



Discussion

Does using a foreign language reduce mental imagery?

Guillermo Montero-Melis^{a,b,*}, Petrus Isaksson^c, Jeroen van Paridon^a, Markus Ostarek^a^a Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, the Netherlands^b Department of Linguistics, Stockholm University, SE-106 91 Stockholm, Sweden^c Centre for Research on Bilingualism, Stockholm University, SE-106 91 Stockholm, Sweden

ARTICLE INFO

Keywords:

Mental imagery
 Visual imagery
 Foreign language
 Conceptual processing
 Language proficiency
 Foreign language effect

ABSTRACT

In a recent article, Hayakawa and Keysar (2018) propose that mental imagery is less vivid when evoked in a foreign than in a native language. The authors argue that reduced mental imagery could even account for moral foreign language effects, whereby moral choices become more utilitarian when made in a foreign language. Here we demonstrate that Hayakawa and Keysar's (2018) key results are better explained by reduced language comprehension in a foreign language than by less vivid imagery. We argue that the paradigm used in Hayakawa and Keysar (2018) does not provide a satisfactory test of reduced imagery and we discuss an alternative paradigm based on recent experimental developments.

1. Introduction

Are our mental images less vivid when speaking a foreign language than when speaking our native tongue? For many models of how words are linked to conceptual representations the answer should be *no*, because bilinguals will access a common semantic system in both of their languages (e.g., Caramazza & Brones, 1980; Francis, 1999; Illes et al., 1999; Kroll & Stewart, 1994; Paradis, 2004), which in turn should lead to common imagery processes. Yet, in a thought-provoking paper recently published in *Cognition*, Hayakawa and Keysar (2018, henceforth HK18) suggest the answer to be *yes*. Based on the results of two experiments (their experiments 1 and 2), HK18 conclude that visual imagery is reduced in a foreign language compared to the native language. In a final experiment, HK18 test whether reduced visual imagery might partially explain the foreign language effect in moral decision making (Costa et al., 2014). Here we focus on their central proposal: If you are a native speaker of English, the word “window” will evoke a clearer, more vivid mental image of a window than its Spanish counterpart *ventana*, even if you speak Spanish and you know perfectly well that *ventana* means window. We argue that this conclusion does not follow from the findings presented by HK18. Instead, their results are better explained by reduced language comprehension in a foreign language.

But why should mental imagery be reduced in a foreign language in the first place? The explanation advanced by HK18 is that vividness of

imagery depends on the amount of sensory memories associated with a word and how easy it is to access these memories. The typical foreign language speaker has interacted with the world through their native language to a far greater extent than through their foreign language. There is evidence that memories are richer and better recalled when the language used at retrieval matches the linguistic environment at encoding (Marian & Fausey, 2006; Marian & Neisser, 2000). In combination with the well-established view that the episodic memory system is used for simulation of hypothetical events (Schacter, Addis, & Buckner, 2007), this makes it plausible that imagery is reduced in L2: To the extent that less detailed memories are triggered by a foreign word (*ventana*) than by its native counterpart (window), the image we form in the mind's eye could be less vivid, even though we may fully grasp the meaning of *ventana*. We find this line of explanation compelling. However, for this argument it is crucial that 1) participants understand the meaning of the words in the foreign language and 2) the paradigm tests mental imagery.

2. Experiment 1 has problems that extend to experiment 3

In Experiment 1, participants were asked to rate on a 7-point scale the vividness of a range of sensory experiences from different modalities (visual, auditory, tactile, kinaesthetic, gustatory, olfactory, organic). Imagery of sensory experiences was always triggered by a verbal description.¹ The crucial manipulation was between subjects:

* Corresponding author at: Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, the Netherlands.

E-mail address: guillermo.monteromelis@mpi.nl (G. Montero-Melis).

¹ An example of a visual item: “Think of some relative or friend whom you frequently see, carefully considering the visual image that comes to your imagination. Rate how vividly you can imagine the precise gait, length of step, etc. in walking.” The exact materials used in Experiments 1–3 were not made publicly available with the original publication, but the corresponding author kindly shared the English version of the questionnaire in Experiment 1 and the items used in Experiment 2 (S. Hayakawa, personal communications, August 10 and September 9, 2018).

Participants were native speakers of English with Spanish as a foreign language and they were randomly assigned to complete the questionnaire in either English (their native language) or Spanish (their foreign language). The analyses show that imagery ratings were significantly lower in the foreign than in the native language.

