Sensorimotor Norms: Perception and Action Strength norms for 40,000 words

Dermot Lynott (d.lynott@lancaster.ac.uk)
Department of Psychology, Lancaster University, Lancaster, UK
Louise Connell (l.connell@lancaster.ac.uk)
Department of Psychology, Lancaster University, Lancaster, UK
Marc Brysbaert (Marc.Brysbaert@UGent.be)
Department of Experimental Psychology, Ghent University, Belgium.
James Brand (james.brand@canterbury.ac.nz)
New Zealand Institute of Language Brain and Behaviour, University of Canterbury, New Zealand
James Carney (James.Carney@brunel.ac.uk)
Department of Arts and Humanities, Brunel University, London, UK.

Abstract

Sensorimotor information plays a fundamental role in cognition. However, datasets of ratings of sensorimotor experience have generally been restricted to several hundred words, leading to limited linguistic coverage and reduced statistical power for more complex analyses. Here, we present modality-specific and effector-specific norms for 39,954 concepts across six sensory modalities (touch, hearing, smell, taste, vision, and interoception) and five action effectors (mouth/throat, hand/arm, foot/leg, head excluding mouth, and torso), which were gathered from 4,557 participants who completed a total of 32,456 surveys using Amazon's Mechanical Turk platform. The dataset therefore represents one of the largest set of semantic norms currently available. We describe the data collection procedures, provide summary descriptives of the data set, demonstrate the utility of the norms in predicting lexical decision times and accuracy, as well as offering new insights and outlining avenues for future research. Our findings will be of interest to researchers in embodied cognition, cognitive semantics, sensorimotor processing, and the psychology of language generally. The scale of this dataset will also facilitate computational modelling and big data approaches to the analysis of language and conceptual representations.

Keywords: embodied cognition; semantics; norms

Background

Sensorimotor information is central to how we experience and navigate the world. We acquire information through our senses, while our bodies provide feedback, as we physically interact with objects, people, and the wider environment. Many theoretical views of cognition describe a fundamental role for such sensorimotor knowledge (e.g., Barsalou, 1999; Connell & Lynott, 2014; Smith & Gasser, 2005), with numerous empirical demonstrations supporting such claims (e.g., Connell, Lynott & Dreyer, 2012; Kaschak et al., 2006; Matlock, 2004; Zwaan & Taylor, 2006).

In order to test such embodied (or grounded) theories of cognition, researchers need appropriate stimuli for empirical tests and for developing mathematical or computational models. Lynott and Connell (2009, 2013) developed a set of modality-specific sensory norms for concepts where each sensory modality (e.g., auditory, gustatory, haptic, olfactory, visual) maps onto distinct cortical regions (e.g., gustatory cortex, auditory cortex etc.). By having individuals provide ratings for each modality separately, the norms capture the extent to which something is experienced across different sensory modalities, without risk of ignoring or distorting the role of particular modalities (Connell & Lynott, 2016). Subsequent empirical studies have found that such modality-specific measures are good predictors of people's performance across a range of cognitive tasks (e.g., lexical decision, word-naming) and often out-performed long-established measures such as concreteness and imageability (e.g., Connell & Lynott, 2012; 2014). For example, in examining performance on lexical decision and word naming (reading aloud) tasks, Connell and Lynott (2012) found that modality-specific experience (and specifically the highest level of perceptual experience on any modality for a given concept, or “max strength”) was a more reliable predictor of performance than either concreteness and imageability.

An added advantage of using measures of sensory experience for specific modalities is that it allows researchers to tap into effects that relate to particular modalities and not others. Connell & Lynott (2010) showed how a processing disadvantage for tactile stimuli observed during perceptual processing (Spence, Nichols & Driver, 2001) was also observed when processing modality-specific words. Connell and Lynott (2014) derived contrasting modality-specific predictions relating to lexical decision and reading aloud for individual words. Thus, for lexical decisions, a visually-focussed task, strength of perceptual experience in the visual modality (but not the auditory modality) was a reliable predictor of performance. By contrast, reading aloud, requires additional attention on the auditory modality (as participants must monitor their speech output to ensure correctly articulated responses). Consistent with this idea, both strength of auditory experience and strength of visual experience were reliable predictors for performance for the reading aloud task. Other semantic measures (such as concreteness or imageability) could not have been used as they do not offer sufficient granularity in terms of sensory experience. Thus, modality-specific measures of sensory experience provide the capacity to generate and test novel predictions related to modality-specific processing and representations.

