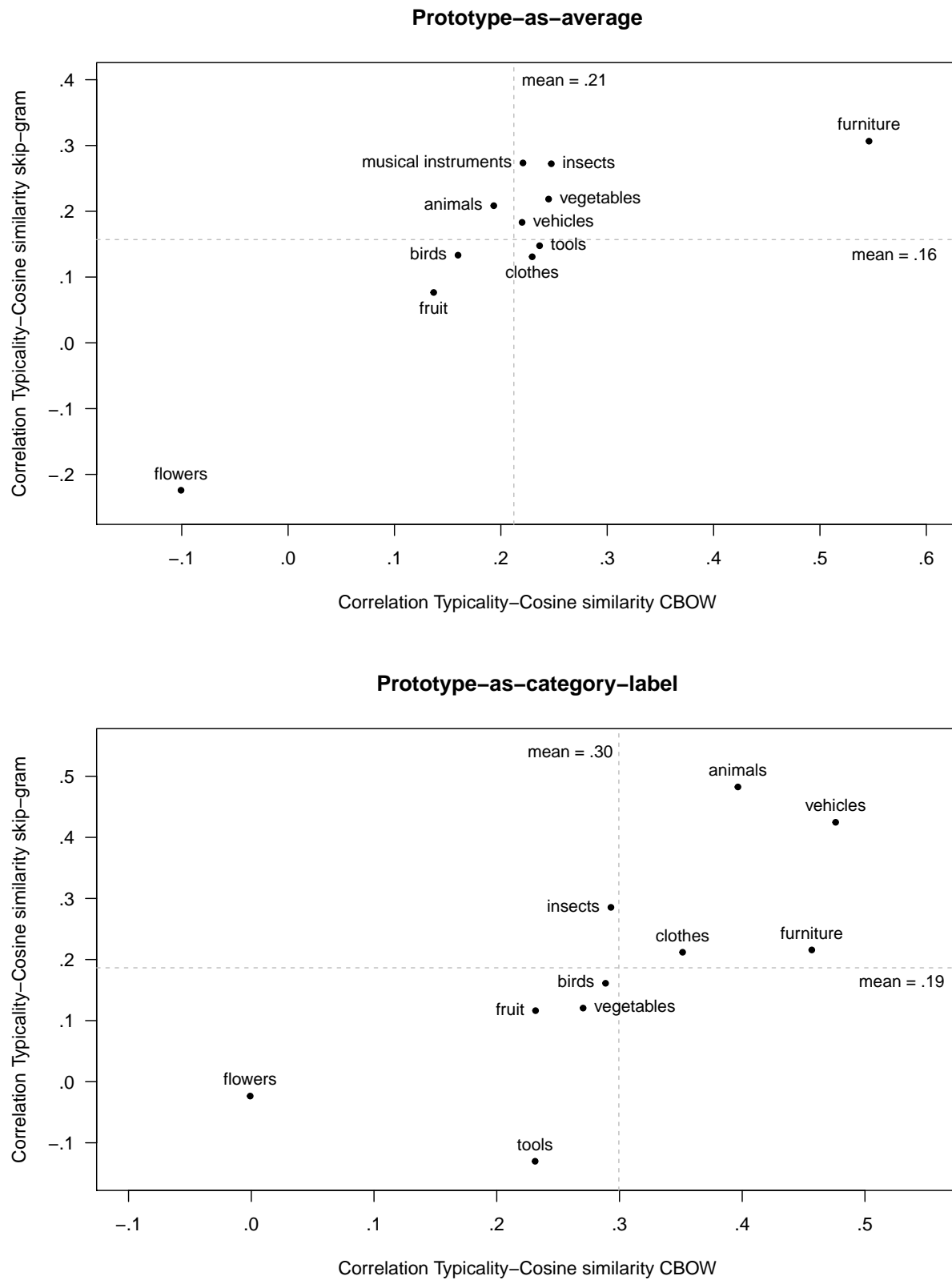


Figure 3. Summary of the results of Experiment 1a for English older adults, see main text for details

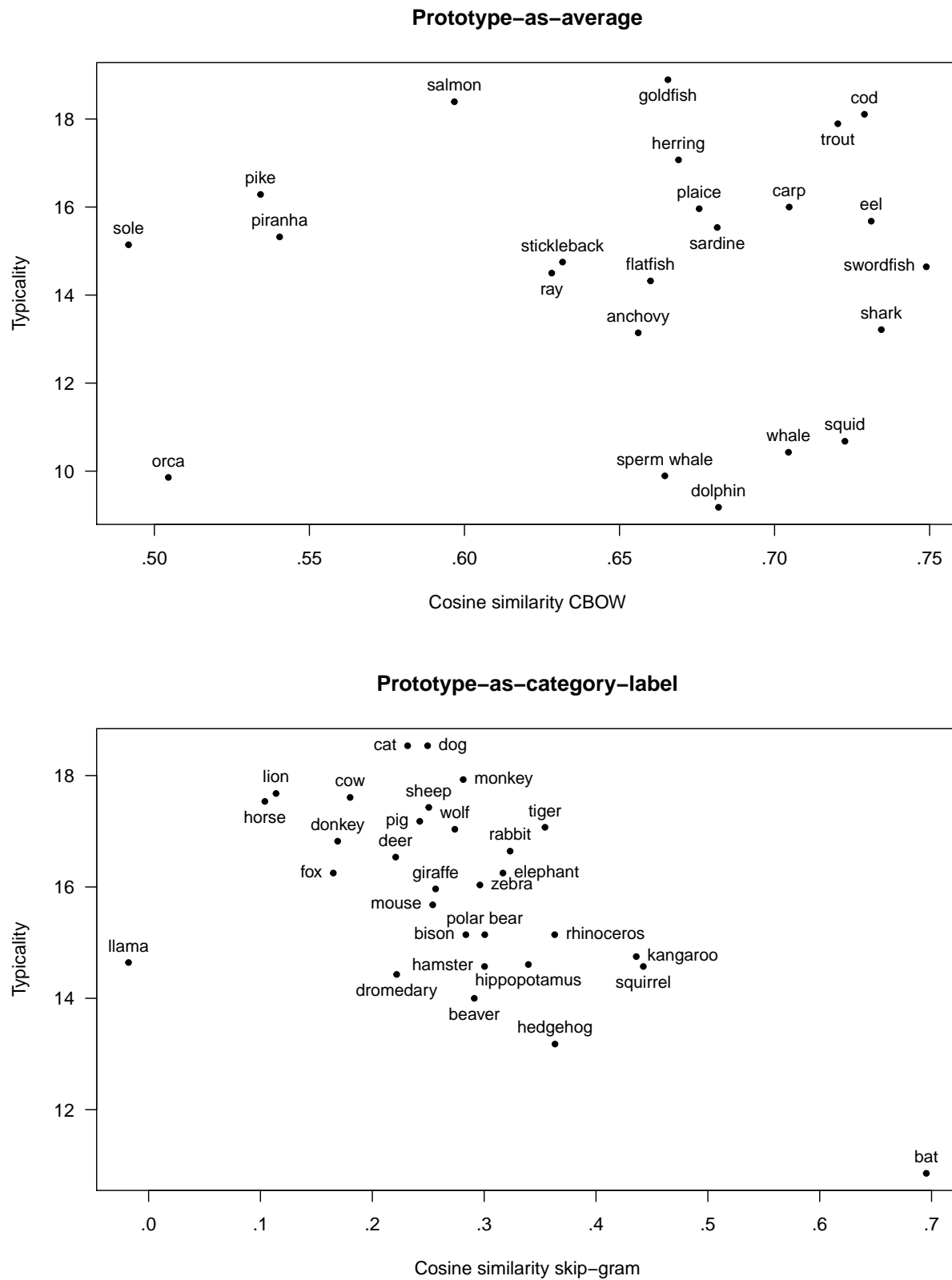


Figure 4. Results for the category fish (upper panel) and mammals (lower panel) in Dutch

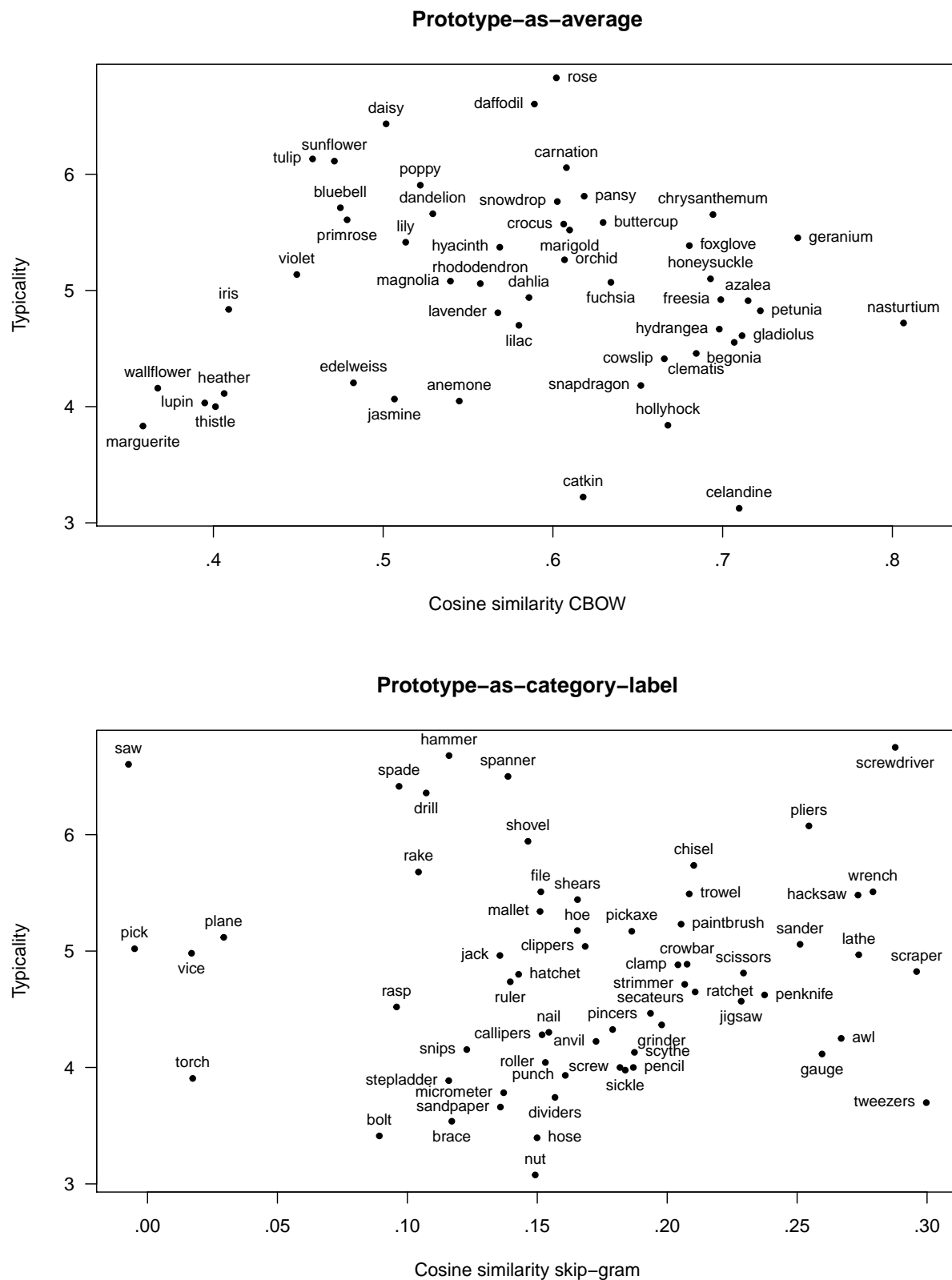


Figure 5. Results for the category flowers (upper panel) and tools (lower panel) in English for the younger adults

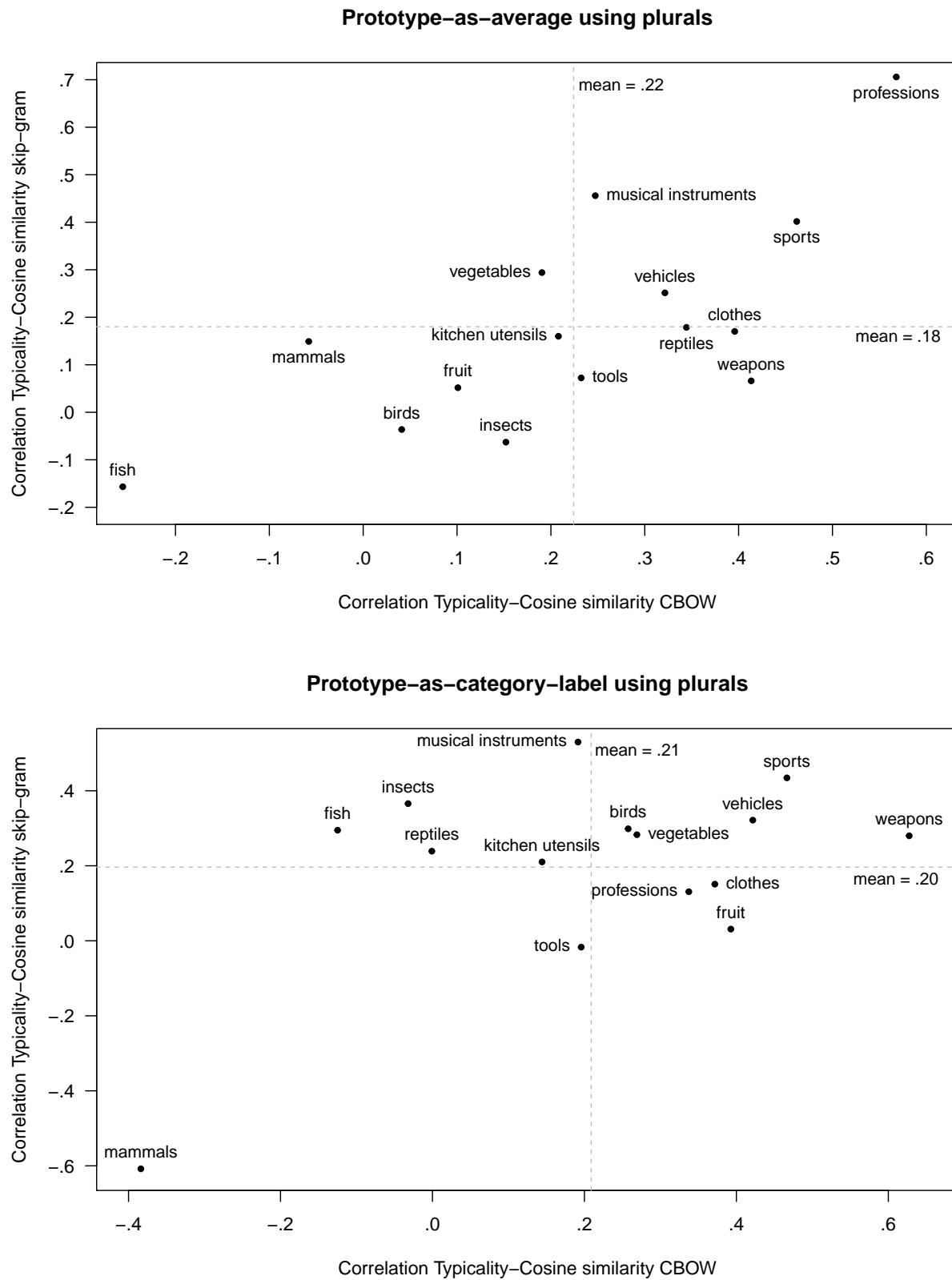


Figure 6. Summary of the results of Experiment 1b for Dutch, see main text for details



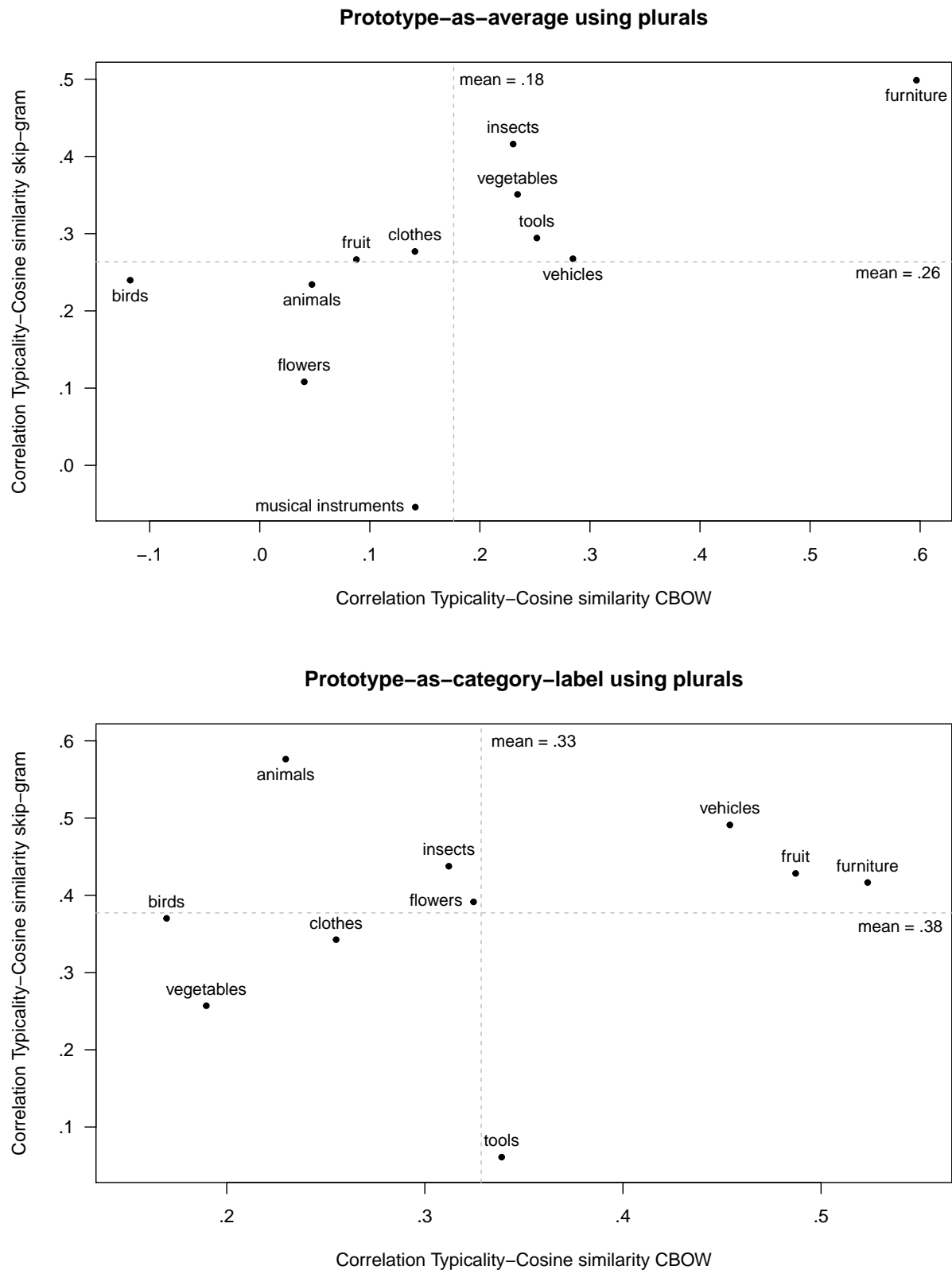


Figure 7. Summary of the results of Experiment 1b for English younger adults, see main text for details