HK18 themselves question whether Experiment 1 constitutes a reliable test of mental imagery. First, they note that scale labels in a foreign language might lead to less extreme ratings because of unfamiliarity with the labels. Second, they question the validity of self-report measures because of their inherent subjectivity and because self-report questionnaires of imagery have been shown to correlate poorly with objective imagery measures. Additionally, there is a potential confound in the design of Experiment 1 that is not addressed by the authors: The language of the questionnaire (English/Spanish) was perfectly confounded with the nativeness of language (native/foreign). Therefore, it is possible that the Spanish version of the experiment could have biased vividness ratings downwards independently of whether Spanish was a participant's native or foreign language (e.g., due to the particular word choice, see Clark & Watson, 1995). In previous studies, Hayakawa and colleagues have counterbalanced their design by crossing the actual language and its native vs foreign status (e.g., Hayakawa, Tannenbaum, Costa, Corey, & Keysar, 2017: Experiments 5 and 6; see also Costa et al., 2014: Experiment 2). Such a design would have allowed for a stronger inference here as well.

More importantly, even assuming the English and Spanish versions were perfectly equivalent, a stronger bias towards the middle of the scale in the foreign language would also be expected if participants did not fully understand all words in a description. Consider the following item from the tactile modality: "Think of 'feeling' or touching... the warmth of a tepid bath". While "warmth" and "bath" are relatively frequent words (4.4 and 31.1 occurrences per million words [pmw], respectively), the adjective "tepid" has a low frequency (0.2 pmw).² For each item, participants had the option to select "do not understand" (p. 9); however, because an item typically consisted of a description (rather than a single word), a participant might comprehend enough not to choose the "do not understand" option, but still have some uncertainty as to the exact meaning of the description. For the previous example, participants might understand "the warmth of a bath" but not the low-frequency adjective "tepid" that modifies it and makes the situation more specific. But our argument is not contingent on the presence of low frequency words – it holds if there are words that just some participants did not know, and even relatively high frequency words might have not been correctly understood occasionally. In general, uncertainty will bias Likert scale ratings towards the middle of the scale (Douven, 2018), which is exactly the tendency that was observed in the foreign language condition in Experiment 1.

Although our comment focuses on the claim of reduced visual imagery itself, the concern above extends to Experiment 3. In this experiment, reduced visual imagery was suggested to underlie more utilitarian choices in a moral dilemma (the footbridge dilemma). If our argument is right, participants who are uncertain about the precise meaning of a prompt in the foreign language will generally show a tendency to answer towards the midpoint of a self-report scale. This should hold whether they are rating the vividness of a mental picture or the likelihood that they would push one man from a footbridge to save five lives. And this is indeed what the results of Experiment 3 show: The effect of rating in a foreign compared to the native language is a drift towards the midpoint of the 7-point scale. For the moral dilemma, mean ratings move from 2.15 in the native language to 2.41 in the foreign language, where 1 is "definitely would not push him" and 7 "definitely would push him" (HK18, p.12). Given the low initial baseline, it is

impossible to tease apart a genuine boost in utilitarian moral choices from an increased central tendency bias. Suboptimal comprehension could be the variable that drives the same midpoint bias for imagery and moral choice in a foreign language.

Returning to our main argument: For the reasons adduced by the authors themselves, and for the additional concerns raised above, Experiment 1 does not provide convincing evidence for reduced visual imagery in a foreign language. But perhaps Experiment 2 does?

3. Experiment 2: Limited understanding of words in the foreign language will impair the shape task more than the category task

In contrast to Experiment 1, HK18 present Experiment 2 as an "objective behavioral measure to investigate the vividness of mental imagery" (p. 12). Participants were native Mandarin speakers with English as a foreign language. The basic task involved one of two versions of an odd-one-out paradigm: In each trial, participants saw groups of three words (e.g., violin, piano and hourglass) and were asked "to click on the one that was least like the other two based on a given attribute" (HK18, p. 10). The attribute could either be *category* (odd one out: hourglass, because it is not a musical instrument) or *shape* (odd one out: piano, because it does not have a double-rounded shape like a violin and an hourglass). The experiment thus consisted of a 2-by-2 design with the factor Language (foreign vs native language, between subjects) and Task (shape vs category, within subjects). A decrease in accuracy in the foreign language condition for the shape task larger than for the category task (i.e., a Language-by-Task interaction) was taken as an indication that visual imagery was reduced in the foreign language despite understanding the words. The critical assumptions made by the authors were that 1) the task measured vividness of mental imagery and 2) the shape and category tasks "required the same level of knowledge of the language" (p. 11). We next argue that both assumptions are problematic.