More recently, Connell, Lynott and Banks (2018) showed that interoception (i.e., sensations inside the body) also plays an important role in semantic representations, and could be a primary grounding mechanism for abstract concepts. It was found that strength of interoceptive experience was higher for abstract concepts, such as *hungry* and *serenity*, compared to more concrete concepts like *capacity* or *rainy*. What's more, interoceptive experience was found to be most important for emotion concepts, especially for negative emotions such as *fear* and *sadness,* with interoceptive experience found to be just as important as other sensory modalities in capturing semantic knowledge.

Finally, speaking to the utility of sensory norms and broader interest in this area, several research groups have extended this earlier work, either by developing modality-specific norms in other languages, including Russian, Serbian, Dutch and Mandarin (Miklashevsky, 2018; Đurđević et al., 2016; Speed & Majid, 2017; Chen et al., 2017), or by applying these norms in novel ways. For example, the original modality-specific norms have been used to examine stylistic differences of authors (Kernot, Bossomaier & Bradley, 2019), test models of lexical representations (Johns & Jones, 2012), and evaluate the iconicity of words (Winter et al., 2017).

Nonetheless, a notable gap in the work discussed above is that it focuses solely on sensory experience, and has not included parallel measures of action or effector-specific experience. However, there is good evidence for the relevance of action experience to people's semantic representations of concepts (e.g., Glenberg & Gallese, 2012; Hauk, Johnsrude & Pulvermuller, 2004). For instance, manual action verbs like *throw* activate some of the same motor circuits as moving the hand (Hauk et al., 2004), and their processing is selectively impaired in patients with Parkinson’s disease, which entails neurodegeneration of the motor system (Boulenger et al., 2008; Fernandino et al., 2013). Critically, the motor basis to semantic knowledge is specific to the bodily effector used to carry out a particular action. Applying transcranial magnetic stimulation (TMS) to hand and leg areas of the motor cortex differentially influences processing of hand- and leg-action words: hand area TMS facilitates lexical decision of hand-action words like *pick* compared to leg-action words like *kick*, whereas this effect is reversed with leg-area TMS (Pulvermueller, Hauk, Nikulin & Ilmoniemi, 2005). Such double dissociations in motor-language facilitation underscore the importance of individually examining separate action effectors when norming the motor basis of words and concepts.

Some existing measures have attempted to capture action knowledge, but have alternatively used feature production methods as opposed to rating dimensions of action (e.g., where people verbally list features associated with concepts: McRae et al., 2005; Vinson & Vigliocco, 2008), focused on generalised action (e.g., body-object interaction: Tillotson, Siakaluk, & Pexman, 2008; relative embodiment: Sidhu, Kwan, Pexman, & Siakaluk, 2014; see Connell & Lynott, 2015, for review), or on a restricted subset of action types (e.g., graspability: Amsel, Urbach, & Kutas, 2012: actions associated with lower limb, upper limb, or head: Binder et al., 2016) that omits other parts of the body involved in action. For example, the action of *pushing* can also involve the torso (Moody & Gennari, 2010), and mouth actions are cortically distinct from other actions of the face (Meier, Aflalo, Kastner, & Graziano, 2008). To our knowledge, therefore, there is no large-scale set of norms that taps into a comprehensive range of effector-specific action experience. In the present work, we address this gap by collecting effector-specific action strength norms for a large number of concepts.

Here, we present sensorimotor norms collected across 11 dimensions for approximately 40,000 concepts, comprising 6 modality-specific dimensions of perceptual strength (auditory, gustatory, haptic, olfactory, visual, interoceptive) and 5 effector-specific dimensions of action strength (head, arm/hand, mouth/throat, leg/foot, torso).