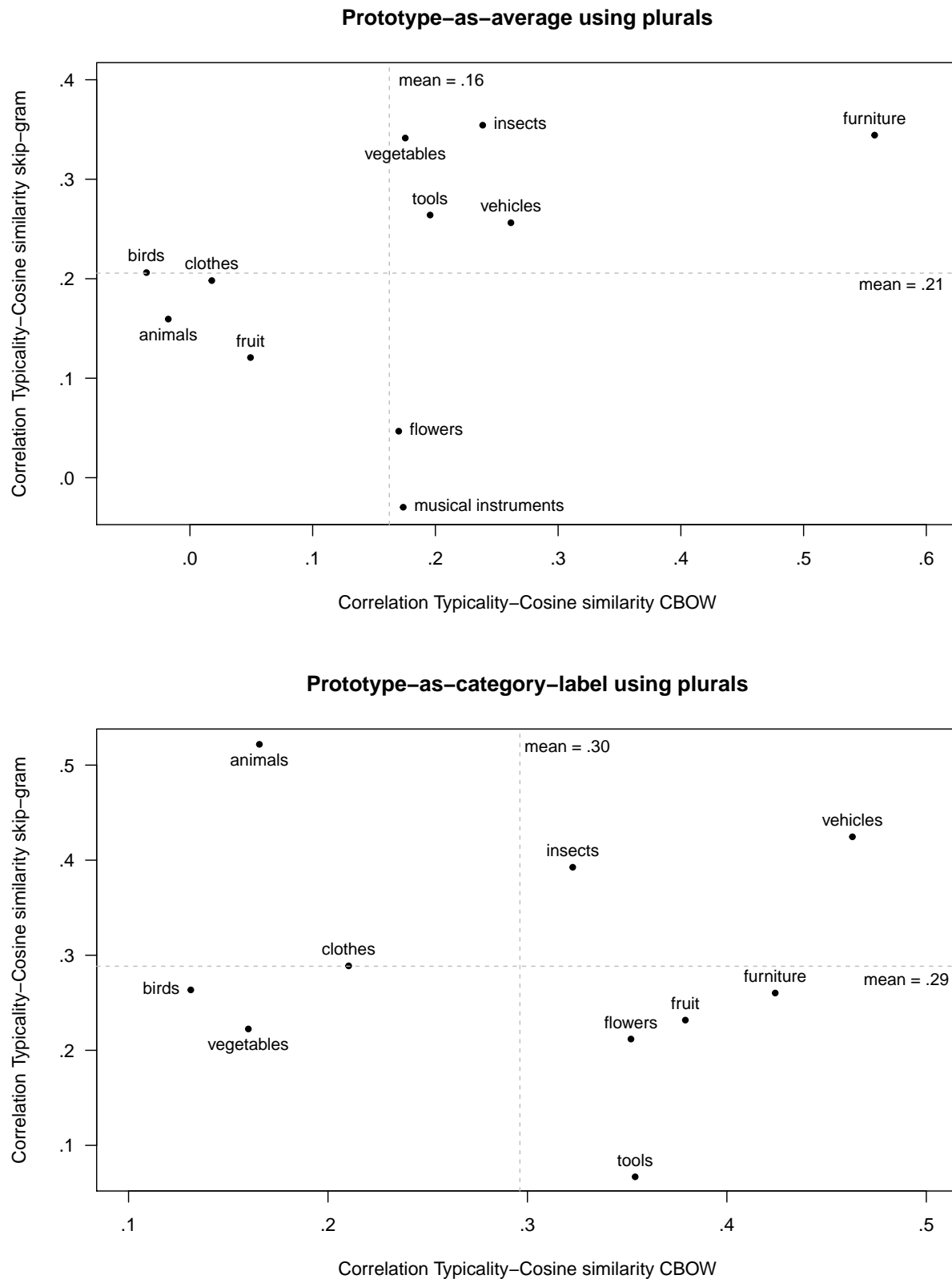
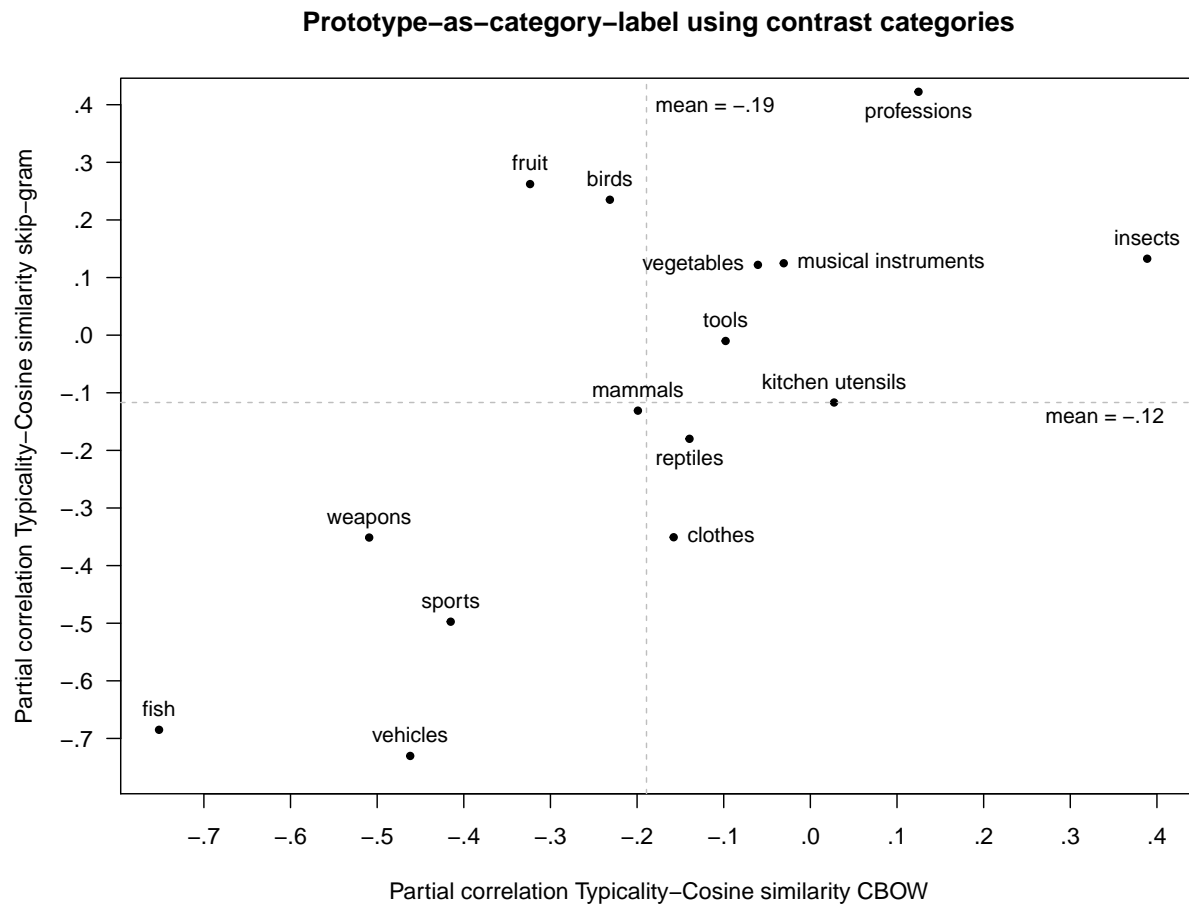


Figure 8. Summary of the results of Experiment 1b for English older adults, see main text for details



*Figure 9.* Summary of the results of Experiment 2a, see main text for details. As there was no skip-gram based partial correlation for the category kitchen utensils, we assigned it the cross-category average in this figure.

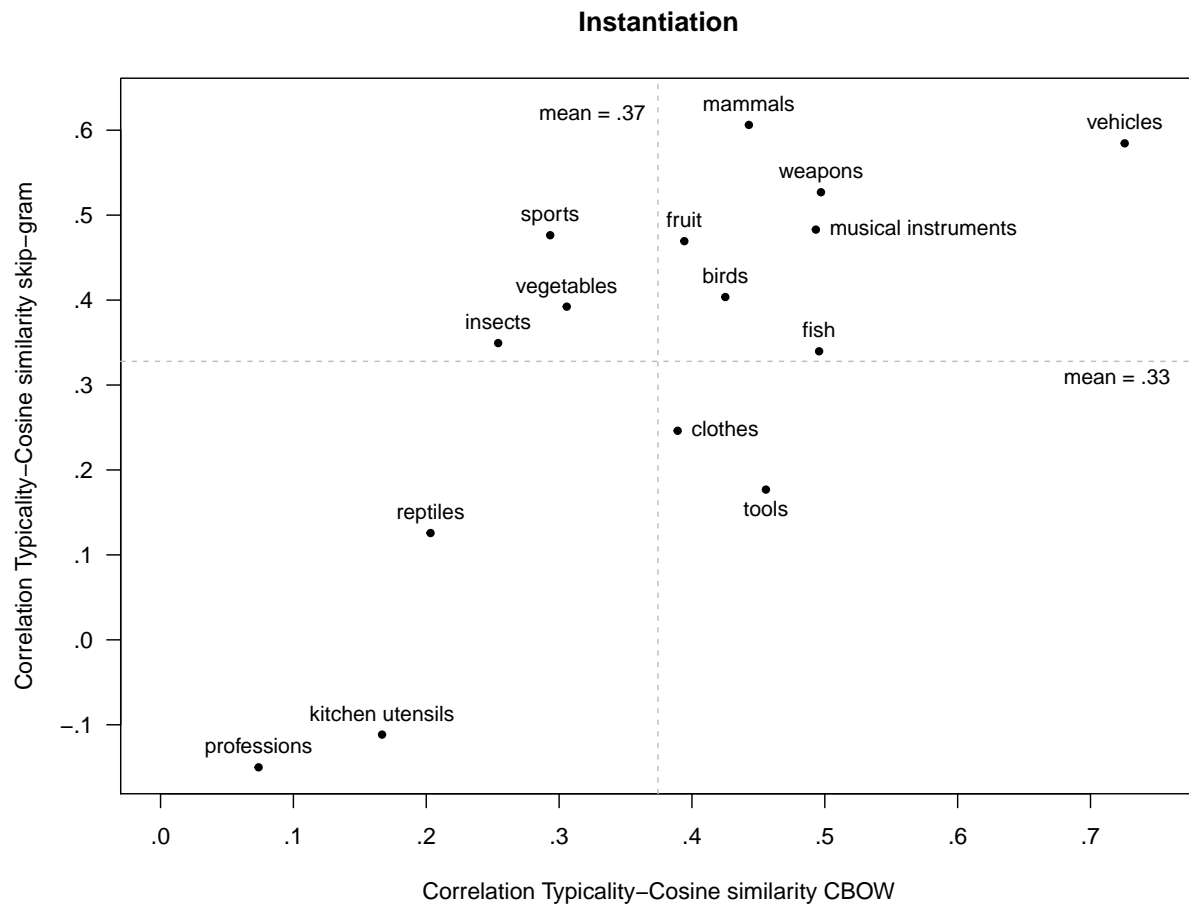


Figure 10. Summary of the results of Experiment 2b, see main text for details

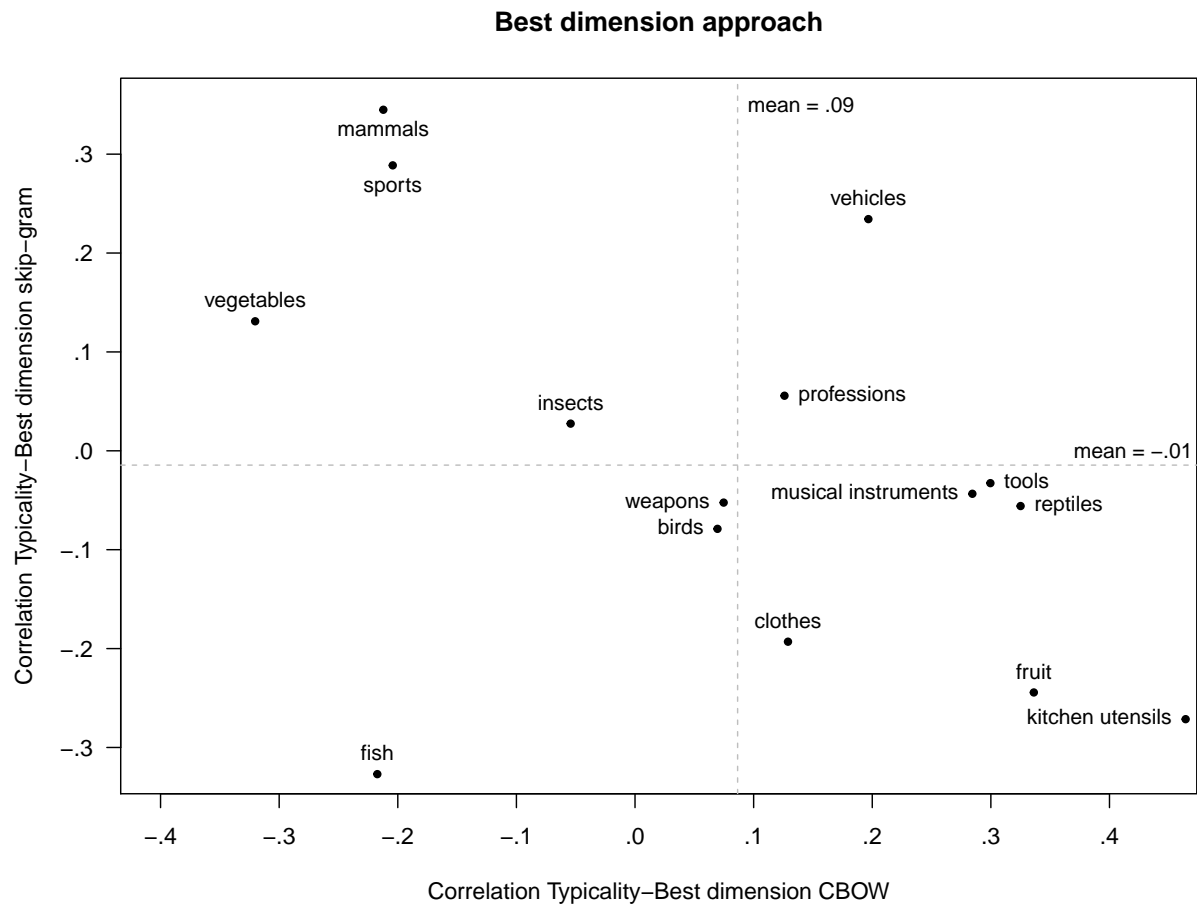


Figure 11. Summary of the results of Experiment 2c for Dutch, see main text for details

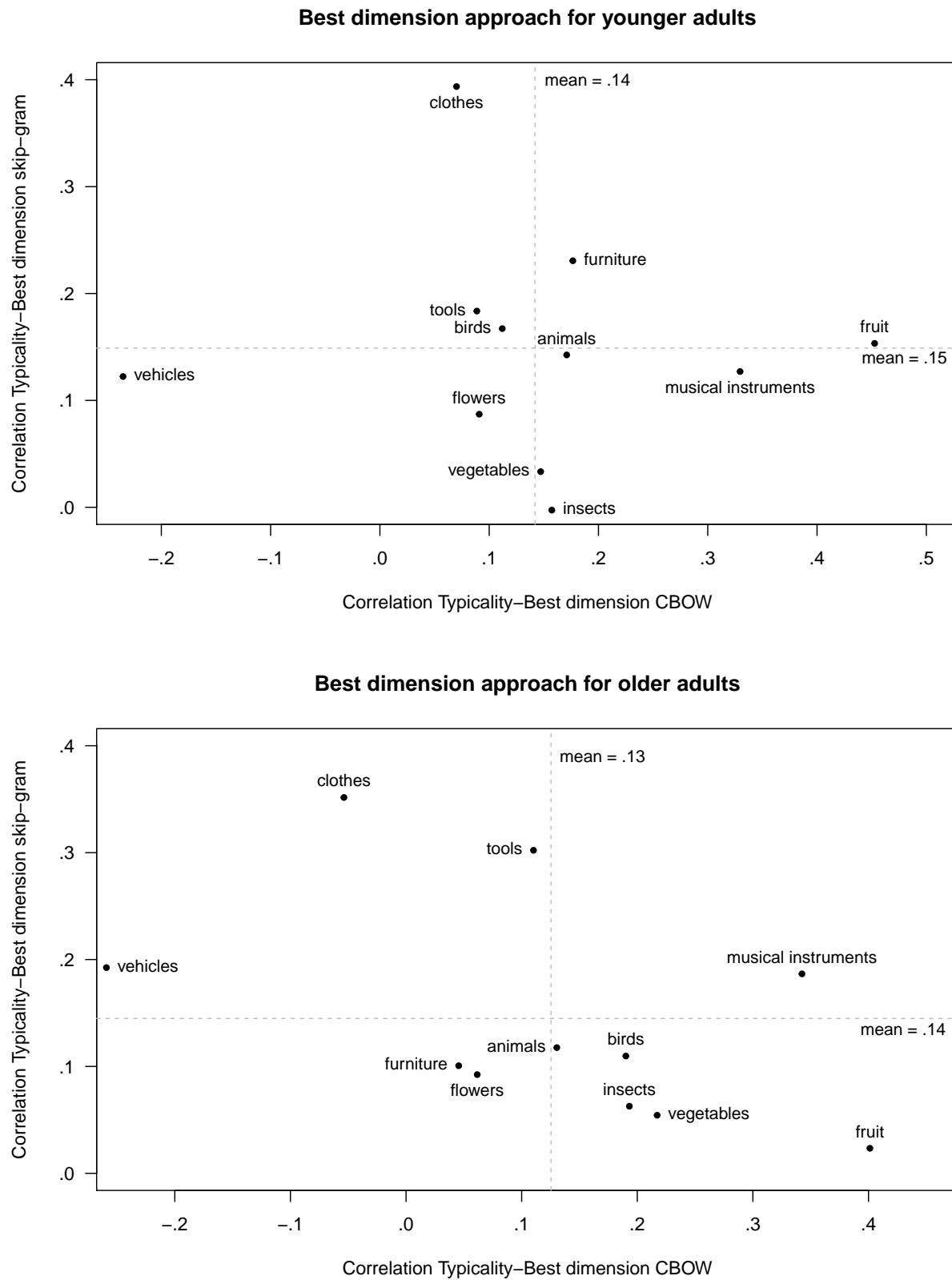


Figure 12. Summary of the results of Experiment 2c for English, see main text for details

## Appendix

Table A1

*Exemplars of the category birds sorted by typicality or similarity to the prototype with the actual values in parentheses for Dutch*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
sparrow (19.18)	crow (.75)	stork (.77)	parrot (.55)	heron (.61)
blackbird (18.82)	robin (.71)	crow (.70)	dove (.55)	woodpecker (.59)
robin (18.32)	magpie (.69)	duck (.70)	parakeet (.55)	stork (.57)
dove (18.14)	dove (.67)	heron (.70)	crow (.54)	parakeet (.56)
crow (18.11)	stork (.67)	dove (.69)	magpie (.52)	cuckoo (.55)
seagull (17.96)	ostrich (.66)	magpie (.67)	robin (.52)	duck (.55)
magpie (17.89)	parakeet (.65)	cuckoo (.67)	duck (.51)	magpie (.53)
canary (17.89)	parrot (.61)	parrot (.66)	seagull (.49)	canary (.52)
swallow (17.86)	woodpecker (.61)	woodpecker (.66)	stork (.49)	seagull (.51)
parakeet (17.64)	heron (.60)	robin (.64)	canary (.47)	dove (.51)
chickadee (17.11)	pheasant (.59)	owl (.64)	sparrow (.43)	crow (.50)
eagle (16.93)	owl (.59)	swan (.63)	ostrich (.43)	parrot (.49)
woodpecker (16.43)	cuckoo (.59)	vulture (.63)	owl (.42)	owl (.48)

heron (16.11)	duck (.59)	pheasant (.62)	woodpecker (.41)	robin (.48)
cuckoo (16.00)	canary (.58)	pelican (.61)	vulture (.39)	swallow (.47)
owl (16.00)	swan (.57)	seagull (.60)	heron (.37)	vulture (.47)
parrot (15.86)	seagull (.57)	parakeet (.59)	chicken (.37)	pheasant (.47)
falcon (15.50)	vulture (.55)	blackbird (.59)	swan (.35)	blackbird (.45)
stork (15.39)	pelican (.55)	ostrich (.57)	swallow (.34)	pelican (.45)
vulture (15.14)	sparrow (.53)	swallow (.57)	pheasant (.32)	ostrich (.40)
pheasant (13.71)	chicken (.53)	eagle (.57)	cuckoo (.31)	swan (.40)
swan (12.82)	swallow (.47)	canary (.56)	eagle (.30)	eagle (.34)
duck (12.79)	eagle (.46)	falcon (.54)	turkey (.30)	chickadee (.33)
pelican (12.57)	blackbird (.44)	chicken (.51)	pelican (.29)	falcon (.32)
peacock (12.29)	turkey (.42)	turkey (.51)	blackbird (.27)	chicken (.30)
turkey (11.68)	chickadee (.41)	rooster (.47)	rooster (.26)	turkey (.29)
chicken (11.57)	peacock (.40)	chickadee (.46)	falcon (.24)	sparrow (.23)
ostrich (11.21)	falcon (.40)	sparrow (.38)	peacock (.21)	rooster (.19)
rooster (11.07)	rooster (.40)	peacock (.38)	chickadee (.20)	peacock (.15)