First, it is important to note that the task tests whether participants *have knowledge* about the semantic category of words and about the typical shapes of their referents. It does not tell us whether participants used *mental imagery* to solve the task, in the sense of activating information in a depictive format (Dijkstra, Bosch, & van Gerven, 2019; Pearson & Kosslyn, 2015). The assumption that vivid imagery is required to solve the shape task is highly speculative, as we argue in the next section. What then explains that L2 speakers performed worse in the shape task? By its very nature, the category task can be solved without knowing the specific meaning of a word; it is enough to know to which superordinate category a word belongs. For instance, a foreign language user might be certain that pianos and violins are both tokens of the broader category of music instruments, without being quite sure which instruments exactly they denote. This alone suffices to exclude "hourglass" and get the category trial right, even if they have no clue what "hourglass" means. However, to accurately respond to the shape task, this approximate knowledge is *not* enough. To match for shape, one needs to know precisely which instruments "piano" and "violin" refer to, and also which kind of object is denoted by "hourglass". Without this specific knowledge about the entities denoted by the words, the shape task cannot be solved. Thus, the shape task requires more specific lexical knowledge (not more vivid imagery) than the category task.

This argument would be ruled out if we could be sure that participants in fact understood all words at the specific level required for the shape task. That is, did learners know which specific instrument is denoted by "violin" rather than just knowing that it is a musical instrument? We cannot tell. HK18 included the following comprehension check for Experiment 2, which suffers from a similar problem as the main experiment: Participants (after the experiment) were shown each of the stimulus words again one by one, together with three line drawings, and were asked to match the word with the correct drawing. If a word was incorrectly matched, all trials containing that word were

² All frequency counts were obtained from the SUBTLEX-US corpus (Brybaert & New, 2009), retrieved 10 June 2019 from <https://www.ugent.be/pp/experimentele-psychologie/en/research/documents/subtlexus/overview.htm>.

removed from the analysis for that participant. First, this does not control for chance-level performance; that is, on average one third of the comprehension checks would be correct if a participant was answering randomly. Second, participants again need not know the exact meaning of a word to solve this comprehension check. Suppose the word is “hourglass” and the learner does not know what it means. If two of the pictures depict objects for which they *do* know the names (e.g., violin and piano), they can easily deduce that the word must denote the object in the 3rd drawing. Knowing the label for just two of the three drawings will always suffice to correctly solve a trial in this task even if one initially has no idea what the word means.

There is an additional reason why participants might have fared worse on the shape task due to suboptimal comprehension. It is well established that there exists a strong relationship between semantic similarity and word co-occurrence patterns (Firth, 1957; Griffiths, Steyvers, & Tenenbaum, 2007; Z. Harris, 1954; Landauer & Dumais, 1997; Mikolov, Sutskever, Chen, Corrado, & Dean, 2013, among others). People are sensitive to co-occurrence patterns and techniques that exploit these purely statistical regularities are remarkably accurate at approximating human performance on a wide variety of conceptual tasks (Landauer & Dumais, 1997; Landauer, Foltz, & Laham, 1998; Pereira, Gershman, Ritter, & Botvinick, 2016). Foreign language learners will have picked up such co-occurrence statistics through exposure to the language (Ellis, 2002). To the extent that it can be shown that category matches (e.g., violin–piano) have a stronger statistical association than shape matches (e.g., violin–hourglass), participants could be solving the category task simply by following their intuitions about which words tend to occur in similar contexts, even if they do not know the precise meaning of each of the words (note that in this case, we do not even need to assume learners know to which superordinate category a word belongs). Fig. 1 plots the pairwise similarity for the stimuli used in HK18's Experiment 2, as computed from co-occurrence statistics in a large corpus of English.³ As one would expect, cosine similarities are consistently higher for category than for shape matches. This means that even just having a hunch that “violin” and “piano” tend to appear in similar contexts, but “hourglass” does not share many contexts with either, will be enough to solve the category task. For the shape task, however, co-occurrence statistics will not come to our rescue.

In sum, we have shown two ways in which reduced *comprehension* in the foreign language could have produced the results in HK18's Experiment 2: First, the category task requires less word knowledge than the shape task and the comprehension checks do not establish that learners understood the stimulus words at a specific level, and second, implicit knowledge of word co-occurrence patterns would lead to accurate responses in the category but not in the shape task. As pointed out by the authors: “After all, it would not be surprising if those using a foreign language visualize scenes less vividly if they are unable to comprehend the descriptions” (p. 9). We think this less exciting explanation is indeed the most likely one.