Study 1: Sensorimotor Norms

Method

Participants A total of 4,557 unique participants completed 32,456 surveys via Amazon’s Mechanical Turk platform (M = 7.12 samples per participant). Data for perceptual strength ratings and action strength ratings were gathered separately. Participants were self-selecting and had English as their first language. We recruited only experienced MTurk users who had already completed over 100 HITS, and high-quality participants who had >97% HIT approval. Participants were remunerated at a rate above minimum wage in the US.

Materials Perceptual and action ratings were collected for a total of 39,954 words. These words were taken from Brysbaert, Warriner, & Kuperman's (2014) work on concreteness ratings, which included 37,058 English lemmas and 2,896 two-word expressions. These words were split into 832 lists of 48 items, along with 5 calibrator words and 5 control words occurring in each. Responses to controls and calibrators (selected for being highly familiar, and low in variance based on previous norms) were used for quality checks, which we describe below in the subsection on Data Quality and Exclusions. Lists were populated to provide words that varied in terms of familiarity (“percentage known” in Brysbaert et al's study) and concreteness.

Procedure Using Qualtrics survey software, a template survey was created that followed procedures developed in Lynott & Connell (2009, 2013). At the start of the survey participants read an information sheet, and indicated their informed consent to continue with the study. Specifically, each concept in a 58-word sample was presented individually on a screen (order randomised by participant) followed by question text. For perceptual strength ratings, the text was “To what extent do you experience WORD,” where WORD was replaced with the concept in question. Underneath were six rating scales, one for each of the perceptual modalities under investigation, labelled “By feeling through touch”, “By hearing”, “By sensations inside your body”, “By smelling” and “By tasting,”. The order of the ratings scales was randomised by sample.

For the action strength ratings, the text read “To what extent do you experience WORD by performing an action with,” followed by choices of “Foot / leg”, “Hand / arm”, “Head excluding mouth”, “Mouth / throat”, “Torso”. For these ratings, each scale also contained an image of a body avatar that highlighted the body part relevant to each intended effector. Participants were asked to rate the extent to which they experience each concept through each of the named senses or effectors; both the sensory and motor components had 6-point scales ranging from 0 (not at all) to 5 (greatly).

There was no default value selected on the scale and participants clicked on a button under the relevant value to select or change their response. Participants were explicitly told there were no right or wrong answers and they should use their own judgment; they were also instructed to select the “I don’t know the meaning of this word” option if the word was unfamiliar to them. Progress to the next item could only occur if values were selected for all perceptual senses or action effectors or the “I don’t know the meaning of this word” option was checked. The study was self-paced and timed to last 18-20 minutes.

Data quality and exclusions In order to ensure the data collected is of sufficiently high quality, we instituted a number of checks, in terms of individual performance, item performance, and agreement for each list of words. Overall, only 0.8% of all responses were removed following data checks. Participants whose scores exhibited a Pearson’s $r < 0.2$ with the controls or who responded ‘don’t know the meaning of this word’ for more than five control and calibrator words, were dropped from the sample. Additionally, there were a small number of participants who completed the same sample of words more than once, when this happened only the earliest submitted responses were retained. Cronbach’s alphas (Cronbach, 1951) were calculated for each modality for all other participants; results were only retained when the mean alpha for all samples was >= 0.8.

Norms Data The final set of norms, results, analyses, and scripts are available on the project’s Open Science Framework page: https://osf.io/7emr6/

Results

Summary statistics were calculated for all valid samples, with 39,707 words included in the overall norms, following exclusion criteria. Each word in a sample is represented by a row that contains ratings for each of the 11 dimensions. Each dimension has separate values for mean score, standard deviation, median score, trimmed mean, trimmed standard deviation by modality/effector, and the percentage of participants who knew the word. Inter-rater reliability by modality/effector was high for both perceptual and action ratings: mean Cronbach’s alphas for perceptual modalities were: auditory 0.93, gustatory 0.96, haptic 0.92, interoceptive 0.92, olfactory 0.94 and visual 0.90; for action effectors mean alphas were: foot 0.93, hand 0.91, head 0.85, mouth 0.92 and torso 0.89.