---



Table A2

*Exemplars of the category clothes sorted by typicality or similarity to the prototype with the actual values in parentheses for Dutch*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
top (18.29)	blouse (.77)	shirt (.83)	blouse (.63)	skirt (.62)
pants (18.25)	coat (.76)	coat (.82)	boots (.56)	shirt (.62)
shirt (18.07)	pants (.76)	boots (.78)	bra (.56)	coat (.57)
jeans (18.00)	skirt (.75)	blouse (.78)	jeans (.56)	shoes (.53)
skirt (17.82)	sweater (.74)	socks (.77)	suit (.56)	blouse (.52)
coat (17.79)	boots (.74)	skirt (.77)	sweater (.55)	boots (.51)
pullover (17.79)	jeans (.73)	scarf (.74)	skirt (.53)	dress (.49)
blouse (17.43)	socks (.71)	hat (.73)	shirt (.53)	socks (.49)
sweater (17.32)	beanie (.70)	shoes (.70)	shoes (.51)	bra (.48)
dress (17.04)	shirt (.70)	beanie (.70)	coat (.51)	suit (.47)
suit (17.04)	shoes (.70)	suit (.64)	pants (.51)	bathing suit (.47)
shorts (16.68)	scarf (.70)	dress (.60)	panties (.48)	scarf (.46)
tracksuit (15.54)	suit (.66)	bra (.60)	scarf (.48)	belt (.45)
socks (14.43)	panties (.65)	belt (.59)	dress (.45)	hat (.43)
dungarees (14.23)	pullover (.65)	pants (.59)	beanie (.45)	beanie (.42)

bra (13.68)	bra (.63)	sweater (.57)	socks (.44)	pants (.37)
panties (13.64)	dress (.63)	pyjamas (.56)	pyjamas (.44)	pyjamas (.37)
pyjamas (13.11)	pyjamas (.63)	top (.54)	pullover (.44)	top (.31)
tie (12.93)	hat (.62)	bathing suit (.50)	hat (.39)	sweater (.30)
boots (12.54)	bathing suit (.60)	jeans (.50)	bathing suit (.39)	jeans (.27)
shoes (12.36)	cap (.53)	mittens (.48)	tie (.39)	panties (.26)
scarf (12.00)	tie (.52)	pullover (.47)	belt (.37)	mittens (.25)
cap (11.93)	belt (.51)	cap (.45)	dungarees (.32)	cap (.24)
bathing suit (11.82)	top (.50)	panties (.43)	cap (.31)	pullover (.23)
beanie (11.46)	shorts (.46)	tracksuit (.37)	shorts (.29)	tracksuit (.15)
hat (11.21)	dungarees (.43)	shorts (.22)	top (.27)	shorts (.11)
mittens (11.21)	tracksuit (.38)	tie (.17)	tracksuit (.24)	tie (-.04)
belt (10.86)	mittens (.28)	dungarees (-)	mittens (.17)	dungarees (-)

---

Table A3

*Exemplars of the category fish sorted by typicality or similarity to the prototype with the actual values in parentheses for Dutch*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
goldfish (18.89)	swordfish (.75)	eel (.81)	shark (.63)	cod (.65)
salmon (18.39)	shark (.73)	cod (.79)	cod (.61)	anchovy (.63)
cod (18.11)	eel (.73)	anchovy (.77)	eel (.59)	herring (.57)
trout (17.89)	cod (.73)	squid (.76)	trout (.59)	eel (.57)
herring (17.07)	squid (.72)	swordfish (.74)	carp (.59)	shark (.56)
pike (16.29)	trout (.72)	carp (.73)	swordfish (.58)	squid (.54)
carp (16.00)	carp (.70)	herring (.71)	goldfish (.58)	carp (.53)
plaice (15.96)	whale (.70)	shark (.70)	plaice (.57)	salmon (.53)
eel (15.68)	dolphin (.68)	salmon (.69)	whale (.56)	swordfish (.51)
sardine (15.54)	sardine (.68)	flatfish (.68)	herring (.55)	flatfish (.51)
piranha (15.32)	plaice (.68)	pike (.67)	anchovy (.55)	sardine (.50)
sole (15.14)	herring (.67)	stickleback (.66)	dolphin (.54)	stickleback (.46)
stickleback (14.75)	goldfish (.67)	plaice (.65)	squid (.54)	plaice (.44)
swordfish (14.64)	sperm whale (.66)	sardine (.63)	sardine (.53)	pike (.44)
ray (14.50)	flatfish (.66)	sperm whale (.61)	sperm whale (.52)	ray (.39)

flatfish (14.32)	anchovy (.66)	dolphin (.57)	flatfish (.52)	sperm whale (.37)
shark (13.21)	stickleback (.63)	orca (.56)	stickleback (.48)	dolphin (.33)
anchovy (13.14)	ray (.63)	ray (.53)	salmon (.47)	whale (.32)
squid (10.68)	salmon (.60)	whale (.50)	ray (.45)	orca (.27)
whale (10.43)	piranha (.54)	trout (.45)	piranha (.44)	trout (.27)
sperm whale (9.89)	pike (.53)	goldfish (.44)	pike (.43)	goldfish (.25)
orca (9.86)	orca (.50)	sole (.36)	orca (.34)	sole (.20)
dolphin (9.18)	sole (.49)	piranha (.35)	sole (.29)	piranha (.14)

---

Table A4

*Exemplars of the category fruit sorted by typicality or similarity to the prototype with the actual values in parentheses for Dutch*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
banana (19.14)	grapefruit (.80)	strawberry (.86)	banana (.47)	dates (.75)
apple (19.04)	orange (.78)	orange (.85)	dates (.41)	pineapple (.71)
orange (18.86)	lime (.77)	peach (.84)	nectarine (.40)	lime (.70)
strawberry (18.25)	apricot (.76)	raspberry (.83)	strawberry (.40)	banana (.70)
pear (18.18)	blackberry (.76)	banana (.82)	blackberry (.40)	coconut (.69)
grape (17.50)	peach (.76)	pineapple (.81)	apple (.39)	orange (.65)
pineapple (17.36)	nectarine (.75)	lychee (.81)	papaya (.38)	blackberry (.63)
cherry (17.21)	raspberry (.75)	lime (.80)	pineapple (.38)	papaya (.62)
kiwi (17.11)	pineapple (.74)	cherry (.78)	grapefruit (.37)	lychee (.60)
peach (17.00)	melon (.74)	papaya (.78)	peach (.37)	strawberry (.58)
plum (16.68)	strawberry (.74)	coconut (.77)	kiwi (.36)	raspberry (.56)
melon (16.54)	plum (.72)	nectarine (.77)	mango (.36)	pumpkin (.56)
raspberry (16.25)	papaya (.71)	fig (.76)	plum (.35)	kiwi (.56)
nectarine (16.25)	lemon (.66)	apricot (.74)	apricot (.35)	peach (.54)
apricot (16.00)	blueberry (.65)	kiwi (.73)	fig (.34)	lemon (.54)

mandarine (15.82)	banana (.64)	plum (.73)	raspberry (.34)	melon (.52)
grapefruit (15.71)	coconut (.64)	blackberry (.72)	melon (.34)	fig (.50)
lemon (15.25)	grape (.63)	grapefruit (.72)	coconut (.33)	cherry (.49)
mango (15.04)	lychee (.62)	dates (.72)	orange (.32)	mango (.47)
blueberry (14.96)	red currant (.62)	lemon (.71)	lychee (.31)	nectarine (.47)
passion fruit (14.93)	fig (.61)	melon (.71)	blueberry (.30)	apricot (.45)
blackberry (14.39)	mango (.61)	passion fruit (.68)	passion fruit (.30)	plum (.45)
lime (13.93)	dates (.60)	apple (.65)	pear (.28)	grapefruit (.41)
lychee (13.46)	kiwi (.58)	red currant (.64)	red currant (.28)	passion fruit (.41)
pumpkin (12.54)	apple (.58)	blueberry (.62)	grape (.28)	apple (.40)
papaya (12.46)	pumpkin (.58)	pumpkin (.62)	lime (.27)	grape (.36)
fig (12.43)	passion fruit (.53)	mango (.60)	lemon (.25)	blueberry (.33)
coconut (12.36)	cherry (.43)	grape (.55)	cherry (.25)	red currant (.32)
red currant (11.75)	pear (.41)	pear (.47)	pumpkin (.25)	pear (.26)
dates (10.86)	mandarine (.34)	mandarine (.16)	mandarine (.15)	mandarine (.06)

---

Table A5

*Exemplars of the category insects sorted by typicality or similarity to the prototype with the actual values in parentheses for Dutch*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
mosquito (18.39)	cockroach (.76)	earwig (.76)	cockroach (.57)	wood louse (.23)
fly (18.21)	grasshopper (.69)	grasshopper (.74)	worm (.55)	grasshopper (.21)
wood louse (17.29)	ladybug (.68)	wasp (.73)	ant (.53)	butterfly (.20)
wasp (17.25)	ant (.65)	cockchafer (.72)	ladybug (.53)	wasp (.19)
bee (17.07)	caterpillar (.65)	fly (.70)	wasp (.52)	leech (.18)
ant (17.00)	wasp (.64)	caterpillar (.69)	caterpillar (.52)	ladybug (.16)
cockroach (16.96)	bumblebee (.62)	beetle (.68)	grasshopper (.50)	bumblebee (.14)
grasshopper (16.61)	worm (.60)	flee (.67)	bumblebee (.47)	beetle (.14)
dragonfly (16.36)	butterfly (.60)	mosquito (.66)	fruit fly (.47)	cockchafer (.13)
cricket (16.32)	flee (.60)	ant (.62)	butterfly (.46)	earwig (.12)
fruit fly (16.25)	fly (.59)	ladybug (.61)	beetle (.45)	cricket (.12)
beetle (16.25)	earwig (.57)	dragonfly (.58)	fly (.44)	worm (.11)
horsefly (15.92)	fruit fly (.56)	cricket (.57)	spider (.40)	cockroach (.10)
bumblebee (15.86)	beetle (.55)	cockroach (.56)	dragonfly (.39)	mosquito (.10)
ladybug (15.39)	dragonfly (.53)	bumblebee (.56)	leech (.38)	caterpillar (.09)

cockchafer (15.04)	spider (.52)	fruit fly (.54)	earwig (.38)	horsefly (.09)
moth (14.68)	leech (.52)	centipede (.52)	flee (.37)	ant (.08)
caterpillar (14.18)	cricket (.48)	leech (.52)	wood louse (.33)	dragonfly (.07)
earwig (13.89)	mosquito (.43)	horsefly (.50)	cockchafer (.30)	fruit fly (.06)
centipede (13.79)	wood louse (.43)	spider (.44)	horsefly (.28)	flee (.06)
flee (13.61)	cockchafer (.40)	worm (.42)	mosquito (.26)	spider (.06)
butterfly (13.39)	horsefly (.39)	moth (.38)	cricket (.26)	centipede (.05)
louse (13.32)	moth (.37)	butterfly (.20)	centipede (.20)	bee (.02)
spider (12.25)	louse (.37)	wood louse (.19)	louse (.19)	louse (.01)
leech (11.36)	centipede (.34)	louse (.10)	moth (.18)	moth (.00)
worm (10.18)	bee (.25)	bee (.09)	bee (.10)	fly (-.02)

---



Table A6

*Exemplars of the category kitchen utensils sorted by typicality or similarity to the prototype with the actual values in parentheses for Dutch*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
stove (18.46)	spatula (.73)	spatula (.72)	toaster (.47)	glass (.55)
pan (18.39)	electric kettle (.73)	spoon (.70)	percolator (.45)	toaster (.49)
oven (17.75)	mixer (.71)	oven (.67)	electric kettle (.43)	place mat (.48)
mixer (17.46)	whisk (.70)	toaster (.67)	mixer (.39)	stove (.47)
microwave oven (17.14)	toaster (.69)	can opener (.66)	microwave oven (.38)	spatula (.47)
wok (17.11)	spoon (.68)	colander (.66)	oven (.34)	grater (.46)
whisk (16.93)	grater (.65)	electric kettle (.64)	stove (.33)	oven (.44)
pot (16.93)	microwave oven (.65)	bottle (.64)	fridge (.33)	colander (.43)
percolator (16.78)	oven (.64)	stove (.63)	can opener (.32)	spoon (.43)
spoon (16.43)	colander (.64)	grater (.62)	spoon (.32)	electric kettle (.43)
knife (16.43)	percolator (.62)	knife (.62)	spatula (.31)	can opener (.41)
colander (16.43)	teaspoon (.59)	glass (.61)	whisk (.30)	wok (.40)
fork (16.36)	wok (.57)	fork (.60)	wok (.27)	teaspoon (.36)
electric kettle (16.36)	fridge (.56)	scissors (.60)	grater (.26)	scissors (.36)

grater (16.21)	scissors (.56)	place mat (.60)	bottle (.26)	bottle (.33)
plate (16.18)	stove (.56)	wok (.59)	mug (.25)	sieve (.32)
bowl (15.89)	knife (.56)	kettle (.58)	pot (.25)	knife (.32)
can opener (15.82)	sieve (.56)	sieve (.56)	colander (.24)	mixer (.31)
fridge (15.82)	plate (.53)	teaspoon (.56)	nutcracker (.24)	fork (.29)
toaster (15.46)	bowl (.52)	towel (.54)	plate (.23)	plate (.28)
sieve (15.18)	can opener (.52)	plate (.53)	place mat (.23)	pot (.27)
spatula (14.86)	pot (.50)	mug (.52)	glass (.23)	towel (.25)
kettle (14.82)	pan (.49)	whisk (.51)	scissors (.22)	whisk (.23)
teaspoon (14.64)	fork (.49)	apron (.50)	knife (.22)	apron (.23)
apron (14.36)	bottle (.46)	mixer (.49)	fork (.17)	kettle (.22)
glass (14.32)	glass (.46)	pot (.49)	pan (.16)	scales (.20)
mug (14.14)	mug (.46)	scales (.43)	sieve (.15)	mug (.20)
towel (12.39)	towel (.42)	pan (.31)	kettle (.15)	nutcracker (.15)
scales (12.39)	kettle (.41)	bowl (.28)	scales (.14)	pan (.12)
nutcracker (12.30)	scales (.38)	nutcracker (.23)	bowl (.13)	microwave oven (.09)
place mat (12.11)	place mat (.38)	microwave oven (.21)	towel (.12)	bowl (.08)
bottle (11.79)	apron (.37)	fridge (.17)	apron (.09)	fridge (.04)
scissors (10.93)	nutcracker (.33)	percolator (-)	teaspoon (.08)	percolator (-)

---

Table A7

*Exemplars of the category mammals sorted by typicality or similarity to the prototype with the actual values in parentheses for Dutch*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
dog (18.54)	rhinoceros (.74)	hippopotamus (.78)	bat (.54)	bat (.70)
cat (18.54)	bison (.73)	dog (.77)	rhinoceros (.52)	squirrel (.44)
monkey (17.93)	giraffe (.73)	rabbit (.76)	hippopotamus (.48)	kangaroo (.44)
lion (17.68)	tiger (.70)	monkey (.76)	kangaroo (.47)	hedgehog (.36)
cow (17.61)	kangaroo (.68)	cat (.75)	hedgehog (.46)	rhinoceros (.36)
horse (17.54)	squirrel (.67)	elephant (.73)	polar bear (.45)	tiger (.35)
sheep (17.43)	polar bear (.65)	donkey (.73)	bison (.45)	hippopotamus (.34)
pig (17.18)	hamster (.64)	tiger (.72)	hamster (.41)	rabbit (.32)
tiger (17.07)	rabbit (.63)	sheep (.70)	monkey (.41)	elephant (.32)
wolf (17.04)	hippopotamus (.63)	polar bear (.69)	giraffe (.40)	polar bear (.30)
donkey (16.82)	hedgehog (.63)	pig (.69)	squirrel (.39)	hamster (.30)
rabbit (16.64)	elephant (.63)	rhinoceros (.69)	tiger (.39)	zebra (.30)
deer (16.54)	cat (.61)	deer (.67)	pig (.38)	beaver (.29)
elephant (16.25)	dromedary (.61)	bison (.66)	elephant (.35)	bison (.28)

fox (16.25)	pig (.61)	squirrel (.64)	dromedary (.35)	monkey (.28)
zebra (16.04)	cow (.60)	wolf (.63)	rabbit (.34)	wolf (.27)
giraffe (15.96)	donkey (.59)	hedgehog (.63)	cow (.32)	giraffe (.26)
mouse (15.68)	deer (.59)	mouse (.60)	cat (.31)	mouse (.25)
bison (15.14)	dog (.59)	giraffe (.59)	sheep (.30)	sheep (.25)
polar bear (15.14)	monkey (.58)	beaver (.59)	deer (.30)	dog (.25)
rhinoceros (15.14)	wolf (.58)	cow (.58)	zebra (.29)	pig (.24)
kangaroo (14.75)	horse (.58)	kangaroo (.58)	mouse (.28)	cat (.23)
llama (14.64)	zebra (.57)	horse (.58)	wolf (.28)	dromedary (.22)
hippopotamus (14.61)	bat (.56)	dromedary (.54)	dog (.27)	deer (.22)
squirrel (14.57)	sheep (.55)	hamster (.54)	lion (.25)	cow (.18)
hamster (14.57)	lion (.55)	fox (.54)	donkey (.22)	donkey (.17)
dromedary (14.43)	mouse (.54)	zebra (.51)	fox (.20)	fox (.17)
beaver (14.00)	llama (.45)	lion (.48)	llama (.18)	lion (.11)
hedgehog (13.18)	fox (.37)	bat (.47)	horse (.16)	horse (.10)
bat (10.86)	beaver (.35)	llama (.16)	beaver (.14)	llama (-.02)

---

Table A8

*Exemplars of the category musical instruments sorted by typicality or similarity to the prototype with the actual values in parentheses for Dutch*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
piano (19.43)	accordion (.80)	double bass (.93)	tambourine (.65)	bagpipe (.67)
guitar (19.25)	saxophone (.79)	accordion (.92)	accordion (.64)	flute (.62)
violin (18.86)	trombone (.77)	German flute (.91)	violin (.63)	accordion (.61)
saxophone (18.46)	German flute (.77)	trumpet (.91)	harpsichord (.62)	tambourine (.61)
trumpet (18.44)	cello (.77)	flute (.90)	guitar (.61)	harp (.60)
German flute (18.00)	clarinet (.76)	violin (.90)	German flute (.60)	German flute (.60)
flute (17.86)	trumpet (.76)	saxophone (.90)	trombone (.60)	double bass (.60)
drum set (17.68)	violin (.75)	piano (.89)	bagpipe (.58)	pan flute (.60)
bass guitar (17.50)	harpsichord (.75)	cello (.89)	piano (.58)	synthesizer (.59)
cello (17.29)	harmonica (.75)	trombone (.89)	cello (.57)	harmonica (.59)
clarinet (17.25)	guitar (.75)	clarinet (.88)	harmonica (.57)	bassoon (.58)
double bass (16.93)	tambourine (.75)	bassoon (.86)	pan flute (.57)	trumpet (.57)
drum (16.61)	double bass (.75)	guitar (.84)	clarinet (.56)	guitar (.55)
organ (16.18)	piano (.75)	harmonica (.83)	saxophone (.56)	drum set (.55)

harp (16.00)	drum set (.70)	harp (.82)	cymbals (.56)	violin (.55)
trombone (15.93)	bass guitar (.69)	tambourine (.82)	trumpet (.55)	cello (.55)
accordion (15.89)	bagpipe (.67)	banjo (.82)	drum set (.55)	harpsichord (.55)
harmonica (15.71)	bassoon (.66)	harpsichord (.79)	double bass (.54)	drum (.54)
synthesizer (15.68)	cymbals (.66)	bass guitar (.75)	synthesizer (.53)	clarinet (.53)
bassoon (15.44)	pan flute (.65)	triangle (.74)	bassoon (.52)	piano (.53)
bagpipe (15.14)	organ (.65)	bagpipe (.72)	organ (.50)	banjo (.53)
banjo (14.86)	harp (.65)	drum set (.71)	harp (.50)	trombone (.53)
triangle (14.54)	synthesizer (.63)	synthesizer (.69)	banjo (.49)	triangle (.52)
harpsichord (14.42)	banjo (.62)	pan flute (.65)	bass guitar (.48)	saxophone (.52)
tambourine (14.33)	flute (.53)	organ (.64)	flute (.40)	cymbals (.52)
cymbals (14.32)	drum (.48)	cymbals (.62)	triangle (.39)	organ (.50)
pan flute (14.25)	triangle (.45)	drum (.50)	drum (.33)	bass guitar (.43)