Following Lynott and Connell (2009; 2013), additional variables of interest were calculated for each of the words in the sensorimotor norms. This included: Exclusivity scores (i.e., a measure of the extent to which a particular concept is experienced through a single dimension, calculated per word as the rating range divided by the sum of the ratings, and extending from 0%, for completely multidimensional, to 100%, for completely unidimensional); separate exclusivity scores were calculated for the perceptual (6 modalities) and action components (5 effectors), in addition to scores calculated across all 11 dimensions. Similarly, each concept was assigned a dominant dimension (i.e., the dimension that had the highest mean rating), for the perceptual, action and the full sensorimotor norms. When the highest mean rating was found in more than 1 dimension (Perceptual: N = 593; Action: N = 706; Sensorimotor: N = 478), a random dimension was assigned.

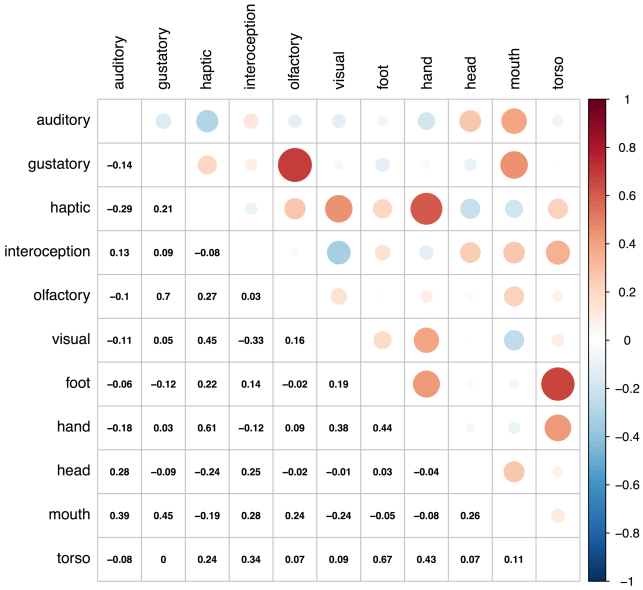

Figure 1 Correlation matrix plot between 11 dimensions for mean ratings of the sensorimotor strength norms (N = 39,707). Larger circles indicate stronger correlations, with red shades being positive, and blue shades being negative.

The norms confirm previous reports that we predominantly experience the world perceptually through our visual modality (Lynott & Connell, 2009; 2013; Winter, Perlman & Majid, 2018 – See Table 1), with the head emerging as the primary action effector. The least prominent dimensions were gustation and olfaction, highlighting the fact that only a small subset of the conceptual system is experienced strongly through these modalities. For the action norms, the head was observed to be the dominant effector and the torso had the least dominance.

Bayesian correlation analysis (Figure 1) between the dimensions showed that almost all the dimensions were significantly correlated with one another, with the exception of gustation~torso, as well as head~vision, with correlations approaching zero. It should be noted however, that a large number of the correlations were very weak, which is to be expected as each dimension is tapping into different aspects of sensorimotor experience. In some cases of course, certain dimensions often co-occur in our sensorimotor experience, with notable relationships found between gustation~olfaction (r = .70), foot/leg~torso (r = 0.67) and hand/arm~haptic (r = 0.62).

	M	*SD*
Perceptual Modality		
Auditory	1.514	0.991
Gustatory	0.324	0.697
Haptic	1.074	0.934
Interoceptive	1.032	0.880
Olfactory	0.390	0.619
Visual	2.897	0.902
Action Effector		
Foot/leg	0.807	0.750
Hand/arm	1.447	0.907
Head	2.276	0.719
Mouth/throat	1.257	0.903
Torso	0.816	0.670

Table 1 Mean Strength Ratings (0–5) and Standard deviations (SD) per Sensorimotor Dimension

Study 2: Modelling Lexical Decision

In Study 2, we address three issues. First, we determine what is the best composite variable (i.e., single value) for representing a concept's sensorimotor profile. Second, we wish to replicate the utility of perceptual strength ratings in modelling people's performance in cognitive tasks (e.g., Connell & Lynott, 2012), and establish the independent utility of action strength as a performance predictor. Third, we will check the generalisability of the findings, by examining performance across two different data sets (i.e., English Lexicon Project, British Lexicon Project).

While an 11-dimension sensorimotor profile is a rich source of semantic information about a particular concept, it can nonetheless be somewhat unwieldy for some uses. It is often useful to aggregate multiple dimensions into a single composite variable, such as for use as a predictor in regression analyses without unnecessarily inflating the number of parameters. A single variable would also facilitate comparisons with other single-variable measures of people's experience (e.g., concreteness, valence etc.) There are many different methods of creating a composite variable. Previous work on perceptual strength has used strength of the dominant modality (i.e., maximum perceptual strength rating across all modalities) as the preferred composite variable (e.g., Connell & Lynott, 2016; Connell, Lynott, & Banks, 2018), finding it offered a better fit than alternatives to visual word recognition performance (Connell & Lynott, 2012). However, work in Serbian (Đurđević, Stijačić & Karapandžić, 2016) found the best fit emerged from summed perceptual strength (i.e., sum of perceptual strength ratings across all modalities) or vector length (i.e., Euclidean distance of the multidimensional vector of perceptual strength ratings from the origin). It is difficult to be certain whether this variability is due to language differences (i.e., English vs Serbian) or sampling differences (i.e., hundreds of words with limited overlap).

We therefore sought to empirically determine the best single composite variable for the 11-dimension sensorimotor profile using a much larger and more representative sample of concepts in English. As with previous studies (e.g., Connell & Lynott, 2012), we judge the "best" variable to be the one that offers the best fit to lexical decision latency, a task where semantic facilitation emerges from automatic and implicit access to the sensorimotor basis of the concept.

Method

Materials A total of 22,297 words were collated, representing the intersection of data available between the sensorimotor strength norms and lexical decision data from the English Lexicon Project (Balota, et al., 2007). A separate set of 11,768 words was also collated from the British Lexicon Project (Keuleers, Lacey, Rastle & Brysbaert, 2012).

Candidate Composite Variables Composite variables were calculated separately for sensorimotor (all 11 dimensions), perception (6 dimensions) and action (5 dimensions) dimensions. Most of the candidate variables we tested are distance metrics in vector space of a particular concept (i.e., an 11-dimension vector) from the origin. Minkowski distance (with exponent parameter *m*) is a generalisation of these distance metrics: roughly speaking, the highest-value dimension always contributes to the calculated distance, and *m* determines the extent to which the other dimensions contribute according to how close their values are to the highest-value dimension. That is, low-value *m* means that all dimensions make noticeable contributions to the calculated distance, whereas high-value *m* means only the highest-value dimension(s) make noticeable contributions to the calculated distance.

For example, for Minkowski 10 distance (Minkowski distance at m = 10 of the vector from the origin), theoretically, it represents sensorimotor strength of the dominant dimension plus an attenuated influence of any other dimensions that are nearly as strong as the dominant dimension. By contrast Minkowski 3 distance represents sensorimotor strength in all dimensions but the influence of weaker dimensions is attenuated.

Our set of candidate variables comprises: maximum strength, Minkowski 3, Minkowski 10, Euclidean vector length, Summed strength, and single PCA component.

Design and Analysis We performed Bayesian linear regressions predicting the dependent variable of zRT (i.e., standardised Lexical Decision RT per participant) and accuracy from 2 datasets: Elexicon (ELP) and the British Lexicon Project (BLP). First, for each dependent variable we built a null model of lexical predictors (log SUBTLEX word frequency, number of letters, number of syllables, orthographic Levenstein Distance), all of which are known to reliably predict lexical decision performance. In subsequent models, we then added one of the candidate composite variables to the model and use Bayes Factors to quantify the evidence in favour of each. In Table 2, we report R-squared change for each model comparison, to allow comparisons with other megastudies in the literature (e.g., Pexman, Muraki, Sidhu, Siakuluk & Yap, 2019).