---

Table A9

*Exemplars of the category professions sorted by typicality or similarity to the prototype with the actual values in parentheses for Dutch*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
teacher (18.89)	police officer (.58)	shop-assistant (.70)	judge (.42)	judge (.63)
baker (18.18)	plumber (.57)	accountant (.69)	lawyer (.39)	lawyer (.47)
police officer (18.14)	doctor (.56)	doctor (.68)	informatician (.29)	accountant (.33)
doctor (17.93)	lawyer (.54)	lawyer (.68)	accountant (.29)	shop-assistant (.27)
plumber (17.89)	accountant (.52)	police officer (.67)	plumber (.26)	veterinarian (.27)
secretary (17.71)	secretary (.51)	secretary (.67)	educator (.24)	pharmacist (.27)
accountant (17.54)	veterinarian (.51)	pharmacist (.64)	police officer (.22)	dentist (.26)
dentist (17.50)	garbage collector (.51)	dentist (.64)	stewardess (.20)	fireman (.26)
postman (17.43)	dentist (.50)	plumber (.64)	secretary (.20)	police officer (.24)
cook (17.32)	teacher (.50)	fireman (.62)	doctor (.19)	secretary (.23)
shop-assistant (17.32)	fireman (.49)	veterinarian (.62)	garbage collector (.19)	garbage collector (.22)
butcher (17.25)	pharmacist (.48)	garbage collector (.62)	teacher (.18)	psychologist (.21)
manager (17.07)	informatician (.48)	butcher (.59)	postman (.18)	teacher (.20)
pharmacist (16.89)	butcher (.48)	psychologist (.58)	actor (.17)	butcher (.20)

veterinarian (16.82)	shop-assistant (.46)	stewardess (.57)	baker (.16)	doctor (.19)
lawyer (16.79)	educator (.45)	postman (.57)	psychologist (.15)	plumber (.18)
fireman (16.75)	postman (.44)	teacher (.57)	fireman (.14)	physiotherapist (.18)
architect (16.36)	physiotherapist (.43)	cook (.53)	architect (.14)	manager (.18)
stewardess (16.21)	baker (.43)	educator (.53)	dentist (.13)	informatician (.16)
physiotherapist (16.18)	actor (.43)	physiotherapist (.52)	cook (.12)	stewardess (.16)
garbage collector (16.18)	manager (.41)	actor (.51)	butcher (.12)	educator (.15)
educator (16.14)	psychologist (.41)	archaeologist (.48)	pharmacist (.10)	baker (.14)
judge (16.04)	stewardess (.41)	manager (.47)	pilot (.09)	postman (.12)
pilot (15.93)	cook (.40)	stallholder (.46)	veterinarian (.08)	minister (.11)
psychologist (15.79)	judge (.39)	baker (.46)	manager (.08)	archaeologist (.10)
informatician (14.61)	pilot (.38)	pilot (.45)	stallholder (.08)	stallholder (.09)
stallholder (14.54)	stallholder (.38)	informatician (.44)	physiotherapist (.07)	architect (.08)
minister (13.57)	architect (.37)	architect (.38)	archaeologist (.06)	cook (.07)
actor (12.96)	archaeologist (.29)	judge (.36)	shop-assistant (.04)	pilot (.05)
archaeologist (12.89)	minister (.19)	minister (.31)	minister (.01)	actor (.05)

---



Table A10

*Exemplars of the category reptiles sorted by typicality or similarity to the prototype with the actual values in parentheses for Dutch*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
lizard (17.68)	lizard (.76)	lizard (.82)	crocodile (.62)	crocodile (.57)
crocodile (17.46)	iguana (.73)	gecko (.81)	lizard (.61)	monitor lizard (.51)
iguana (17.04)	crocodile (.73)	frog (.79)	alligator (.60)	alligator (.46)
chameleon (16.64)	tortoise (.71)	snake (.77)	snake (.59)	lizard (.45)
alligator (16.43)	snake (.69)	salamander (.76)	iguana (.58)	tortoise (.44)
snake (15.93)	alligator (.66)	tortoise (.76)	dinosaur (.57)	blindworm (.42)
tortoise (15.18)	dinosaur (.63)	iguana (.76)	tortoise (.55)	dinosaur (.41)
gecko (14.44)	blindworm (.62)	viper (.76)	blindworm (.55)	snake (.40)
monitor lizard (14.42)	frog (.60)	blindworm (.75)	monitor lizard (.50)	caiman (.39)
python (13.96)	gecko (.59)	monitor lizard (.75)	viper (.45)	viper (.38)
cobra (13.93)	monitor lizard (.54)	caiman (.72)	gecko (.44)	iguana (.38)
salamander (13.93)	viper (.54)	crocodile (.71)	frog (.43)	turtle (.37)
boa (13.86)	boa (.54)	chameleon (.65)	turtle (.41)	gecko (.37)
caiman (13.76)	salamander (.51)	turtle (.60)	boa (.38)	salamander (.37)

viper (13.56)	cobra (.47)	alligator (.60)	cobra (.36)	frog (.34)
turtle (13.32)	turtle (.45)	toad (.49)	chameleon (.35)	chameleon (.27)
dinosaur (11.96)	python (.45)	dinosaur (.46)	python (.35)	toad (.23)
frog (11.41)	chameleon (.44)	boa (.35)	salamander (.35)	boa (.15)
toad (11.11)	caiman (.38)	cobra (.30)	caiman (.30)	cobra (.12)
blindworm (5.14)	toad (.35)	python (.27)	toad (.23)	python (.10)

---

Table A11

*Exemplars of the category sports sorted by typicality or similarity to the prototype with the actual values in parentheses for Dutch*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
soccer (18.96)	ice hockey (.69)	gymnastics (.89)	tennis (.55)	volleyball (.60)
tennis (18.68)	billiards (.69)	table tennis (.88)	ice hockey (.54)	ice hockey (.59)
basketball (18.57)	tennis (.69)	volleyball (.87)	squash (.51)	gymnastics (.56)
swimming (18.57)	squash (.69)	basketball (.86)	rugby (.51)	table tennis (.54)
running (18.36)	gymnastics (.68)	handball (.84)	soccer (.50)	squash (.53)
volleyball (18.25)	badminton (.67)	boxing (.83)	cycling (.50)	handball (.52)
cycling (17.89)	handball (.65)	badminton (.82)	baseball (.49)	judo (.51)
gymnastics (17.32)	basketball (.64)	squash (.81)	gymnastics (.47)	basketball (.51)
judo (17.07)	table tennis (.64)	ice hockey (.79)	billiards (.47)	boxing (.51)
handball (16.82)	judo (.64)	judo (.77)	basketball (.47)	horseback riding (.50)
squash (16.61)	rugby (.62)	horseback riding (.74)	judo (.44)	rugby (.50)
baseball (16.46)	volleyball (.62)	running (.70)	table tennis (.44)	badminton (.47)
badminton (16.21)	baseball (.61)	billiards (.70)	boxing (.43)	billiards (.44)
ice hockey (16.11)	horseback riding (.58)	tennis (.69)	horseback riding (.42)	soccer (.42)
long jump (16.00)	boxing (.56)	cycling (.68)	badminton (.42)	tennis (.42)

rugby (15.75)	swimming (.54)	rugby (.68)	running (.42)	running (.40)
boxing (15.61)	running (.53)	swimming (.62)	volleyball (.41)	cycling (.38)
surfing (15.21)	chess (.51)	chess (.62)	handball (.41)	walking (.36)
horseback riding (15.04)	soccer (.51)	surfing (.61)	golfing (.38)	surfing (.34)
ballet (14.54)	surfing (.49)	soccer (.60)	chess (.35)	baseball (.32)
table tennis (14.54)	shot-put (.48)	sailing (.52)	surfing (.33)	sport fishing (.30)
fencing (13.75)	cycling (.48)	walking (.51)	swimming (.29)	fencing (.28)
shot-put (13.00)	golfing (.48)	sport fishing (.50)	shot-put (.26)	swimming (.25)
sailing (12.93)	long jump (.42)	fencing (.49)	walking (.23)	sailing (.25)
golfing (12.39)	ballet (.41)	shot-put (.43)	sport fishing (.21)	chess (.23)
walking (12.18)	walking (.38)	baseball (.42)	long jump (.18)	shot-put (.20)
billiards (9.79)	sailing (.35)	long jump (.41)	sailing (.17)	long jump (.19)
sport fishing (8.18)	fencing (.27)	golfing (.28)	fencing (.16)	golfing (.16)
chess (7.71)	sport fishing (.24)	ballet (.23)	ballet (.14)	ballet (.13)

---

Table A12

*Exemplars of the category tools sorted by typicality or similarity to the prototype with the actual values in parentheses for Dutch*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
hammer (18.43)	screwdriver (.76)	chisel (.76)	drill (.56)	rope (.59)
saw (17.96)	drill (.71)	pickaxe (.73)	screwdriver (.53)	filling-knife (.58)
screwdriver (17.57)	saw (.70)	rope (.71)	saw (.50)	chisel (.57)
shovel (17.54)	grinding disc (.66)	knife (.70)	lawn mower (.48)	file (.51)
drill (17.36)	chisel (.66)	screwdriver (.69)	grinding disc (.47)	grinding disc (.51)
tongs (16.86)	hammer (.64)	file (.67)	pickaxe (.46)	screwdriver (.48)
chisel (16.11)	file (.62)	drill (.66)	file (.42)	knife (.47)
wrench (16.04)	wrench (.62)	crowbar (.66)	crowbar (.41)	vacuum cleaner (.46)
axe (15.96)	pickaxe (.61)	hammer (.65)	hammer (.38)	level (.46)
grinding disc (15.32)	knife (.61)	grinding disc (.63)	wheelbarrow (.38)	drill (.45)
pickaxe (14.96)	crowbar (.61)	anvil (.63)	chisel (.36)	crowbar (.45)
crowbar (14.89)	lawn mower (.58)	vacuum cleaner (.61)	knife (.35)	anvil (.43)
crowbar (14.57)	anvil (.56)	wrench (.60)	wrench (.35)	pickaxe (.42)
wheelbarrow (14.57)	axe (.55)	filling-knife (.60)	nail (.35)	wrench (.41)
filling-knife (14.14)	nail (.55)	axe (.59)	vacuum cleaner (.34)	wire brush (.36)

level (14.14)	rope (.52)	level (.59)	level (.33)	hammer (.36)
plane (13.12)	level (.52)	nail (.57)	crowbar (.29)	wheelbarrow (.34)
paint brush (13.00)	wheelbarrow (.50)	crowbar (.55)	rope (.29)	saw (.33)
plough (12.71)	shovel (.50)	wheelbarrow (.54)	anvil (.28)	nail (.32)
clamp (12.26)	wire brush (.50)	wire brush (.52)	shovel (.26)	plane (.32)
lawn mower (12.18)	filling-knife (.49)	plane (.48)	plane (.26)	axe (.31)
file (12.04)	vacuum cleaner (.44)	shovel (.47)	axe (.25)	crowbar (.27)
nail (11.96)	crowbar (.42)	saw (.45)	filling-knife (.24)	lawn mower (.20)
knife (11.75)	tongs (.39)	lawn mower (.39)	wire brush (.24)	shovel (.19)
wire brush (11.64)	paint brush (.33)	tongs (.24)	tongs (.20)	tongs (.16)
anvil (11.14)	plane (.30)	plough (.10)	paint brush (.18)	plough (-.04)
oilcan (10.78)	plough (.09)	clamp (-)	oilcan (.06)	clamp (-)
rope (9.46)	oilcan (.03)	oilcan (-)	plough (.01)	oilcan (-)
vacuum cleaner (8.21)	clamp (-)	paint brush (-)	clamp (-)	paint brush (-)

---

Table A13

*Exemplars of the category vegetables sorted by typicality or similarity to the prototype with the actual values in parentheses for Dutch*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
lettuce (18.86)	zucchini (.82)	cucumber (.88)	cauliflower (.56)	cucumber (.76)
carrot (18.86)	radish (.81)	zucchini (.87)	spinach (.56)	leek (.74)
leek (18.25)	cauliflower (.80)	spinach (.86)	peas (.53)	spinach (.74)
cauliflower (18.00)	chervil (.80)	parsley (.86)	tomato (.51)	zucchini (.73)
spinach (18.00)	black salsify (.79)	leek (.86)	radish (.50)	gherkins (.73)
beans (17.96)	water cress (.78)	tomato (.85)	eggplant (.50)	beans (.72)
endive (17.68)	parsley (.78)	gherkins (.85)	beet (.50)	cauliflower (.72)
tomato (17.32)	spinach (.77)	eggplant (.85)	cucumber (.50)	mushrooms (.71)
peas (17.19)	endive (.77)	cauliflower (.84)	pepper (.49)	lettuce (.70)
Brussels sprouts (17.14)	peas (.76)	mushrooms (.83)	leek (.48)	tomato (.69)
cucumber (16.82)	leek (.74)	potato (.83)	water cress (.48)	parsley (.68)
celery (16.71)	mushrooms (.73)	celery (.80)	black salsify (.47)	potato (.68)
asparagus (16.43)	cucumber (.72)	beans (.79)	zucchini (.46)	eggplant (.68)
zucchini (15.79)	tomato (.72)	radish (.79)	endive (.45)	peas (.67)

eggplant (15.71)	pepper (.71)	peas (.77)	chervil (.45)	asparagus (.67)
pepper (15.14)	eggplant (.71)	asparagus (.77)	mushrooms (.43)	celery (.63)
black salsify (14.68)	asparagus (.69)	lettuce (.74)	gherkins (.42)	radish (.61)
radish (14.54)	beet (.68)	beet (.69)	beans (.42)	Brussels sprouts (.53)
beet (14.04)	onions (.65)	water cress (.69)	parsley (.39)	beet (.51)
mushrooms (13.54)	beans (.64)	black salsify (.66)	Brussels sprouts (.37)	black salsify (.51)
chervil (13.48)	gherkins (.61)	Brussels sprouts (.64)	asparagus (.36)	water cress (.46)
onions (12.89)	potato (.59)	chervil (.64)	potato (.36)	onions (.46)
gherkins (12.82)	Brussels sprouts (.58)	onions (.61)	onions (.36)	chervil (.44)
water cress (12.54)	celery (.53)	endive (.57)	lettuce (.30)	endive (.39)
parsley (12.00)	carrot (.47)	carrot (.50)	carrot (.29)	carrot (.32)
potato (10.71)	lettuce (.39)	garlic (.20)	celery (.27)	garlic (.14)
garlic (9.46)	garlic (.34)	pepper (-)	garlic (.15)	pepper (-)

---



Table A14

*Exemplars of the category vehicles sorted by typicality or similarity to the prototype with the actual values in parentheses for Dutch*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
car (19.68)	car (.75)	truck (.80)	car (.56)	tractor (.63)
bus (18.50)	jeep (.65)	car (.77)	jeep (.56)	truck (.61)
van (18.18)	bus (.63)	bicycle (.71)	truck (.44)	car (.57)
jeep (17.93)	truck (.61)	moped (.69)	scooter (.42)	airplane (.57)
truck (17.89)	sled (.61)	tractor (.68)	helicopter (.41)	subway train (.51)
train (17.54)	truck (.60)	helicopter (.66)	bus (.39)	bicycle (.51)
taxi (17.50)	moped (.60)	airplane (.64)	sled (.36)	rocket (.47)
motorbike (17.32)	bicycle (.59)	cart (.64)	moped (.36)	scooter (.47)
moped (17.25)	taxi (.58)	boat (.64)	motorbike (.35)	sled (.45)
truck (17.07)	scooter (.57)	jeep (.61)	truck (.35)	helicopter (.45)
bicycle (17.07)	train (.56)	van (.59)	taxi (.34)	hovercraft (.44)
scooter (16.75)	tractor (.55)	train (.59)	airplane (.34)	train (.43)
tram (16.57)	airplane (.55)	bus (.58)	trailer (.33)	moped (.43)
airplane (15.04)	motorbike (.55)	carriage (.57)	tractor (.32)	jeep (.42)

subway train (13.81)	boat (.54)	(hot air) balloon (.57)	train (.32)	bus (.41)
boat (12.75)	skateboard (.54)	subway train (.56)	subway train (.31)	cart (.40)
helicopter (12.68)	hovercraft (.54)	scooter (.55)	bicycle (.30)	skateboard (.40)
trailer (12.61)	helicopter (.53)	hovercraft (.55)	van (.30)	van (.37)
tractor (12.11)	(hot air) balloon (.50)	taxi (.55)	hovercraft (.30)	truck (.36)
carriage (11.86)	van (.50)	skateboard (.53)	rocket (.29)	boat (.34)
kick scooter (10.32)	carriage (.49)	sled (.53)	carriage (.29)	tram (.32)
cart (10.21)	subway train (.48)	trailer (.51)	submarine (.26)	carriage (.30)
go-cart (9.96)	trailer (.48)	submarine (.51)	boat (.24)	submarine (.29)
hovercraft (9.52)	cart (.48)	rocket (.50)	skateboard (.23)	trailer (.27)
(hot air) balloon (9.29)	go-cart (.47)	truck (.46)	tram (.23)	(hot air) balloon (.26)
skateboard (8.54)	rocket (.45)	go-cart (.44)	go-cart (.23)	taxi (.25)
Zeppelin (8.37)	tram (.43)	tram (.43)	cart (.19)	go-cart (.22)
sled (8.07)	submarine (.41)	kick scooter (.31)	(hot air) balloon (.16)	motorbike (.18)
submarine (7.32)	Zeppelin (.30)	motorbike (.28)	Zeppelin (.15)	kick scooter (.14)
rocket (6.50)	kick scooter (.23)	Zeppelin (.26)	kick scooter (.05)	Zeppelin (.08)

---

Table A15

*Exemplars of the category weapons sorted by typicality or similarity to the prototype with the actual values in parentheses for Dutch*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
pistol (19.43)	rifle (.72)	rifle (.74)	pistol (.83)	shield (.61)
rifle (19.39)	pistol (.70)	sword (.72)	rifle (.76)	sword (.46)
machine gun (18.86)	sword (.68)	club (.70)	shotgun (.54)	rifle (.42)
bazooka (18.50)	dagger (.66)	pistol (.69)	dagger (.52)	pistol (.32)
grenade (17.32)	grenade (.61)	dagger (.68)	bazooka (.49)	slingshot (.32)
dagger (17.21)	shotgun (.61)	spear (.66)	grenade (.47)	canon (.32)
sword (17.07)	bazooka (.60)	grenade (.64)	sword (.47)	grenade (.30)
canon (16.96)	canon (.60)	whip (.64)	canon (.43)	spear (.27)
shotgun (16.57)	slingshot (.59)	axe (.63)	slingshot (.40)	dagger (.25)
club (14.39)	spear (.58)	rope (.62)	knuckle dusters (.40)	shotgun (.24)
spear (13.89)	stick (.58)	stick (.62)	shield (.37)	stick (.24)
tank (13.14)	axe (.54)	bazooka (.59)	rope (.34)	axe (.24)
axe (12.89)	rope (.54)	canon (.59)	club (.34)	knuckle dusters (.23)
slingshot (11.57)	bow (.53)	slingshot (.55)	stick (.33)	club (.23)

bow (11.36)	shield (.52)	knuckle dusters (.54)	axe (.33)	whip (.21)
knuckle dusters (11.11)	tank (.52)	bow (.53)	tank (.31)	bazooka (.20)
whip (9.86)	knuckle dusters (.50)	shotgun (.50)	bow (.30)	bow (.19)
stick (9.00)	club (.50)	shield (.50)	spear (.29)	tank (.18)
shield (7.11)	whip (.40)	machine gun (.48)	whip (.20)	machine gun (.15)
rope (5.21)	machine gun (.09)	tank (.45)	machine gun (.13)	rope (.13)