Results

Overall, we found that the sensorimotor norms reliably predicted lexical decision performance for both response times and accuracy, and in both the English Lexicon and British Lexicon datasets. Each of the six composite measures accounted for a significant amount of additional variance, over and above the basic model of lexical variables (see Table 2). Log Bayes Factors for each variable (ranging from 50 to 228 for zRT, and from 29 to 138 for accuracy) revealed very strong support for their inclusion in the models. Minkowski 3 was the best performing measure in both the ELP and BLP datasets, while PCA, although still considerably improving model fit over the basic model, was the weakest performing composite measure.

Subsequently, using Minkowski 3 as the best predictor, we also found that the inclusion of action effector ratings improved model fit over and above perceptual ratings alone (see Table 3). Furthermore, adding action effector ratings provided better model fit for both ELP and BLP datasets, across both reaction time and accuracy measures.

In summary, these findings replicate the finding that perceptual information is a good predictor of people's performance in lexical decision tasks, provides new support for the utility of action effector experience in modelling cognitive performance, and shows that the findings generalise over more than one largescale data set.

General Discussion

We present a set of almost 40,000 words, normed for perceptual and action strength across 11 dimensions. The first study shows that these sensorimotor norms provide a rich dataset, with the data revealing complex patterns between various dimensions. The second study provides support for the utility of modality-specific and effector-specific sensorimotor information in modelling human performance in classic psycholinguistic tasks.

While these norms extend earlier modality-specific norms, they also quantify important new relations, such as between specific effectors and particular perceptual modalities, as well as including often ignored perceptual dimensions, such as interoception Connell, Lynott & Banks, 2018). What's more, we show that effector-specific information is also predictive of data from lexical decision tasks, over and above using perceptual-specific information alone. These findings provide evidence for a broad role for perceptual and action information in terms of their possible involvement in conceptual representations and their recruitment during cognitive processes.

A notable difference in the new set of norms is the identification of a different single composite variable that could be used in place of the full multi-dimensional vector. In the previous sets of norms (Connell & Lynott, 2012), Maximum Perceptual Strength (i.e., the highest value of any single dimension) was identified as the best single value predicting lexical decision data. In the current analyses, although max strength continued to perform very well, it was outperformed by the Minkowski 3 measure. This is an interesting pattern to emerge, as Minkowski 3 has previously been identified as an optimal parameter when modelling the integration of multiple perceptual cues (To, Baddeley, Troscianko, & Tolhurst, 2011), suggesting greater weighting to higher value dimensions. To and colleagues provided evidence that Minkowski values around 3 actually represent a general principle for perceptual integration, and may reflect the summation of neural responses to perceptual stimuli.

The current norms provide a rich source of information, and provide lexical coverage that reflects a grown adult's conceptual system. As such, we hope that they will provide many avenues for further research. There is much scope for combining the current norms with other data sets to provide even broader coverage of the human conceptual system. These and other data could then be useful for predicting human performance in a diverse array of cognitive tasks. With the increased size of the norms, they may be amenable to some machine learning techniques, for example to acquire semantic representations that could be used in robotics, or perhaps as diagnostic tools (as has been used by Kernot, Bossomaier & Bradley, 2019). Those interested in linguistics, could further investigate the role of grammatical differences in people's sensorimotor experience, and there are also opportunities to extend these norms to other languages and populations, which will enable researchers to consider cross-cultural similarities and individual differences.

Acknowledgments

This research was supported by the Leverhulme Trust, UK (RPG-2015-412, to DL, LC & MB) and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No. 682848 to LC. We thank Rosie Meers and Peidong Mei who provided help with data collection and quality control during the project, and David Balota for permitting us to include data from the English Lexicon Project in our deposited analysis files.