---

Table A16

*Exemplars of the category animals sorted by typicality, according to the younger adults, or similarity to the prototype with the actual values in parentheses (older adults' ratings after the forward slash) for English*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
dog (6.92/6.98)	ocelot (.72)	rabbit (.78)	chimpanzee (.49)	dog (.64)
cat (6.89/6.82)	anteater (.71)	lizard (.76)	alligator (.46)	cat (.59)
horse (6.83/6.85)	muskrat (.69)	otter (.75)	hippopotamus (.46)	pig (.56)
cow (6.69/6.85)	wallaby (.67)	squirrel (.74)	hyena (.45)	llama (.55)
rabbit (6.64/6.38)	tapir (.63)	toad (.74)	elephant (.44)	coyote (.54)
lion (6.48/6.61)	dingo (.63)	possum (.72)	dog (.44)	goat (.53)
sheep (6.47/6.75)	alligator (.63)	fox (.71)	polecat (.44)	rabbit (.53)
pig (6.45/6.63)	baboon (.62)	frog (.71)	iguana (.44)	alligator (.51)
bear (6.21/6.23)	stoat (.62)	coyote (.71)	baboon (.44)	tiger (.51)
elephant (6.21/6.39)	warthog (.62)	turtle (.70)	rabbit (.43)	chimpanzee (.50)
tiger (6.19/6.33)	hippopotamus (.62)	crocodile (.70)	ocelot (.43)	deer (.50)
monkey (6.15/6.23)	antelope (.61)	panther (.69)	deer (.43)	orangutan (.50)
fox (6.04/5.94)	hyena (.61)	cat (.69)	raccoon (.42)	iguana (.50)
pony (5.96/6.34)	raccoon (.61)	snake (.69)	lion (.41)	elephant (.50)

goat (5.91/6.10)	polecat (.60)	iguana (.69)	tapir (.41)	cougar (.49)
gorilla (5.89/6.02)	rabbit (.60)	wolf (.69)	ape (.40)	raccoon (.49)
mouse (5.85/5.94)	elephant (.59)	tortoise (.69)	cheetah (.40)	wallaby (.49)
ape (5.74/6.19)	mongoose (.58)	lemur (.69)	cat (.40)	wolf (.48)
chimpanzee (5.74/5.98)	snake (.58)	alligator (.68)	giraffe (.40)	snake (.48)
donkey (5.70/6.02)	frog (.57)	tapir (.68)	orangutan (.40)	panther (.47)
zebra (5.62/5.57)	lion (.57)	monkey (.68)	ferret (.39)	sheep (.47)
deer (5.62/6.13)	lemur (.57)	tiger (.68)	anteater (.39)	monkey (.47)
squirrel (5.58/6.02)	cougar (.57)	giraffe (.68)	monkey (.38)	turtle (.46)
wolf (5.57/5.89)	iguana (.57)	cougar (.68)	chinchilla (.38)	otter (.46)
hippopotamus (5.53/5.62)	tiger (.57)	elephant (.68)	muskrat (.38)	skunk (.46)
giraffe (5.47/5.67)	rhinoceros (.56)	stoat (.68)	rhinoceros (.38)	whale (.46)
hedgehog (5.47/5.54)	monkey (.56)	anteater (.68)	dingo (.38)	elk (.46)
leopard (5.46/5.74)	giraffe (.56)	gorilla (.68)	gerbil (.38)	bison (.45)
panda (5.43/5.38)	weasel (.56)	moose (.67)	badger (.37)	cow (.45)
rhinoceros (5.43/5.58)	squirrel (.55)	wallaby (.67)	lemur (.37)	jaguar (.45)
hamster (5.42/5.17)	wildebeest (.55)	hippopotamus (.67)	snake (.37)	horse (.45)
kangaroo (5.32/5.46)	goat (.55)	beaver (.67)	wallaby (.37)	anteater (.45)
rat (5.30/5.70)	cat (.55)	antelope (.67)	gorilla (.37)	panda (.45)

cheetah (5.28/5.34)	deer (.55)	porcupine (.67)	whale (.37)	dolphin (.45)
badger (5.25/5.19)	possum (.55)	leopard (.66)	pig (.37)	possum (.45)
hare (5.23/5.62)	cheetah (.55)	raccoon (.66)	antelope (.37)	gorilla (.44)
frog (5.21/5.11)	wolf (.55)	hyena (.65)	rat (.36)	tortoise (.44)
snake (5.17/4.89)	dog (.55)	cheetah (.65)	porpoise (.36)	cheetah (.44)
gerbil (5.17/4.69)	rat (.54)	porpoise (.65)	warthog (.36)	lizard (.44)
camel (5.04/5.51)	leopard (.54)	zebra (.65)	wildebeest (.36)	zebra (.44)
crocodile (4.96/5.17)	lizard (.54)	baboon (.65)	goat (.36)	fox (.43)
panther (4.90/5.46)	gerbil (.53)	rhinoceros (.65)	horse (.35)	lemur (.43)
koala (4.89/5.21)	chipmunk (.53)	llama (.64)	reindeer (.35)	moose (.43)
alligator (4.81/4.94)	chimpanzee (.53)	goat (.64)	meerkat (.35)	koala (.43)
orangutan (4.81/5.22)	wombat (.53)	dingo (.64)	armadillo (.35)	squirrel (.43)
mole (4.79/5.00)	orangutan (.53)	chipmunk (.64)	elk (.35)	beaver (.43)
otter (4.77/5.26)	terrapin (.52)	jaguar (.63)	cow (.35)	hyena (.43)
seal (4.77/5.27)	meerkat (.52)	ape (.63)	cougar (.35)	rat (.43)
dolphin (4.75/4.51)	jackal (.52)	pig (.63)	wolf (.35)	badger (.42)
jaguar (4.75/5.38)	elk (.52)	wombat (.63)	frog (.35)	giraffe (.42)
tortoise (4.75/5.30)	panda (.51)	dolphin (.63)	mouse (.35)	hippopotamus (.42)
whale (4.74/4.98)	moose (.51)	whale (.62)	mongoose (.35)	kangaroo (.42)

beaver (4.70/4.91)	kangaroo (.51)	hedgehog (.62)	hamster (.35)	antelope (.42)
baboon (4.68/5.24)	armadillo (.51)	python (.62)	dolphin (.35)	hamster (.42)
buffalo (4.68/5.29)	pig (.51)	terrapin (.62)	llama (.34)	baboon (.42)
ox (4.68/5.30)	turtle (.51)	impala (.62)	tortoise (.34)	ape (.42)
gazelle (4.65/5.29)	gorilla (.51)	salamander (.62)	coyote (.34)	leopard (.42)
mule (4.62/5.28)	ape (.50)	deer (.62)	buffalo (.34)	toad (.41)
reindeer (4.60/5.54)	bear (.50)	lynx (.62)	bear (.34)	puma (.41)
antelope (4.57/5.46)	koala (.50)	buffalo (.62)	kangaroo (.33)	mink (.41)
weasel (4.57/5.21)	otter (.50)	dog (.61)	squirrel (.33)	terrapin (.40)
turtle (4.55/4.45)	fox (.50)	warthog (.61)	jackal (.33)	rhinoceros (.40)
ass (4.49/5.53)	badger (.50)	armadillo (.61)	crocodile (.33)	frog (.40)
goldfish (4.47/4.10)	salamander (.50)	rat (.61)	panda (.33)	polecat (.40)
puma (4.47/5.23)	beaver (.50)	puma (.61)	turtle (.33)	crocodile (.39)
bison (4.45/5.02)	coyote (.50)	polecat (.61)	shrew (.33)	buffalo (.39)
wildebeest (4.42/4.71)	shark (.50)	vole (.61)	stoat (.32)	porpoise (.39)
hyena (4.42/4.98)	toad (.49)	panda (.61)	weasel (.31)	dingo (.39)
shark (4.40/4.20)	hedgehog (.49)	kangaroo (.61)	tiger (.31)	hedgehog (.39)
lizard (4.40/4.60)	porcupine (.49)	orangutan (.60)	sloth (.31)	armadillo (.39)
raccoon (4.40/4.53)	mouse (.49)	elk (.60)	wombat (.31)	porcupine (.38)



walrus (4.40/4.55)	crocodile (.48)	skunk (.60)	donkey (.31)	donkey (.38)
ferret (4.38/4.85)	porpoise (.48)	wildebeest (.60)	hedgehog (.31)	python (.38)
moose (4.38/5.02)	whale (.48)	mongoose (.59)	moose (.31)	chinchilla (.37)
bat (4.34/4.30)	sloth (.48)	shark (.59)	possum (.31)	impala (.37)
skunk (4.31/4.28)	impala (.48)	badger (.59)	leopard (.30)	stoat (.37)
impala (4.30/4.36)	ferret (.48)	koala (.59)	bison (.30)	meerkat (.37)
lynx (4.30/4.76)	hamster (.47)	meerkat (.58)	vole (.30)	bear (.37)
shrew (4.29/4.66)	dolphin (.47)	walrus (.58)	koala (.30)	tapir (.37)
stoat (4.29/5.04)	shrew (.47)	hare (.58)	pony (.30)	lynx (.36)
mink (4.22/4.50)	buffalo (.46)	muskrat (.58)	chipmunk (.30)	ocelot (.36)
toad (4.21/4.74)	tortoise (.46)	sheep (.57)	shark (.30)	ferret (.36)
jackal (4.17/4.90)	cow (.46)	hamster (.57)	lizard (.30)	shark (.36)
aardvark (4.12/3.69)	reindeer (.46)	newt (.57)	impala (.29)	camel (.36)
llama (4.08/4.83)	walrus (.46)	chimpanzee (.56)	mule (.29)	wildebeest (.35)
wombat (4.07/4.09)	skunk (.46)	donkey (.56)	jellyfish (.29)	walrus (.35)
cougar (4.06/4.83)	gazelle (.46)	jackal (.56)	salamander (.28)	salamander (.35)
chinchilla (4.05/4.58)	crab (.46)	bison (.55)	otter (.28)	lion (.35)
armadillo (4.04/4.56)	donkey (.45)	reindeer (.55)	mink (.28)	goldfish (.35)
iguana (4.04/3.90)	zebra (.45)	goldfish (.55)	beaver (.28)	wombat (.35)

octopus (4.04/3.84)	octopus (.45)	cow (.55)	skunk (.27)	warthog (.34)
vole (4.02/4.78)	mink (.45)	ocelot (.55)	octopus (.27)	mongoose (.34)
coyote (4.02/4.48)	bat (.44)	eel (.55)	zebra (.27)	muskrat (.33)
crab (4.02/4.62)	bison (.44)	gazelle (.54)	hare (.27)	gerbil (.33)
chipmunk (4.00/4.46)	vole (.44)	chinchilla (.54)	walrus (.26)	chipmunk (.32)
python (4.00/4.32)	pony (.44)	lion (.54)	panther (.26)	jackal (.31)
porcupine (3.98/4.51)	chinchilla (.43)	gerbil (.53)	porcupine (.26)	ox (.31)
elk (3.98/5.00)	panther (.43)	shrew (.52)	sheep (.26)	mouse (.31)
lemur (3.94/4.68)	shrimp (.43)	aardvark (.51)	fox (.26)	vole (.31)
wallaby (3.86/4.63)	horse (.43)	weasel (.51)	bat (.26)	reindeer (.30)
possum (3.86/4.15)	llama (.42)	gopher (.51)	jaguar (.24)	shrew (.30)
polecat (3.85/4.33)	ass (.42)	mink (.50)	toad (.24)	gazelle (.29)
anteater (3.85/4.11)	jellyfish (.41)	jellyfish (.50)	gazelle (.24)	pony (.29)
warthog (3.80/4.13)	lobster (.41)	crab (.50)	terrapin (.23)	lobster (.28)
dingo (3.78/4.35)	hare (.41)	salmon (.49)	python (.23)	mule (.28)
mongoose (3.76/4.22)	camel (.41)	trout (.49)	salmon (.23)	trout (.28)
gopher (3.73/3.92)	newt (.41)	bear (.48)	goldfish (.23)	eel (.28)
porpoise (3.71/4.50)	goldfish (.40)	camel (.48)	crab (.22)	sloth (.27)
ocelot (3.69/4.67)	trout (.39)	lobster (.48)	camel (.21)	jellyfish (.27)

salmon (3.68/4.13)	puma (.39)	octopus (.46)	shrimp (.21)	newt (.26)
meerkat (3.67/3.71)	eel (.38)	ox (.46)	trout (.21)	hare (.26)
lemming (3.66/4.29)	salmon (.37)	sloth (.45)	seal (.20)	aardvark (.26)
terrapin (3.61/4.14)	mule (.37)	horse (.44)	lobster (.20)	crab (.26)
sloth (3.60/4.16)	mole (.37)	herring (.44)	ass (.19)	gopher (.25)
newt (3.55/3.98)	lemming (.37)	shrimp (.43)	puma (.18)	salmon (.25)
jellyfish (3.53/3.72)	sheep (.36)	mouse (.42)	aardvark (.18)	weasel (.25)
gnu (3.52/3.95)	python (.36)	cod (.41)	mole (.18)	shrimp (.24)
muskrat (3.50/3.60)	seal (.35)	mule (.40)	newt (.17)	octopus (.24)
lobster (3.49/4.02)	lynx (.35)	pony (.39)	eel (.15)	herring (.21)
tapir (3.27/4.08)	gopher (.34)	mole (.38)	gopher (.13)	seal (.20)
eel (3.17/3.59)	aardvark (.33)	gnu (.38)	ox (.13)	gnu (.19)
salamander (3.16/3.69)	jaguar (.31)	ferret (.38)	lemming (.12)	lemming (.18)
shrimp (3.15/3.84)	ox (.28)	haddock (.38)	lynx (.12)	bat (.18)
herring (3.12/3.92)	herring (.28)	lemming (.36)	sole (.11)	haddock (.16)
trout (3.08/3.90)	haddock (.27)	bat (.30)	herring (.08)	cod (.12)
cod (3.04/4.10)	gnu (.24)	ass (.30)	haddock (.07)	mole (.12)
haddock (2.94/4.37)	cod (.23)	seal (.26)	gnu (.07)	ass (.11)
sole (2.76/3.70)	sole (.16)	sole (.06)	cod (.06)	sole (.04)

---

Table A17

*Exemplars of the category birds sorted by typicality, according to the younger adults, or similarity to the prototype with the actual values in parentheses (older adults' ratings after the forward slash) for English*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
crow (6.58/6.46)	oystercatcher (.70)	heron (.78)	parrot (.59)	owl (.68)
blackbird (6.53/6.65)	wagtail (.70)	robin (.77)	mynah (.59)	falcon (.66)
robin (6.51/6.57)	chaffinch (.68)	owl (.76)	cockatiel (.55)	robin (.65)
pigeon (6.38/6.42)	parakeet (.68)	wren (.75)	parakeet (.53)	pelican (.65)
sparrow (6.38/6.83)	jackdaw (.67)	curlew (.75)	crow (.51)	heron (.64)
seagull (6.17/6.54)	cormorant (.66)	kingfisher (.74)	jackdaw (.49)	hummingbird (.63)
bluetit (6.06/5.75)	woodpecker (.66)	sparrow (.72)	pigeon (.48)	goose (.63)
duck (6.04/6.17)	mynah (.66)	falcon (.72)	eagle (.48)	finch (.62)
budgie (5.98/6.11)	cockatiel (.63)	sandpiper (.71)	owl (.48)	pigeon (.61)
eagle (5.98/5.66)	lovebird (.62)	blackbird (.71)	sparrow (.48)	sparrow (.61)
parrot (5.96/5.74)	yellowhammer (.62)	gannet (.71)	blackbird (.46)	osprey (.60)
swallow (5.87/6.11)	skylark (.61)	buzzard (.70)	skylark (.46)	blackbird (.60)
owl (5.83/5.98)	owl (.60)	pelican (.70)	chaffinch (.46)	kestrel (.60)
hawk (5.81/5.71)	thrush (.60)	oystercatcher (.70)	duck (.46)	woodpecker (.59)

magpie (5.79/6.04)	parrot (.60)	finch (.70)	pheasant (.46)	starling (.59)
starling (5.76/6.58)	sparrow (.60)	wagtail (.70)	budgie (.46)	bluebird (.58)
swan (5.73/5.96)	cuckoo (.59)	kestrel (.69)	wagtail (.45)	kingfisher (.58)
canary (5.72/6.02)	kingfisher (.58)	starling (.69)	chicken (.45)	buzzard (.58)
dove (5.68/5.83)	blackbird (.58)	swan (.68)	cormorant (.44)	seagull (.58)
thrush (5.65/6.44)	stork (.58)	raven (.68)	dove (.43)	swan (.58)
bluebird (5.63/5.02)	curlew (.58)	seagull (.68)	stork (.43)	flamingo (.57)
finch (5.60/5.90)	gannet (.57)	woodpecker (.68)	hawk (.43)	cormorant (.57)
raven (5.55/5.78)	starling (.57)	hummingbird (.67)	cuckoo (.42)	parrot (.57)
chaffinch (5.35/6.00)	tern (.57)	widgeon (.67)	lark (.42)	curlew (.57)
wren (5.32/5.80)	pheasant (.57)	jackdaw (.67)	woodpecker (.42)	wren (.56)
woodpecker (5.30/5.41)	lark (.57)	osprey (.67)	swallow (.42)	pheasant (.55)
cuckoo (5.25/5.70)	sandpiper (.57)	parrot (.67)	vulture (.42)	gannet (.55)
kingfisher (5.20/5.37)	dove (.56)	pigeon (.66)	albatross (.41)	oystercatcher (.55)
lark (5.20/6.02)	magpie (.55)	mynah (.66)	lovebird (.41)	sandpiper (.53)
chicken (5.15/6.17)	buzzard (.54)	chaffinch (.66)	thrush (.41)	crow (.53)
falcon (5.15/5.23)	crow (.54)	goose (.66)	magpie (.41)	raven (.53)
turkey (5.13/5.67)	coot (.54)	cormorant (.65)	oystercatcher (.40)	turkey (.53)
rook (5.09/5.79)	eagle (.54)	skylark (.65)	coot (.39)	magpie (.51)

hummingbird (5.04/4.77)	widgeon (.54)	flamingo (.64)	gannet (.38)	duck (.51)
kestrel (5.02/5.32)	albatross (.53)	puffin (.64)	yellowhammer (.38)	peacock (.51)
stork (5.02/5.23)	duck (.53)	peacock (.64)	kestrel (.38)	cockatiel (.51)
peacock (5.00/5.29)	hawk (.52)	bluebird (.64)	osprey (.38)	puffin (.50)
goose (4.98/5.48)	heron (.52)	yellowhammer (.63)	starling (.37)	wagtail (.50)
swift (4.98/5.31)	osprey (.52)	magpie (.63)	ostrich (.36)	vulture (.50)
pheasant (4.89/5.46)	vulture (.51)	crow (.62)	finch (.36)	condor (.50)
puffin (4.89/5.36)	finch (.51)	budgie (.62)	dodo (.36)	widgeon (.50)
nightingale (4.87/5.28)	kestrel (.51)	partridge (.62)	kite (.36)	jackdaw (.49)
jackdaw (4.86/5.72)	hummingbird (.51)	cockatiel (.61)	raven (.36)	skylark (.49)
osprey (4.86/4.85)	falcon (.50)	stork (.60)	falcon (.36)	budgie (.49)
heron (4.79/5.21)	pigeon (.49)	pheasant (.59)	seagull (.35)	chaffinch (.49)
buzzard (4.77/4.74)	partridge (.49)	nightingale (.58)	peacock (.35)	mynah (.48)
curlew (4.72/5.39)	swallow (.49)	cuckoo (.58)	partridge (.35)	parakeet (.48)
skylark (4.71/5.83)	raven (.48)	duck (.58)	hummingbird (.35)	eagle (.48)
grouse (4.65/5.54)	pelican (.48)	parakeet (.57)	widgeon (.35)	grouse (.47)
gannet (4.64/4.98)	flamingo (.47)	penguin (.57)	buzzard (.34)	emu (.47)
pelican (4.64/4.63)	goose (.46)	lovebird (.57)	swan (.34)	partridge (.47)
cockatiel (4.64/4.30)	wren (.46)	grouse (.56)	kingfisher (.34)	stork (.45)

wagtail (4.63/5.44)	chicken (.43)	eagle (.56)	flamingo (.32)	yellowhammer (.44)
vulture (4.60/5.17)	bluebird (.43)	vulture (.56)	crane (.32)	penguin (.44)
parakeet (4.57/4.59)	ostrich (.43)	tern (.56)	heron (.32)	tern (.44)
partridge (4.55/5.09)	seagull (.43)	thrush (.56)	curlew (.32)	cuckoo (.43)
flamingo (4.55/5.02)	grouse (.42)	coot (.55)	goose (.31)	hawk (.43)
oystercatcher (4.54/5.33)	swan (.42)	emu (.54)	tern (.31)	nightingale (.43)
albatross (4.51/4.72)	rook (.41)	condor (.52)	puffin (.30)	chicken (.42)
jay (4.48/4.79)	budgie (.41)	turkey (.51)	grouse (.30)	lovebird (.41)
penguin (4.45/4.96)	kite (.41)	jay (.50)	turkey (.29)	thrush (.40)
cormorant (4.45/5.25)	nightingale (.41)	hawk (.48)	nightingale (.29)	coot (.40)
yellowhammer (4.39/4.90)	crane (.40)	ostrich (.47)	pelican (.29)	jay (.37)
mynah (4.36/4.58)	puffin (.39)	chicken (.46)	canary (.29)	ostrich (.34)
kite (4.34/4.70)	swift (.39)	dodo (.44)	penguin (.28)	kite (.34)
tern (4.32/5.02)	peacock (.37)	teal (.43)	rook (.27)	dove (.33)
ostrich (4.28/5.15)	penguin (.37)	albatross (.43)	wren (.27)	albatross (.31)
crane (4.25/4.86)	robin (.36)	dove (.41)	sandpiper (.26)	crane (.30)
coot (4.10/4.56)	teal (.36)	canary (.40)	condor (.25)	teal (.30)
condor (4.02/4.44)	dodo (.34)	kite (.39)	robin (.25)	dodo (.29)
lovebird (4.02/4.82)	canary (.34)	lark (.34)	jay (.25)	canary (.29)

emu (3.98/4.47)	jay (.33)	crane (.31)	swift (.25)	lark (.19)
widgeon (3.83/4.08)	condor (.32)	rook (.26)	bluebird (.23)	rook (.14)
sandpiper (3.78/4.96)	turkey (.31)	swallow (.20)	teal (.18)	swallow (.11)
dodo (3.59/3.14)	emu (.16)	swift (.08)	emu (.04)	swift (.01)
teal (3.39/4.20)	bluetit (-)	bluetit (-)	bluetit (-)	bluetit (-)

---



Table A18

*Exemplars of the category clothes sorted by typicality, according to the younger adults, or similarity to the prototype with the actual values in parentheses (older adults' ratings after the forward slash) for English*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
trousers (6.91/6.93)	dungarees (.77)	pants (.76)	shirt (.66)	jeans (.65)
jacket (6.89/6.89)	shirt (.75)	trousers (.76)	dress (.64)	dress (.59)
jeans (6.89/6.35)	jacket (.73)	jacket (.75)	pants (.60)	pants (.54)
jumper (6.87/6.68)	dress (.73)	blouse (.73)	sarong (.58)	trousers (.53)
coat (6.75/6.83)	leotard (.73)	corduroys (.73)	jacket (.58)	knickers (.51)
shirt (6.74/6.87)	blouse (.73)	cardigan (.73)	sweater (.53)	blouse (.50)
dress (6.68/6.76)	pullover (.73)	blazer (.73)	suit (.52)	sweater (.50)
skirt (6.62/6.62)	pants (.72)	sweater (.73)	trousers (.52)	dungarees (.49)
sweatshirt (6.57/6.15)	trousers (.72)	tights (.72)	nightdress (.52)	shirt (.49)
sweater (6.51/6.65)	nightdress (.71)	jeans (.72)	bathrobe (.51)	corduroys (.49)
sock (6.38/6.52)	sweater (.71)	scarf (.72)	dungarees (.51)	jacket (.49)
pullover (6.25/6.46)	sarong (.70)	waistcoat (.72)	pyjamas (.51)	nightdress (.48)
suit (6.25/6.77)	skirt (.69)	dress (.71)	blouse (.50)	scarf (.48)
knickers (6.23/6.44)	suspender (.69)	knickers (.70)	coat (.50)	robe (.47)

cardigan (6.17/6.30)	bathrobe (.68)	dungarees (.70)	raincoat (.49)	tuxedo (.46)
jersey (6.13/6.17)	sweatshirt (.67)	shirt (.69)	jeans (.49)	shoe (.46)
blouse (6.04/6.57)	scarf (.67)	jodhpurs (.68)	robe (.49)	tights (.46)
pants (6.04/6.48)	parka (.67)	slacks (.68)	gown (.47)	sari (.46)
anorak (6.00/6.28)	chemise (.67)	raincoat (.68)	overcoat (.46)	salopettes (.46)
shorts (6.00/6.13)	raincoat (.67)	overcoat (.68)	chemise (.46)	overalls (.45)
pyjamas (5.70/6.47)	jerkin (.66)	sweatshirt (.67)	swimsuit (.45)	smock (.45)
fleece (5.68/4.40)	coat (.66)	robe (.67)	leotard (.45)	bathrobe (.45)
hat (5.64/6.22)	jodhpurs (.65)	tuxedo (.67)	skirt (.45)	sweatshirt (.44)
overcoat (5.60/6.41)	waistcoat (.64)	culottes (.66)	scarf (.45)	raincoat (.43)
shoe (5.51/6.70)	smock (.64)	smock (.66)	hat (.44)	sarong (.43)
scarf (5.45/5.83)	overcoat (.64)	cape (.66)	duffel (.44)	coat (.43)
raincoat (5.43/6.60)	hat (.64)	petticoat (.66)	shawl (.44)	chemise (.43)
tie (5.43/6.21)	shawl (.64)	cagoule (.65)	shorts (.44)	cardigan (.42)
brassiere (5.40/6.26)	bloomers (.63)	leotard (.65)	pullover (.43)	slacks (.42)
vest (5.38/5.94)	underskirt (.63)	coat (.65)	knickers (.43)	parka (.42)
nightdress (5.36/6.61)	gown (.62)	pullover (.65)	overalls (.42)	overcoat (.42)
tights (5.23/6.45)	cagoule (.62)	suspender (.65)	sweatshirt (.42)	shawl (.42)
cap (5.21/5.93)	suit (.62)	nightdress (.65)	vest (.42)	culottes (.42)

glove (5.21/6.30)	knickers (.62)	parka (.65)	jodhpurs (.41)	swimsuit (.42)
briefs (5.19/6.21)	shorts (.61)	mackintosh (.65)	shoe (.41)	kilt (.41)
duffel (5.19/5.26)	slacks (.61)	chemise (.64)	slacks (.41)	jodhpurs (.41)
dungarees (5.04/5.76)	duffel (.61)	poncho (.64)	tights (.40)	sandal (.41)
blazer (5.00/6.33)	vest (.61)	skirt (.63)	handkerchief (.40)	poncho (.41)
corduroys (4.98/5.59)	salopettes (.60)	corset (.63)	bloomers (.39)	sock (.40)
cagoule (4.94/5.30)	jeans (.60)	underskirt (.63)	tuxedo (.39)	gown (.40)
waistcoat (4.92/5.60)	pyjamas (.60)	sarong (.63)	parka (.39)	petticoat (.40)
boot (4.91/6.15)	robe (.59)	shawl (.63)	waistcoat (.38)	handkerchief (.39)
legging (4.87/5.10)	blazer (.58)	hat (.63)	bikini (.36)	bikini (.39)
slacks (4.79/6.39)	tie (.58)	gown (.62)	underskirt (.36)	waistcoat (.38)
trainer (4.75/5.61)	swimsuit (.57)	bathrobe (.62)	corset (.36)	shorts (.38)
belt (4.67/5.56)	beret (.57)	salopettes (.62)	smock (.36)	blazer (.38)
stocking (4.64/6.49)	tights (.56)	shorts (.62)	suspender (.36)	mackintosh (.38)
kilt (4.58/5.44)	kilt (.56)	balaclava (.61)	sari (.36)	duffel (.37)
mackintosh (4.57/5.87)	corset (.55)	kilt (.61)	tie (.35)	fleece (.37)
parka (4.55/4.94)	balaclava (.55)	handkerchief (.61)	cagoule (.35)	leotard (.37)
bikini (4.53/5.02)	poncho (.55)	sock (.61)	salopettes (.35)	pullover (.36)
bathrobe (4.51/5.23)	brassiere (.55)	beret (.60)	jerkin (.34)	cagoule (.36)

slipper (4.51/5.92)	petticoat (.54)	sari (.59)	petticoat (.34)	hat (.35)
overalls (4.40/5.68)	stiletto (.54)	overalls (.59)	earmuffs (.33)	balaclava (.35)
swimsuit (4.36/5.40)	apron (.54)	anorak (.59)	apron (.33)	earmuffs (.34)
cloak (4.21/4.45)	sock (.53)	brassiere (.58)	glove (.33)	skirt (.33)
mitten (4.21/4.82)	belt (.53)	garter (.58)	sock (.33)	trunks (.33)
tuxedo (4.21/4.15)	handkerchief (.53)	bikini (.58)	poncho (.32)	anorak (.32)
slip (4.17/5.38)	overalls (.53)	fleece (.57)	kilt (.32)	brassiere (.32)
underskirt (4.15/5.83)	tuxedo (.52)	apron (.56)	cloak (.31)	corset (.32)
culottes (4.06/4.41)	cap (.52)	sweatband (.55)	beret (.31)	cape (.32)
petticoat (4.06/5.41)	shoe (.52)	girdle (.55)	jumper (.31)	slipper (.32)
robe (4.04/4.98)	bikini (.51)	sandal (.55)	brassiere (.31)	apron (.32)
jodhpurs (4.02/4.26)	earmuffs (.50)	shoe (.54)	belt (.30)	underskirt (.31)
trunks (4.02/5.62)	cloak (.50)	duffel (.54)	stiletto (.30)	jersey (.31)
gown (3.94/5.25)	sandal (.49)	slipper (.54)	balaclava (.30)	bloomers (.30)
jerkin (3.83/5.29)	jumper (.49)	swimsuit (.53)	sweatband (.29)	sweatband (.30)
sandal (3.79/5.61)	girdle (.49)	stiletto (.53)	blazer (.29)	beret (.29)
shawl (3.79/4.72)	glove (.48)	vest (.53)	muffler (.29)	stocking (.29)
balaclava (3.77/4.61)	muffler (.47)	earmuffs (.52)	trunks (.28)	legging (.29)
chemise (3.77/4.35)	fleece (.47)	jerkin (.50)	sandal (.27)	jerkin (.29)

leotard (3.75/4.43)	sari (.46)	mitten (.50)	cap (.26)	suspender (.29)
salopettes (3.70/3.89)	garter (.44)	legging (.48)	legging (.26)	garter (.28)
sari (3.67/4.56)	slipper (.44)	jersey (.47)	jersey (.26)	girdle (.28)
sarong (3.47/3.91)	legging (.43)	bloomers (.46)	mitten (.25)	vest (.27)
stiletto (3.46/3.56)	braces (.42)	glove (.46)	garter (.25)	braces (.27)
apron (3.40/5.69)	cardigan (.42)	trunks (.45)	slipper (.24)	mitten (.27)
suspender (3.34/4.04)	mitten (.40)	cloak (.45)	cardigan (.24)	cloak (.26)
beret (3.27/4.93)	sweatband (.40)	boot (.42)	briefs (.23)	boot (.24)
cape (3.26/4.35)	boot (.40)	belt (.42)	braces (.22)	glove (.23)
braces (3.19/4.74)	anorak (.39)	muffler (.40)	anorak (.21)	stiletto (.22)
corset (3.19/4.41)	slip (.37)	suit (.39)	fleece (.21)	suit (.22)
smock (3.18/4.15)	trunks (.37)	braces (.35)	boot (.21)	muffler (.20)
earmuffs (3.17/3.82)	cape (.37)	cap (.34)	culottes (.19)	belt (.20)
garter (3.16/3.89)	jersey (.37)	jumper (.33)	cape (.19)	briefs (.18)
bloomers (3.12/4.60)	culottes (.36)	stocking (.31)	stocking (.18)	jumper (.17)
muffler (3.02/4.96)	stocking (.35)	briefs (.30)	girdle (.18)	cap (.14)
poncho (3.00/4.12)	briefs (.30)	tie (.29)	slip (.17)	tie (.09)
clog (2.89/3.85)	trainer (.22)	slip (.21)	trainer (.12)	slip (.08)
girdle (2.83/4.66)	clog (.19)	trainer (.14)	mackintosh (.11)	trainer (.07)

sweatband (2.74/3.94)	mackintosh (.15)	clog (.12)	clog (.10)	clog (.06)
handkerchief (2.51/5.44)	corduroys (-)	pyjamas (-)	corduroys (-)	pyjamas (-)

---

Table A19

*Exemplars of the category flowers sorted by typicality, according to the younger adults, or similarity to the prototype with the actual values in parentheses (older adults' ratings after the forward slash) for English*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
rose (6.83/6.87)	nasturtium (.81)	hydrangea (.82)	nasturtium (.60)	orchid (.69)
daffodil (6.60/6.72)	geranium (.74)	hollyhock (.80)	freesia (.56)	tulip (.66)
daisy (6.43/6.63)	petunia (.72)	nasturtium (.79)	chrysanthemum (.56)	chrysanthemum (.65)
tulip (6.13/6.57)	azalea (.72)	clematis (.78)	rose (.55)	dahlia (.65)
sunflower (6.11/5.55)	gladiolus (.71)	foxglove (.78)	honeysuckle (.54)	hydrangea (.65)
carnation (6.06/6.35)	celandine (.71)	geranium (.76)	orchid (.54)	daffodil (.64)
poppy (5.91/6.20)	begonia (.71)	petunia (.76)	azalea (.53)	marigold (.62)
gypsophila (5.85/5.06)	freesia (.70)	dahlia (.76)	carnation (.52)	carnation (.61)
pansy (5.81/6.39)	hydrangea (.70)	catkin (.75)	clematis (.51)	geranium (.61)
snowdrop (5.76/6.46)	chrysanthemum (.69)	lily (.75)	geranium (.50)	lily (.60)
bluebell (5.71/6.20)	honeysuckle (.69)	marigold (.74)	daffodil (.50)	hollyhock (.59)
dandelion (5.66/5.89)	clematis (.68)	cowslip (.74)	begonia (.49)	lavender (.59)
chrysanthemum (5.65/6.33)	foxglove (.68)	rhododendron (.74)	foxglove (.49)	gladiolus (.59)
primrose (5.61/6.36)	hollyhock (.67)	chrysanthemum (.73)	gladiolus (.48)	petunia (.58)

buttercup (5.58/6.20)	cowslip (.67)	azalea (.73)	hydrangea (.48)	clematis (.57)
crocus (5.57/6.20)	snapdragon (.65)	lavender (.73)	hollyhock (.47)	daisy (.57)
marigold (5.52/6.43)	fuchsia (.63)	crocus (.73)	crocus (.47)	snowdrop (.57)
geranium (5.45/6.28)	buttercup (.63)	daffodil (.73)	fuchsia (.46)	azalea (.56)
lily (5.42/6.06)	pansy (.62)	dandelion (.73)	petunia (.46)	foxglove (.56)
foxglove (5.39/5.64)	catkin (.62)	snowdrop (.72)	pansy (.46)	catkin (.56)
hyacinth (5.37/6.02)	marigold (.61)	magnolia (.72)	buttercup (.46)	lilac (.55)
orchid (5.26/5.28)	carnation (.61)	freesia (.72)	cowslip (.46)	rhododendron (.55)
violet (5.14/5.91)	orchid (.61)	lilac (.72)	celandine (.45)	fuchsia (.54)
honeysuckle (5.10/5.74)	crocus (.61)	buttercup (.72)	snapdragon (.45)	magnolia (.54)
magnolia (5.08/5.08)	snowdrop (.60)	orchid (.71)	catkin (.45)	nasturtium (.54)
fuchsia (5.07/5.74)	rose (.60)	honeysuckle (.71)	jasmine (.44)	dandelion (.53)
rhododendron (5.06/5.79)	daffodil (.59)	primrose (.71)	hyacinth (.44)	hyacinth (.53)
dahlia (4.94/6.04)	dahlia (.59)	celandine (.71)	snowdrop (.42)	crocus (.53)
freesia (4.92/5.68)	lilac (.58)	tulip (.69)	lavender (.39)	cowslip (.52)
azalea (4.91/5.61)	hyacinth (.57)	anemone (.69)	dahlia (.39)	violet (.51)
iris (4.84/5.96)	lavender (.57)	bluebell (.69)	daisy (.39)	freesia (.50)
petunia (4.83/5.81)	rhododendron (.56)	violet (.68)	marigold (.39)	bluebell (.50)
lavender (4.81/5.35)	anemone (.54)	daisy (.68)	lilac (.39)	celandine (.49)



nasturtium (4.72/6.08)	magnolia (.54)	hyacinth (.68)	dandelion (.38)	jasmine (.49)
lilac (4.70/5.72)	dandelion (.53)	gladiolus (.67)	poppy (.38)	sunflower (.49)
hydrangea (4.67/6.04)	poppy (.52)	snapdragon (.66)	anemone (.37)	primrose (.48)
gladiolus (4.61/5.50)	lily (.51)	fuchsia (.66)	tulip (.37)	anemone (.47)
begonia (4.55/5.61)	jasmine (.51)	thistle (.65)	lily (.36)	edelweiss (.47)
clematis (4.46/5.56)	daisy (.50)	jasmine (.63)	magnolia (.35)	honeysuckle (.45)
cowslip (4.41/5.39)	edelweiss (.48)	edelweiss (.63)	sunflower (.35)	poppy (.45)
edelweiss (4.21/4.74)	primrose (.48)	heather (.62)	edelweiss (.34)	iris (.45)
snapdragon (4.18/5.26)	bluebell (.47)	carnation (.61)	rhododendron (.34)	buttercup (.44)
wallflower (4.16/5.98)	sunflower (.47)	sunflower (.59)	bluebell (.27)	thistle (.42)
heather (4.11/5.74)	tulip (.46)	iris (.58)	lupin (.27)	snapdragon (.41)
jasmine (4.07/4.90)	violet (.45)	pansy (.56)	iris (.27)	pansy (.41)
anemone (4.05/5.67)	iris (.41)	poppy (.48)	primrose (.26)	heather (.37)
lupin (4.03/6.06)	heather (.41)	wallflower (.36)	violet (.25)	wallflower (.25)
thistle (4.00/5.60)	thistle (.40)	rose (.20)	thistle (.24)	rose (.13)
hollyhock (3.84/5.35)	lupin (.39)	begonia (-)	wallflower (.22)	begonia (-)
marguerite (3.83/5.92)	wallflower (.37)	gypsophila (-)	marguerite (.20)	gypsophila (-)
catkin (3.22/4.63)	marguerite (.36)	lupin (-)	heather (.19)	lupin (-)
celandine (3.12/4.71)	gypsophila (-)	marguerite (-)	gypsophila (-)	marguerite (-)