			Elexicon Project (N=22,297)			British Lexicon Project (N=11,768)		
Measurement	Stage	Specification	logBF$_{10}$	R^2	ΔR^2	logBF$_{10}$	R^2	ΔR^2
RT	0	Lexical predictors		0.591			0.485	
	1	PCA	54.532	0.593	0.002	50.747	0.490	0.005
		Summed strength	142.813	0.596	0.005	136.651	0.497	0.012
		Max strength	165.420	0.597	0.006	175.912	0.501	0.016
		Minkowski 10	177.86	0.598	0.007	193.544	0.502	0.017
		Euclidean	192.129	0.598	0.007	210.165	0.503	0.018
		Minkowski 3	**202.551**	**0.598**	**0.007**	**228.285**	**0.505**	**0.020**
Accuracy	0	Lexical predictors		0.237			0.286	
	1	PCA	43.377	0.241	0.004	29.613	0.290	0.004
		Summed strength	69.975	0.242	0.005	92.785	0.298	0.012
		Max strength	88.240	0.244	0.007	105.252	0.299	0.013
		Euclidean	93.375	0.244	0.007	131.714	0.302	0.016
		Minkowski 10	93.828	0.244	0.007	114.398	0.300	0.014
		Minkowski 3	**102.067**	**0.245**	**0.008**	**138.688**	**0.303**	**0.017**

Table 2 Bayesian linear regression results for Elexicon and British Lexicon Projects lexical decision data (Study 2). Lexical predictors were added to a null model at Stage 0 (LogSUBTLEX-US word frequency, orthographic length, number of syllables and orthographic Levenshtein distance)

			Elexicon Project (N=22,297)			British Lexicon Project (N=11,768)		
Measurement	Model	Specification	logBF$_{10}$	R^2	ΔR^2	logBF$_{10}$	R^2	ΔR^2
RT	Null	Lexical predictors		0.591			0.485	
	BF$_{10}$	Minkowski 3 perception	141.337	0.596	0.005	182.839	0.501	0.016
	BF$_{20}$	Minkowski 3 action	149.563	0.597	0.006	138.873	0.497	0.012
	BF$_{21}$	Comparison of simple effects	8.226			-43.966		
	BF$_{31}$		65.757	0.599		47.366	0.505	
Accuracy	Null	Lexical predictors		0.237			0.286	
	BF$_{10}$	Minkowski 3 perception	66.951	0.242	0.005	98.251	0.298	0.012
	BF$_{20}$	Minkowski 3 action	81.215	0.243	0.006	105.771	0.299	0.013
	BF$_{21}$	Comparison of simple effects	14.264			7.520		
	BF$_{31}$		37.815	0.245		46.023	0.304	

Table 3 Bayesian linear regression results for Elexicon and British Lexicon Projects lexical decision data. As above, Lexical predictors were added to a null model at Stage 0.

References

Amsel, B. D., Urbach, T. P., & Kutas, M. (2012). Perceptual and motor attribute ratings for 559 object concepts. *Behavior Research Methods*, *44*(4), 1028-1041.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., ... & Treiman, R. (2007). The English lexicon project. *Behavior research methods*, *39*(3), 445-459.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577–660.

Binder, J. R., Conant, L. L., Humphries, C. J., Fernandino, L., Simons, S. B., Aguilar, M., & Desai, R. H. (2016). Toward a brain-based componential semantic representation. *Cognitive neuropsychology*, *33*, 130-174.

Boulenger, V., Hauk, O., & Pulvermüller, F. (2008). Grasping ideas with the motor system: semantic somatotopy in idiom comprehension. *Cerebral cortex, 19(8)*, 1905-1914.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. Behavior Research Methods, 46(3), 904–11.

Chen, I. H., Chao, Q., Wang, S., Long, Y., & Huang, C. R. (2017). Exclusivity and competition of sensory modalities: evidence from Mandarin synaesthesia. *International Cognitive Linguistics Conference*, Tartu, Estonia.

Connell, L., & Lynott, D. (2010). Look but don't touch: tactile disadvantage in processing modality-specific words. *Cognition, 115*, 1-9.

Connell, L., & Lynott, D. (2012). Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition*, *125*(3), 452-465.

Connell, L., & Lynott, D. (2014). I see/hear what you mean: Semantic activation in visual word recognition depends on perceptual attention. *Journal of Experimental Psychology: General, 143*, 527-533.

Connell, L., & Lynott, D. (2015). Embodied semantic effects in visual word recognition. *Foundations of embodied cognition*, *2*, 71-89.

Connell, L., & Lynott, D. (2016). Do we know what we're simulating? Information loss on transferring unconscious perceptual simulation to conscious imagery. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(8), 1218.