---

Table A20

*Exemplars of the category fruit sorted by typicality, according to the younger adults, or similarity to the prototype with the actual values in parentheses (older adults' ratings after the forward slash) for English*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
apple (6.96/6.91)	lychee (.75)	pear (.82)	pomegranate (.52)	mango (.66)
orange (6.91/6.89)	redcurrant (.71)	apricot (.81)	pineapple (.50)	berry (.65)
banana (6.74/6.81)	blackcurrant (.71)	berry (.81)	blackcurrant (.50)	apple (.64)
strawberry (6.45/6.59)	pistachio (.70)	strawberry (.79)	cherry (.49)	pear (.63)
pear (6.44/6.57)	pomegranate (.69)	mango (.78)	guava (.49)	melon (.62)
peach (6.30/6.36)	nectarine (.69)	peach (.77)	lychee (.48)	grape (.62)
lemon (6.25/6.50)	guava (.67)	blueberry (.76)	banana (.47)	strawberry (.61)
grape (6.17/6.30)	peach (.67)	melon (.76)	apple (.47)	tomato (.60)
tangerine (6.13/6.34)	pecan (.67)	tomato (.75)	gooseberry (.47)	guava (.59)
raspberry (5.98/6.43)	apricot (.66)	avocado (.74)	grape (.47)	peach (.59)
satsuma (5.92/5.67)	elderberry (.66)	apple (.74)	melon (.45)	nectarine (.57)
pineapple (5.72/5.74)	gooseberry (.66)	almond (.74)	nectarine (.44)	blueberry (.57)
melon (5.68/6.24)	hazelnut (.65)	grape (.74)	breadfruit (.43)	grapefruit (.57)
berry (5.66/5.96)	avocado (.65)	nectarine (.73)	pear (.43)	pomegranate (.56)

nectarine (5.62/5.58)	melon (.64)	rhubarb (.73)	pistachio (.43)	pineapple (.54)
plum (5.58/6.37)	pineapple (.64)	pineapple (.73)	peach (.43)	apricot (.54)
grapefruit (5.57/6.37)	cherry (.64)	clementine (.73)	grapefruit (.42)	lychee (.53)
cherry (5.51/5.85)	breadfruit (.63)	loganberry (.73)	redcurrant (.40)	watermelon (.53)
kiwi (5.42/5.19)	cranberry (.62)	guava (.72)	pecan (.40)	banana (.52)
apricot (5.30/5.38)	kumquat (.62)	raspberry (.72)	watermelon (.40)	papaya (.52)
mandarin (5.30/5.83)	raspberry (.62)	kumquat (.72)	papaya (.39)	clementine (.52)
blackcurrant (5.25/5.89)	lemon (.62)	satsuma (.71)	strawberry (.38)	loganberry (.52)
watermelon (5.15/5.51)	banana (.62)	pomegranate (.71)	avocado (.38)	avocado (.51)
clementine (5.11/5.72)	currant (.61)	pawpaw (.71)	apricot (.38)	pawpaw (.51)
lime (5.08/4.80)	watermelon (.61)	damson (.71)	berry (.38)	fig (.51)
blackberry (4.96/5.60)	apple (.60)	lychee (.70)	orange (.37)	almond (.50)
tomato (4.83/6.52)	orange (.60)	hazelnut (.70)	kiwi (.36)	kumquat (.50)
mango (4.72/4.53)	strawberry (.60)	grapefruit (.69)	kumquat (.36)	olive (.50)
bramble (4.51/5.67)	blueberry (.60)	lemon (.69)	currant (.36)	rhubarb (.50)
raisin (4.46/5.59)	grapefruit (.60)	persimmon (.68)	prune (.35)	breadfruit (.49)
blueberry (4.43/4.62)	damson (.59)	cranberry (.67)	raspberry (.35)	raspberry (.48)
gooseberry (4.38/5.78)	pear (.59)	dewberry (.67)	tomato (.34)	pistachio (.48)
damson (4.34/5.32)	plum (.58)	gooseberry (.67)	coconut (.34)	hazelnut (.47)

rhubarb (4.34/5.64)	papaya (.58)	currant (.67)	elderberry (.34)	satsuma (.47)
cranberry (4.31/4.96)	coconut (.57)	watermelon (.67)	plum (.34)	blackcurrant (.47)
redcurrant (4.31/5.50)	tomato (.56)	papaya (.67)	damson (.33)	cherry (.47)
sultana (4.26/5.46)	mango (.55)	pecan (.66)	raisin (.33)	pumpkin (.47)
prune (4.21/5.30)	raisin (.55)	blackcurrant (.66)	satsuma (.33)	coconut (.47)
greengage (4.19/5.00)	grape (.54)	pistachio (.66)	almond (.33)	pecan (.46)
currant (4.19/5.52)	satsuma (.53)	banana (.66)	hazelnut (.33)	lemon (.46)
avocado (4.15/4.87)	almond (.53)	olive (.65)	olive (.32)	raisin (.45)
fig (4.08/4.82)	walnut (.53)	fig (.65)	mango (.32)	damson (.45)
pomegranate (4.06/4.81)	tangerine (.52)	walnut (.64)	fig (.29)	gooseberry (.44)
elderberry (4.02/4.80)	pumpkin (.51)	tangerine (.63)	cranberry (.29)	persimmon (.44)
papaya (3.84/4.06)	lime (.51)	raisin (.63)	peanut (.28)	dewberry (.42)
loganberry (3.83/4.80)	olive (.50)	elderberry (.63)	rhubarb (.28)	lime (.41)
date (3.82/5.48)	rhubarb (.50)	redcurrant (.62)	pumpkin (.27)	walnut (.41)
guava (3.75/4.18)	berry (.49)	pumpkin (.62)	blueberry (.26)	cranberry (.41)
dewberry (3.67/3.80)	prune (.49)	lime (.62)	lemon (.26)	blackberry (.40)
pawpaw (3.66/4.28)	chestnut (.46)	coconut (.62)	walnut (.26)	currant (.40)
coconut (3.51/4.83)	sultana (.44)	blackberry (.61)	sultana (.23)	peanut (.40)
lychee (3.44/4.09)	peanut (.44)	cherry (.59)	tangerine (.22)	elderberry (.38)

pumpkin (3.40/4.10)	kiwi (.42)	sultana (.59)	lime (.22)	prune (.38)
kumquat (3.24/4.08)	blackberry (.41)	breadfruit (.55)	blackberry (.21)	mandarin (.38)
olive (3.19/4.40)	fig (.36)	mandarin (.55)	mandarin (.20)	sultana (.37)
breadfruit (3.15/3.48)	bramble (.35)	peanut (.51)	dewberry (.18)	tangerine (.36)
hazelnut (2.92/4.53)	mandarin (.34)	prune (.51)	acorn (.16)	acorn (.35)
almond (2.83/4.37)	clementine (.32)	acorn (.49)	chestnut (.15)	kiwi (.33)
persimmon (2.80/3.20)	acorn (.26)	orange (.49)	bramble (.13)	redcurrant (.33)
pecan (2.69/3.49)	persimmon (.26)	chestnut (.49)	clementine (.13)	orange (.32)
walnut (2.69/4.58)	dewberry (.25)	bramble (.46)	date (.12)	plum (.30)
chestnut (2.59/4.24)	date (.23)	plum (.46)	persimmon (.06)	bramble (.29)
pistachio (2.47/3.58)	greengage (-)	kiwi (.43)	greengage (-)	chestnut (.26)
acorn (2.44/4.39)	loganberry (-)	date (.06)	loganberry (-)	date (.08)
peanut (2.35/4.64)	pawpaw (-)	greengage (-)	pawpaw (-)	greengage (-)

---

Table A21

*Exemplars of the category furniture sorted by typicality, according to the younger adults, or similarity to the prototype with the actual values in parentheses (older adults' ratings after the forward slash) for English*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
chair (6.92/6.91)	sofa (.75)	sofa (.82)	sofa (.47)	sofa (.50)
table (6.87/6.93)	settee (.70)	settee (.75)	bookcase (.43)	carpet (.48)
armchair (6.77/6.83)	chair (.67)	dresser (.73)	sideboard (.43)	dresser (.46)
settee (6.68/6.55)	washstand (.66)	bookcase (.73)	carpet (.40)	futon (.46)
couch (6.64/6.33)	bed (.66)	couch (.71)	washstand (.40)	bookcase (.44)
sofa (6.64/6.63)	sideboard (.66)	ottoman (.70)	wardrobe (.39)	fireplace (.42)
wardrobe (6.49/6.83)	couch (.63)	fireplace (.69)	chair (.38)	settee (.42)
bed (6.45/6.91)	table (.61)	footstool (.69)	settee (.38)	washstand (.41)
desk (6.36/5.90)	footstool (.60)	divan (.67)	table (.36)	ottoman (.40)
stool (5.94/6.07)	dresser (.59)	washstand (.66)	footstool (.35)	sideboard (.40)
cabinet (5.91/5.96)	bookcase (.59)	sideboard (.66)	fireplace (.34)	wardrobe (.39)
cupboard (5.89/6.13)	desk (.58)	futon (.64)	dresser (.34)	footstool (.37)
bookcase (5.81/6.38)	cupboard (.56)	bed (.64)	couch (.34)	couch (.35)
sideboard (5.58/6.70)	fireplace (.56)	fridge (.61)	stool (.31)	pouffe (.34)

dresser (5.51/5.49)	fridge (.52)	cupboard (.60)	futon (.31)	cupboard (.32)
bench (5.13/4.93)	heater (.52)	pouffe (.59)	bed (.31)	bath (.32)
footstool (4.68/5.17)	armchair (.51)	cot (.59)	heater (.30)	divan (.32)
bureau (4.51/5.76)	cushion (.51)	cooker (.56)	fridge (.28)	typewriter (.31)
lamp (4.46/5.74)	lamp (.51)	stool (.55)	desk (.28)	cabinet (.29)
tallboy (4.46/6.04)	wardrobe (.50)	toilet (.55)	cabinet (.28)	lamp (.29)
futon (4.36/3.91)	stool (.50)	desk (.55)	cupboard (.27)	bed (.28)
cot (4.15/5.77)	toilet (.50)	lamp (.54)	lamp (.27)	chair (.28)
fridge (3.98/5.78)	divan (.49)	bath (.53)	typewriter (.27)	cot (.28)
pouffe (3.98/4.83)	carpet (.48)	heater (.52)	divan (.26)	stool (.27)
ottoman (3.97/4.53)	bath (.47)	table (.52)	cooker (.24)	heater (.27)
cushion (3.91/5.21)	cooker (.47)	armchair (.51)	cot (.23)	toilet (.27)
television (3.87/6.11)	futon (.46)	tallboy (.51)	bench (.23)	desk (.26)
bath (3.79/5.59)	sink (.45)	carpet (.50)	armchair (.22)	fridge (.25)
divan (3.78/5.87)	bench (.44)	typewriter (.50)	toilet (.22)	cooker (.25)
cooker (3.74/5.83)	cot (.43)	chair (.47)	ottoman (.20)	armchair (.22)
sink (3.67/5.32)	cabinet (.41)	wardrobe (.44)	picture (.20)	tallboy (.21)
carpet (3.55/6.35)	typewriter (.41)	cabinet (.41)	sink (.18)	television (.20)
fireplace (3.44/5.44)	picture (.40)	bench (.37)	cushion (.18)	table (.20)

picture (3.42/5.26)	curtain (.40)	curtain (.34)	tallboy (.16)	fire (.18)
curtain (3.40/5.53)	television (.37)	sink (.31)	television (.16)	cushion (.15)
toilet (3.30/5.38)	fire (.35)	cushion (.31)	fire (.15)	picture (.14)
heater (3.25/4.91)	tallboy (.34)	television (.27)	bath (.14)	sink (.14)
washstand (2.79/4.44)	ottoman (.31)	picture (.25)	curtain (.13)	bench (.13)
radio (2.63/5.40)	radio (.27)	fire (.24)	bureau (.09)	bureau (.11)
fire (2.56/4.58)	bureau (.24)	radio (.20)	plant (.08)	curtain (.10)
typewriter (2.21/4.11)	plant (.19)	bureau (.17)	radio (.06)	radio (.08)
plant (2.12/2.78)	pouffe (-)	plant (.13)	pouffe (-)	plant (.06)

---



Table A22

*Exemplars of the category insects sorted by typicality, according to the younger adults, or similarity to the prototype with the actual values in parentheses (older adults' ratings after the forward slash) for English*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
fly (6.77/6.67)	greenfly (.65)	wasp (.80)	mosquito (.54)	beetle (.77)
ant (6.55/6.11)	cockroach (.65)	moth (.77)	cockroach (.52)	aphid (.74)
bee (6.55/6.56)	weevil (.62)	beetle (.75)	beetle (.49)	wasp (.72)
wasp (6.49/6.37)	spider (.62)	spider (.72)	aphid (.49)	moth (.72)
beetle (6.32/6.15)	wasp (.61)	ant (.71)	greenfly (.49)	ant (.67)
spider (6.30/6.35)	woodlouse (.61)	caterpillar (.71)	wasp (.48)	caterpillar (.67)
ladybird (5.96/5.74)	aphid (.61)	aphid (.70)	dragonfly (.48)	weevil (.66)
bluebottle (5.88/6.11)	grasshopper (.61)	weevil (.69)	butterfly (.46)	mosquito (.64)
midge (5.75/5.83)	beetle (.60)	grasshopper (.69)	mite (.46)	locust (.61)
mosquito (5.75/5.30)	tarantula (.60)	dragonfly (.69)	spider (.45)	termite (.61)
grasshopper (5.74/5.39)	centipede (.60)	earwig (.68)	weevil (.45)	dragonfly (.60)
earwig (5.68/5.67)	scorpion (.60)	cockroach (.68)	ant (.44)	spider (.58)
butterfly (5.62/6.28)	ant (.59)	snail (.68)	tarantula (.43)	greenfly (.58)
centipede (5.57/4.92)	silverfish (.58)	ladybird (.67)	woodlouse (.43)	grasshopper (.57)

flea (5.56/5.21)	louse (.57)	greenfly (.67)	silverfish (.43)	midge (.56)
woodlouse (5.51/4.62)	bedbug (.57)	mosquito (.67)	bee (.42)	mite (.54)
caterpillar (5.51/5.37)	earwig (.57)	midge (.66)	moth (.42)	ladybird (.54)
moth (5.49/5.81)	worm (.57)	locust (.65)	bedbug (.41)	bee (.54)
cockroach (5.47/5.35)	butterfly (.56)	cicada (.65)	scorpion (.40)	earwig (.54)
slater (5.42/4.95)	mite (.56)	mantis (.64)	midge (.38)	cicada (.54)
greenfly (5.33/5.50)	dragonfly (.56)	horsefly (.64)	centipede (.38)	cockroach (.53)
dragonfly (5.19/5.46)	millipede (.55)	scorpion (.64)	grasshopper (.38)	snail (.52)
cricket (4.89/4.59)	moth (.55)	millipede (.63)	louse (.37)	hornet (.52)
millipede (4.85/4.64)	ladybird (.54)	bluebottle (.63)	worm (.37)	millipede (.50)
locust (4.81/4.96)	hornet (.54)	hornet (.62)	millipede (.36)	scorpion (.50)
worm (4.72/5.22)	bee (.52)	tarantula (.62)	earwig (.35)	bedbug (.49)
aphid (4.70/5.30)	mantis (.52)	bee (.61)	flea (.35)	silverfish (.49)
louse (4.68/4.85)	mosquito (.51)	worm (.60)	mantis (.34)	tarantula (.48)
termite (4.61/4.25)	caterpillar (.51)	mite (.60)	termite (.34)	mantis (.47)
tick (4.61/4.40)	gnat (.49)	termite (.59)	locust (.34)	worm (.47)
tarantula (4.55/4.67)	snail (.48)	centipede (.59)	hornet (.34)	centipede (.47)
hornet (4.54/4.63)	locust (.47)	silverfish (.58)	caterpillar (.34)	louse (.45)
snail (4.53/4.44)	flea (.47)	louse (.57)	ladybird (.33)	bluebottle (.44)

horsefly (4.50/4.84)	termite (.46)	maggot (.53)	gnat (.33)	horsefly (.44)
cleg (4.46/5.02)	cicada (.46)	butterfly (.52)	cicada (.33)	butterfly (.43)
slug (4.45/4.91)	fly (.46)	bedbug (.52)	snail (.31)	firefly (.42)
mantis (4.37/4.02)	slug (.45)	gnat (.52)	fly (.29)	gnat (.39)
maggot (4.36/4.80)	midge (.45)	firefly (.50)	slug (.27)	maggot (.36)
gnat (4.35/5.13)	maggot (.42)	leech (.45)	leech (.25)	tick (.29)
mite (4.33/4.38)	tick (.38)	slug (.39)	tick (.24)	leech (.26)
firefly (4.19/4.40)	bluebottle (.38)	tick (.38)	cricket (.22)	flea (.25)
bedbug (4.02/4.06)	leech (.36)	flea (.37)	maggot (.21)	slug (.22)
scorpion (3.98/4.33)	cricket (.35)	slater (.34)	bluebottle (.21)	fly (.20)
silverfish (3.62/3.61)	firefly (.31)	gadfly (.31)	firefly (.18)	gadfly (.16)
gadfly (3.50/4.40)	slater (.24)	fly (.27)	slater (.08)	cricket (.14)
weevil (3.50/4.25)	gadfly (.22)	cricket (.23)	gadfly (.08)	slater (.09)
leech (3.35/3.54)	cleg (-)	cleg (-)	cleg (-)	cleg (-)
cicada (3.33/4.33)	horsefly (-)	woodlouse (-)	horsefly (-)	woodlouse (-)

---

Table A23

*Exemplars of the category musical instruments sorted by typicality, according to the younger adults, or similarity to the prototype with the actual values in parentheses (older adults' ratings after the forward slash) for English*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
piano (6.79/6.65)	piano (.75)	flute (.86)	-	-
guitar (6.60/6.06)	vibraphone (.73)	trombone (.85)	-	-
violin (6.55/6.57)	flute (.73)	piano (.83)	-	-
trumpet (6.45/6.52)	violin (.73)	saxophone (.83)	-	-
drum (6.25/6.24)	clarinet (.72)	violin (.83)	-	-
flute (6.21/5.87)	zither (.72)	cello (.83)	-	-
trombone (6.09/6.15)	timpani (.72)	clarinet (.83)	-	-
keyboard (6.08/5.79)	saxophone (.71)	tuba (.80)	-	-
saxophone (6.06/6.28)	xylophone (.71)	harmonica (.80)	-	-
fiddle (6.02/6.50)	glockenspiel (.71)	bassoon (.79)	-	-
clarinet (5.94/6.09)	bassoon (.70)	oboe (.79)	-	-
cello (5.91/6.00)	marimba (.70)	mandolin (.78)	-	-
recorder (5.73/5.52)	dulcimer (.70)	viola (.78)	-	-
organ (5.66/5.98)	guitar (.70)	guitar (.78)	-	-

oboe (5.29/5.63)	trombone (.69)	banjo (.78)	-	-
tuba (5.23/5.08)	harp (.69)	xylophone (.78)	-	-
viola (5.18/5.50)	harmonium (.69)	zither (.76)	-	-
xylophone (5.13/5.00)	trumpet (.68)	vibraphone (.75)	-	-
cymbal (5.09/5.06)	lute (.68)	trumpet (.75)	-	-
harp (5.06/5.52)	harmonica (.67)	harpsichord (.75)	-	-
tambourine (5.04/4.80)	tambourine (.67)	lute (.75)	-	-
harmonica (4.98/5.00)	bagpipes (.66)	cornet (.74)	-	-
horn (4.96/5.52)	harpsichord (.66)	marimba (.74)	-	-
bagpipes (4.92/5.31)	oboe (.66)	euphonium (.74)	-	-
accordion (4.88/5.60)	ukulele (.65)	harp (.74)	-	-
banjo (4.85/5.33)	accordion (.65)	sousaphone (.74)	-	-
cornet (4.52/5.31)	euphonium (.65)	harmonium (.73)	-	-
synthesiser (4.48/3.93)	flugelhorn (.65)	dulcimer (.72)	-	-
bassoon (4.45/5.23)	fiddle (.64)	timpani (.72)	-	-
bugle (4.38/5.21)	celesta (.64)	flugelhorn (.72)	-	-
timpani (4.38/4.92)	cello (.64)	tambourine (.71)	-	-
kettledrum (4.35/5.33)	mandolin (.63)	ukulele (.71)	-	-
bongo (4.33/3.83)	castanets (.63)	piccolo (.70)	-	-

mandolin (4.33/4.54)	kazoo (.63)	sitar (.70)	-	-
bell (4.32/4.22)	banjo (.63)	glockenspiel (.69)	-	-
glockenspiel (4.31/4.16)	tuba (.62)	accordion (.68)	-	-
piccolo (4.31/5.12)	lyre (.61)	celesta (.68)	-	-
triangle (4.28/4.70)	sitar (.61)	clavichord (.68)	-	-
harpsichord (4.21/4.75)	clavichord (.61)	didgeridoo (.66)	-	-
panpipes (4.20/3.93)	cornet (.60)	horn (.66)	-	-
maraca (3.96/3.98)	spinet (.60)	bagpipes (.65)	-	-
harmonium (3.91/4.83)	panpipes (.60)	panpipes (.65)	-	-
cembalo (3.86/3.00)	didgeridoo (.58)	ocarina (.63)	-	-
chime (3.77/3.96)	concertina (.56)	spinet (.63)	-	-
didgeridoo (3.75/3.77)	drum (.55)	kettledrum (.62)	-	-
concertina (3.73/4.70)	viola (.52)	lyre (.62)	-	-
ukulele (3.73/5.02)	cymbal (.52)	fiddle (.61)	-	-
vibraphone (3.70/3.96)	horn (.51)	kazoo (.61)	-	-
fife (3.69/4.20)	ocarina (.51)	castanets (.61)	-	-
lute (3.67/4.31)	piccolo (.50)	cymbal (.61)	-	-
castanets (3.65/4.39)	organ (.49)	bugle (.59)	-	-
euphonium (3.63/4.67)	keyboard (.48)	maraca (.58)	-	-

whistle (3.62/4.56)	bugle (.47)	bongo (.56)	-	-
flugelhorn (3.61/3.92)	synthesiser (.46)	keyboard (.55)	-	-
celesta (3.54/3.61)	whistle (.44)	concertina (.55)	-	-
clavichord (3.48/3.77)	bongo (.38)	drum (.52)	-	-
lyre (3.38/4.06)	recorder (.36)	organ (.50)	-	-
spinet (3.33/3.94)	bell (.36)	fife (.47)	-	-
ocarina (3.27/4.00)	chime (.32)	whistle (.38)	-	-
sitar (3.13/3.87)	fife (.17)	chime (.35)	-	-
zither (3.11/4.16)	triangle (.14)	bell (.33)	-	-
kazoo (3.09/3.28)	cembalo (-)	recorder (.33)	-	-
marimba (3.07/3.29)	kettledrum (-)	triangle (.22)	-	-
dulcimer (3.00/4.33)	maraca (-)	cembalo (-)	-	-
sousaphone (2.73/3.92)	sousaphone (-)	synthesiser (-)	-	-

---

Table A24

*Exemplars of the category tools sorted by typicality, according to the younger adults, or similarity to the prototype with the actual values in parentheses (older adults' ratings after the forward slash) for English*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
screwdriver (6.75/6.67)	hacksaw (.68)	screwdriver (.73)	chisel (.45)	tweezers (.30)
hammer (6.68/6.79)	chisel (.67)	awl (.71)	scraper (.38)	scraper (.30)
saw (6.60/6.59)	crowbar (.66)	pliers (.67)	screwdriver (.36)	screwdriver (.29)
spanner (6.50/6.50)	pickaxe (.64)	chisel (.67)	crowbar (.35)	wrench (.28)
spade (6.42/6.56)	penknife (.64)	mallet (.66)	trowel (.35)	lathe (.27)
drill (6.36/6.06)	screwdriver (.64)	secateurs (.66)	wrench (.34)	hacksaw (.27)
pliers (6.08/6.35)	secateurs (.64)	hacksaw (.66)	sandpaper (.33)	awl (.27)
shovel (5.94/6.30)	hammer (.63)	scissors (.65)	scissors (.32)	gauge (.26)
axe (5.83/6.12)	sandpaper (.60)	sander (.64)	axe (.32)	pliers (.25)
chisel (5.74/6.32)	scraper (.60)	trowel (.63)	pickaxe (.32)	sander (.25)
rake (5.68/6.13)	trowel (.59)	shears (.63)	hacksaw (.32)	penknife (.24)
file (5.51/6.04)	pincers (.59)	crowbar (.62)	secateurs (.32)	scissors (.23)
wrench (5.51/5.20)	scissors (.59)	tweezers (.62)	hammer (.31)	jigsaw (.23)
trowel (5.49/6.17)	wrench (.58)	clippers (.62)	grinder (.31)	ratchet (.21)



hacksaw (5.48/5.65)	tweezers (.57)	trimmer (.61)	penknife (.30)	chisel (.21)
shears (5.44/5.82)	trimmer (.56)	pickaxe (.61)	shovel (.30)	trowel (.21)
mallet (5.34/5.46)	shovel (.56)	shovel (.61)	paintbrush (.29)	crowbar (.21)
paintbrush (5.23/5.70)	axe (.56)	penknife (.61)	shears (.29)	trimmer (.21)
hoe (5.18/6.07)	bolt (.55)	wrench (.60)	trimmer (.27)	paintbrush (.21)
pickaxe (5.17/5.83)	spanner (.53)	scraper (.60)	lathe (.27)	clamp (.20)
plane (5.12/6.00)	rasp (.53)	lathe (.59)	pencil (.27)	grinder (.20)
sander (5.06/5.25)	grinder (.52)	grinder (.58)	tweezers (.26)	secateurs (.19)
clippers (5.04/5.34)	shears (.52)	paintbrush (.57)	spanner (.26)	scythe (.19)
pick (5.02/5.67)	pliers (.51)	pencil (.55)	clippers (.26)	pencil (.19)
vice (4.98/5.72)	drill (.51)	screw (.55)	rasp (.26)	pickaxe (.19)
lathe (4.97/5.22)	scythe (.50)	pincers (.54)	pincers (.26)	sickle (.18)
jack (4.96/5.44)	lathe (.50)	hammer (.54)	stepladder (.26)	screw (.18)
crowbar (4.89/5.09)	pencil (.50)	sandpaper (.54)	pliers (.25)	pincers (.18)
clamp (4.88/5.32)	clippers (.49)	hoe (.53)	nail (.24)	anvil (.17)
scraper (4.82/5.21)	screw (.48)	spade (.53)	bolt (.24)	clippers (.17)
scissors (4.81/6.28)	spade (.48)	scythe (.51)	hose (.23)	shears (.17)
hatchet (4.80/5.79)	micrometer (.48)	hose (.50)	clamp (.22)	hoe (.17)
ruler (4.74/5.96)	stepladder (.47)	bolt (.49)	ratchet (.22)	punch (.16)

trimmer (4.71/5.35)	hatchet (.47)	stepladder (.48)	hatchet (.22)	dividers (.16)
ratchet (4.65/4.55)	nail (.46)	hatchet (.47)	mallet (.20)	nail (.15)
penknife (4.62/5.68)	ratchet (.45)	roller (.47)	scythe (.20)	roller (.15)
jigsaw (4.57/4.72)	hoe (.45)	spanner (.46)	hoe (.19)	callipers (.15)
blowlamp (4.53/5.07)	clamp (.45)	sickle (.45)	rake (.19)	file (.15)
rasp (4.52/5.00)	mallet (.44)	nail (.44)	drill (.19)	mallet (.15)
secateurs (4.46/5.66)	punch (.44)	jack (.43)	micrometer (.19)	hose (.15)
grinder (4.37/4.82)	rake (.44)	rasp (.42)	spade (.19)	nut (.15)
pincers (4.33/5.40)	callipers (.42)	snips (.42)	torch (.19)	shovel (.15)
nail (4.30/5.71)	hose (.42)	callipers (.42)	screw (.18)	hatchet (.14)
callipers (4.28/4.62)	nut (.40)	anvil (.41)	snips (.18)	ruler (.14)
awl (4.25/5.04)	snips (.39)	rake (.41)	anvil (.18)	spanner (.14)
anvil (4.22/4.70)	pick (.39)	nut (.38)	nut (.17)	micrometer (.14)
snips (4.15/4.89)	paintbrush (.39)	jigsaw (.37)	awl (.17)	sandpaper (.14)
scythe (4.13/4.52)	torch (.38)	punch (.35)	sander (.17)	jack (.14)
gauge (4.12/4.74)	anvil (.38)	micrometer (.35)	ruler (.16)	snips (.12)
roller (4.04/4.20)	jack (.37)	dividers (.33)	punch (.16)	brace (.12)
pencil (4.00/5.37)	roller (.36)	clamp (.33)	roller (.16)	hammer (.12)
screw (4.00/5.76)	brace (.35)	torch (.33)	jack (.14)	stepladder (.12)

sickle (3.98/4.85)	dividers (.34)	drill (.31)	dividers (.14)	drill (.11)
punch (3.93/4.71)	awl (.34)	brace (.31)	sickle (.14)	rake (.10)
torch (3.91/4.94)	saw (.34)	ratchet (.29)	jigsaw (.14)	spade (.10)
stepladder (3.89/5.61)	gauge (.33)	ruler (.26)	pick (.12)	rasp (.10)
micrometer (3.78/4.24)	sickle (.31)	pick (.25)	gauge (.12)	bolt (.09)
dividers (3.74/4.39)	sander (.28)	gauge (.25)	brace (.11)	plane (.03)
tweezers (3.70/4.94)	plane (.27)	plane (.20)	callipers (.10)	torch (.02)
sandpaper (3.66/4.96)	ruler (.26)	saw (.17)	saw (.09)	vice (.02)
brace (3.54/5.29)	jigsaw (.25)	file (.12)	file (.08)	pick (-.00)
bolt (3.41/4.64)	file (.20)	vice (.09)	plane (.02)	saw (-.01)
hose (3.40/4.55)	vice (.15)	axe (-)	vice (.00)	axe (-)
nut (3.08/5.02)	blowlamp (-)	blowlamp (-)	blowlamp (-)	blowlamp (-)

---

Table A25

*Exemplars of the category vegetables sorted by typicality, according to the younger adults, or similarity to the prototype with the actual values in parentheses (older adults' ratings after the forward slash) for English*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
carrot (6.75/6.87)	courgette (.76)	cauliflower (.83)	cabbage (.44)	potato (.63)
potato (6.75/6.96)	spinach (.71)	cabbage (.83)	broccoli (.42)	broccoli (.61)
pea (6.40/6.41)	celery (.69)	asparagus (.83)	asparagus (.40)	onion (.59)
cabbage (6.21/6.57)	cabbage (.68)	celery (.80)	courgette (.40)	cauliflower (.59)
onion (6.13/6.68)	asparagus (.68)	broccoli (.79)	spinach (.39)	cabbage (.57)
cauliflower (6.09/6.46)	cauliflower (.67)	radish (.79)	beetroot (.39)	asparagus (.56)
broccoli (6.00/5.87)	radish (.67)	onion (.78)	lettuce (.37)	garlic (.55)
cucumber (5.89/5.72)	aubergine (.67)	courgette (.77)	potato (.34)	cucumber (.55)
leek (5.87/6.41)	onion (.67)	cucumber (.77)	artichoke (.33)	celery (.54)
turnip (5.83/6.63)	beetroot (.66)	garlic (.77)	cauliflower (.32)	lettuce (.54)
lettuce (5.74/6.47)	garlic (.64)	artichoke (.77)	celery (.32)	spinach (.53)
sprout (5.68/6.44)	broccoli (.64)	potato (.76)	turnip (.32)	yam (.53)
mushroom (5.42/5.72)	cucumber (.63)	lettuce (.76)	onion (.31)	courgette (.53)
bean (5.34/6.09)	potato (.63)	parsnip (.76)	sprout (.30)	radish (.51)

pepper (5.30/5.13)	lettuce (.62)	leek (.75)	radish (.30)	artichoke (.49)
courgette (5.25/5.13)	parsnip (.62)	spinach (.72)	pea (.29)	bean (.48)
corn (5.23/5.06)	turnip (.62)	beetroot (.68)	aubergine (.29)	corn (.47)
parsnip (5.19/5.73)	artichoke (.58)	bean (.67)	garlic (.27)	mushroom (.46)
celery (5.02/5.32)	pea (.57)	pea (.66)	carrot (.27)	beetroot (.46)
swede (4.91/6.34)	carrot (.55)	turnip (.65)	parsnip (.27)	leek (.46)
spinach (4.67/5.15)	mushroom (.54)	mushroom (.64)	cucumber (.26)	parsnip (.44)
beetroot (4.60/5.96)	pepper (.53)	swede (.64)	corn (.23)	turnip (.44)
asparagus (4.49/5.07)	sprout (.52)	aubergine (.62)	bean (.22)	aubergine (.41)
aubergine (4.40/4.83)	bean (.50)	pepper (.60)	yam (.22)	gourd (.41)
radish (4.32/4.87)	corn (.47)	yam (.60)	pepper (.22)	swede (.40)
marrow (4.27/4.92)	swede (.45)	gherkin (.56)	gherkin (.21)	pea (.39)
artichoke (3.77/4.36)	yam (.43)	corn (.56)	gourd (.20)	gherkin (.38)
garlic (3.74/4.72)	gourd (.43)	gourd (.55)	mushroom (.19)	pepper (.37)
gherkin (3.58/4.40)	gherkin (.41)	carrot (.53)	leek (.18)	carrot (.31)
yam (3.43/4.25)	leek (.41)	sprout (.44)	swede (.17)	sprout (.27)
gourd (1.56/2.62)	marrow (.36)	marrow (.40)	marrow (.14)	marrow (.24)

---

Table A26

*Exemplars of the category vehicles sorted by typicality, according to the younger adults, or similarity to the prototype with the actual values in parentheses (older adults' ratings after the forward slash) for English*

Typicality	Prototype-as-average		Prototype-as-category-label	
	CBOW	skip-gram	CBOW	skip-gram
car (6.98/6.98)	boat (.65)	boat (.74)	car (.65)	car (.78)
bus (6.85/6.94)	truck (.65)	motorboat (.72)	truck (.59)	truck (.70)
train (6.73/6.89)	car (.65)	scooter (.69)	lorry (.51)	van (.65)
aeroplane (6.47/6.73)	motorboat (.64)	speedboat (.69)	jeep (.51)	tractor (.59)
coach (6.32/6.43)	speedboat (.58)	canoe (.65)	trailer (.44)	scooter (.58)
taxi (6.32/6.57)	jeep (.57)	truck (.65)	minibus (.43)	jeep (.57)
van (6.23/6.67)	hydroplane (.55)	van (.65)	tractor (.41)	moped (.53)
underground (6.19/5.74)	scooter (.55)	motorbike (.65)	tanker (.38)	motorbike (.51)
bicycle (6.17/6.41)	helicopter (.54)	hovercraft (.65)	ambulance (.38)	minibus (.49)
motorbike (6.04/6.00)	bus (.54)	tractor (.64)	bus (.38)	bicycle (.49)
boat (5.94/6.46)	hydrofoil (.54)	dinghy (.64)	horsebox (.38)	motorboat (.47)
ship (5.75/6.60)	aeroplane (.53)	kayak (.64)	wagon (.36)	trailer (.47)
minibus (5.74/5.87)	motorbike (.53)	car (.63)	van (.36)	bus (.46)
truck (5.74/6.22)	wagon (.53)	jeep (.63)	tank (.35)	lorry (.45)

ferry (5.66/5.58)	sidecar (.53)	bicycle (.62)	motorbike (.35)	taxi (.44)
lorry (5.60/6.74)	tractor (.52)	lorry (.62)	helicopter (.35)	boat (.43)
jet (5.26/6.00)	minibus (.51)	moped (.61)	speedboat (.34)	horsebox (.42)
jeep (5.17/5.32)	canoe (.51)	bus (.61)	scooter (.32)	helicopter (.40)
moped (4.83/4.42)	ship (.51)	hydrofoil (.61)	taxi (.32)	speedboat (.40)
helicopter (4.81/5.83)	bicycle (.51)	trolley (.61)	caravan (.32)	wagon (.40)
scooter (4.62/4.68)	trailer (.51)	ferry (.61)	rickshaw (.31)	rickshaw (.39)
ambulance (4.53/5.74)	yacht (.50)	rickshaw (.60)	spacecraft (.30)	tricycle (.37)
liner (4.33/6.02)	train (.50)	glider (.60)	juggernaut (.30)	caravan (.37)
motorboat (4.21/5.15)	lorry (.50)	minibus (.60)	aeroplane (.30)	ambulance (.36)
tram (4.21/5.06)	buggy (.50)	helicopter (.59)	moped (.30)	tanker (.35)
yacht (4.15/4.91)	biplane (.49)	tricycle (.58)	motorboat (.29)	buggy (.35)
hovercraft (4.11/4.81)	airship (.49)	wagon (.58)	ship (.29)	glider (.35)
caravan (4.09/5.52)	hovercraft (.49)	sleigh (.57)	boat (.28)	cart (.34)
pantechnicon (4.00/4.81)	dinghy (.49)	taxi (.56)	bicycle (.28)	hydroplane (.33)
speedboat (3.92/4.58)	moped (.49)	barge (.56)	sidecar (.28)	kayak (.33)
tractor (3.91/5.11)	unicycle (.49)	horsebox (.56)	carriage (.27)	sleigh (.32)
carriage (3.85/4.35)	rickshaw (.48)	trawler (.56)	rocket (.26)	canoe (.31)
juggernaut (3.80/5.25)	horsebox (.48)	yacht (.55)	hydroplane (.26)	airship (.30)

wagon (3.70/4.93)	carriage (.48)	buggy (.55)	hovercraft (.26)	trolley (.30)
hydrofoil (3.64/4.76)	glider (.48)	microlight (.55)	hydrofoil (.26)	jet (.29)
tanker (3.63/5.20)	barge (.47)	cart (.54)	train (.25)	train (.29)
cart (3.57/4.57)	ferry (.47)	schooner (.54)	buggy (.25)	spacecraft (.29)
barge (3.53/4.63)	trolley (.47)	airship (.53)	tricycle (.25)	hovercraft (.29)
buggy (3.51/4.26)	taxi (.46)	train (.52)	tram (.25)	microlight (.29)
monorail (3.48/3.82)	skis (.46)	pram (.52)	trolley (.24)	wheelchair (.29)
hydroplane (3.48/4.13)	tricycle (.46)	surfboard (.52)	monorail (.24)	trawler (.28)
pram (3.38/5.74)	rocket (.46)	biplane (.52)	trawler (.22)	carriage (.28)
trailer (3.34/5.11)	sledge (.45)	tram (.51)	submarine (.21)	skateboard (.28)
raft (3.31/3.53)	surfboard (.45)	tanker (.51)	barge (.21)	barge (.28)
canoe (3.30/4.43)	tanker (.45)	skateboard (.51)	cart (.21)	dinghy (.27)
dinghy (3.30/4.70)	sleigh (.45)	ship (.51)	pram (.20)	yacht (.27)
sledge (3.28/4.09)	kayak (.44)	caravan (.50)	sledge (.20)	ferry (.26)
biplane (3.26/3.85)	monorail (.44)	wheelchair (.50)	skateboard (.20)	tank (.26)
trawler (3.25/4.63)	ambulance (.44)	unicycle (.50)	microlight (.19)	surfboard (.26)
airship (3.25/4.04)	tank (.44)	trailer (.49)	balloon (.19)	rocket (.25)
tricycle (3.25/4.35)	cart (.43)	submarine (.49)	dinghy (.19)	submarine (.24)
submarine (3.21/4.33)	balloon (.42)	carriage (.49)	yacht (.19)	sidecar (.24)



tandem (3.19/4.41)	pram (.42)	hydroplane (.48)	biplane (.19)	ship (.24)
schooner (3.17/4.02)	tram (.42)	sidecar (.48)	ferry (.19)	hydrofoil (.23)
horsebox (3.13/4.45)	raft (.41)	skis (.48)	wheelchair (.18)	pram (.23)
sleigh (3.13/4.24)	submarine (.41)	jet (.48)	kayak (.18)	biplane (.23)
kayak (3.08/3.63)	microlight (.41)	sledge (.47)	airship (.18)	schooner (.22)
balloon (3.04/3.76)	schooner (.40)	ambulance (.46)	glider (.17)	balloon (.20)
wheelchair (3.00/4.53)	wheelchair (.40)	balloon (.43)	canoe (.16)	skis (.20)
rickshaw (2.97/4.04)	jet (.40)	spacecraft (.41)	trap (.15)	tram (.19)
skis (2.92/4.34)	skateboard (.39)	raft (.40)	liner (.15)	sledge (.19)
sidecar (2.91/3.54)	juggernaut (.38)	monorail (.39)	sleigh (.15)	unicycle (.19)
tank (2.89/3.91)	spacecraft (.38)	rocket (.36)	schooner (.15)	monorail (.15)
skateboard (2.85/3.48)	caravan (.38)	tank (.35)	jet (.13)	raft (.15)
rocket (2.83/3.02)	trawler (.37)	liner (.27)	unicycle (.13)	trap (.14)
glider (2.82/4.27)	tandem (.36)	trap (.23)	skis (.13)	liner (.14)
spacecraft (2.81/3.72)	van (.36)	juggernaut (.22)	coach (.12)	juggernaut (.12)
microlight (2.73/3.43)	trap (.35)	underground (.18)	surfboard (.12)	underground (.11)
surfboard (2.55/3.13)	coach (.33)	tandem (.17)	tandem (.09)	tandem (.05)
trolley (2.55/4.64)	liner (.33)	coach (.13)	underground (.08)	coach (.02)
unicycle (2.49/2.84)	underground (.21)	aeroplane (-)	raft (.08)	aeroplane (-)

trap (2.31/3.48)	pantechnicon (-)	pantechnicon (-)	pantechnicon (-)	pantechnicon (-)
------------------	------------------	------------------	------------------	------------------

---