Connell, L. M., Lynott, D. J., & Banks, B. (2018). Interoception: the forgotten modality in perceptual grounding of abstract and concrete concepts. *Philosophical Transactions B: Biological Sciences*.

Đurđević, D. F., Popović Stijačić, M., & Karapandžić, J. (2016) A quest for sources of perceptual richness: Several candidates. In S. Halupka-Rešetar and S. Martínez-Ferreiro (Eds.) *Studies in Language and Mind* (pp. 187-238). RS, Novi Sad.

Fernandino, L., Conant, L. L., Binder, J. R., Blindauer, K., Hiner, B., Spangler, K., & Desai, R. H. (2013). Where is the action? Action sentence processing in Parkinson's disease. *Neuropsychologia, 51(8),* 1510-1517.

Hauk, O., Johnsrude, I. & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41*, 301-307.

Kaschak, M. P., Zwaan, R. A., Aveyard, M., & Yaxley, R. H. (2006). Perception of auditory motion affects language processing. *Cognitive Science*, *30*(4), 733-744.

Kernot, D., Bossomaier, T., & Bradbury, R. (2019). The stylometric impacts of ageing and life events on identity. *Journal of Quantitative Linguistics*, *26*(1), 1-21.

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior research methods*, *44*(1), 287-304.

Lynott, D. & Connell, L. (2009). Modality exclusivity norms for 423 object properties. *Behavior Research Methods, 41*, 558-564.

Lynott, D., & Connell, L. (2013). Modality exclusivity norms for 400 nouns: The relationship between perceptual experience and surface word form. *Behavior Research Methods, 45*, 516-526.

Meier, J. D., Aflalo, T. N., Kastner, S., & Graziano, M. S. (2008). Complex organization of human primary motor cortex: a high-resolution fMRI study. *Journal of neurophysiology, 100(4)*, 1800-1812.

Miklashevsky, A. (2018). Perceptual Experience Norms for 506 Russian Nouns: Modality Rating, Spatial Localization, Manipulability, Imageability and Other Variables. *Journal of psycholinguistic research*, *47*, 641-661.

Pexman, P. M., Muraki, E., Sidhu, D. M., Siakaluk, P. D., & Yap, M. J. (2019). Quantifying sensorimotor experience: Body–object interaction ratings for more than 9,000 English words. *Behavior research methods*, 1-14.

Pulvermüller, F., Hauk, O., Nikulin, V. V., & Ilmoniemi, R. J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience*, *21*, 793-797.

Smith, L., & Gasser, M. (2005). The development of embodied cognition: Six lessons from babies. *Artificial life*, *11*(1-2), 13-29.

Speed, L. J., & Majid, A. (2017). Dutch modality exclusivity norms: Simulating perceptual modality in space. *Behavior research methods*, *49*(6), 2204-2218.

Spence, C., Nicholls, M. E., & Driver, J. (2001). The cost of expecting events in the wrong sensory modality. *Perception & Psychophysics*, *63*(2), 330-336.

Tillotson, S. M., Siakaluk, P. D., & Pexman, P. M. (2008). Body—object interaction ratings for 1,618 monosyllabic nouns. *Behavior Research Methods*, *40*(4), 1075-1078.

To, M. P. S., Baddeley, R. J., Troscianko, T., & Tolhurst, D. J. (2010). A general rule for sensory cue summation: evidence from photographic, musical, phonetic and cross-modal stimuli. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1710), 1365-1372.

Vinson, D. P., & Vigliocco, G. (2008). Semantic feature production norms for a large set of objects and events. *Behavior Research Methods*, *40*(1), 183-190.

Winter, B., Perlman, M., Perry, L. K., & Lupyan, G. (2017). Which words are most iconic?. *Interaction Studies*, *18*(3), 443-464.

Winter, B., Perlman, M., & Majid, A. (2018). Vision dominates in perceptual language: English sensory vocabulary is optimized for usage. *Cognition*, *179*, 213-220.

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: motor resonance in language comprehension. *Journal of Experimental Psychology: General, 135*, 1-11.