4. Looking forward: How should we test (reduced) visual imagery in a foreign language?

We agree with HK18 that reduced visual imagery in a foreign language would be an intriguing finding of great theoretical relevance. We have nevertheless argued that this claim is not justified by their empirical evidence, as there is a simpler and more compelling alternative explanation for their results. What, then, would be a better test of

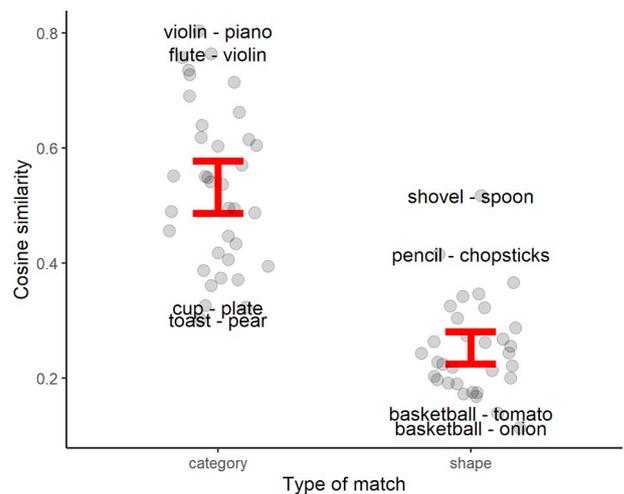


Fig. 1. Cosine similarity of shape and category matches in the stimuli from HK18 Experiment 2. Each point represents a pair of words that was either a category match (left) or a shape match (right). Cosine similarity was computed using subs2vec word vectors (van Paridon & Thompson, 2019). Error bars show non-parametric 95% confidence intervals. The two pairs with the highest and lowest cosine similarities are shown for each type of match.

reduced visual imagery?

It is useful to first consider the phenomenology of the claim: What would it mean for visual imagery to be reduced in a foreign language *despite full understanding*? Imagine you know that the Spanish word *plátano* means banana. It is unlikely that you could know this and yet fail to determine that the shape of a *plátano* is more like that of a flute than that of an apple. You do not need much “access to episodic memory” (p. 13) to solve this. In our opinion, this makes the test in HK18 a poor candidate: Comparing accuracy in an untimed odd-one-out task means one is essentially testing if participants are *at all* capable of accessing information about the shape of a word referent, irrespective of how they are accessing it. There is no compelling reason to believe that the paradigm measures mental imagery, as it merely tests whether participants know the semantic category and shape of word referents. We think that, if an effect exists, it probably manifests itself as a subtle difference rather than a failure to access a highly schematic visual representation of the referent (which is all that was required for HK18's Experiment 2). Therefore, appropriate tests need to be sensitive to small differences in visual imagery and, critically, they need to unambiguously tap into visual imagery.

In recent years, sophisticated paradigms have been developed to study language-induced top-down effects on visual processing. Continuous flash suppression (CFS; Tsuchiya & Koch, 2005), a binocular rivalry technique where a picture presented to one eye is suppressed by presenting rapidly changing flashes of distractor-images to the other eye, is a particularly promising method for these purposes. The ability to detect a suppressed picture in CFS is determined by how efficiently suppressed pictures are processed in the visual cortex (Yuval-Greenberg & Heeger, 2013). Increasing evidence suggests that conscious access to visual information, as required for visual imagery, involves top-down activation of the visual cortex (Dijkstra et al., 2019; Kok, Bains, van Mourik, Norris, & de Lange, 2016; Lawrence et al., 2018; Pearson & Kosslyn, 2015) leading to enhanced processing of compatible visual stimuli (Gayet, Paffen, & Van der Stigchel, 2013; Harris & Mrcic-Flogel, 2013). Recent studies specifically show that hearing words can enhance detection sensitivity in CFS (Lupyan & Ward, 2013; Ostarek & Huettig, 2017). Thus, on the assumption that vividness of imagery modulates the degree to which visual cortex is activated (Dijkstra, Bosch, & van Gerven, 2017; Kosslyn, Ganis, & Thompson, 2001; Kosslyn, Thompson, Klm, & Alpert, 1995), a critical test of reduced imagery in a foreign language is that words in the native

³ Specifically, we computed cosine similarities for word vectors (van Paridon & Thompson, 2019) derived from a large corpus of subtitles (Lison & Tiedemann, 2016) and Wikipedia articles, using the fastText algorithm (Bojanowski, Grave, Joulin, & Mikolov, 2016). The word vectors are numerical representations of single words, therefore multiword noun phrases were excluded from the analysis.

language should enhance detection in CFS more strongly than words in a foreign language. If one implements comprehension checks that unambiguously ensure word comprehension (e.g., having participants translate the words to their native language after the experiment, ideally under time pressure to minimize strategic guessing⁴), such paradigms can be considered a strong test of potential differences in vividness of imagery in native and foreign language because differential detection sensitivity is only expected if imagery vividness indeed differs by language.

Funding

This work was supported by the Swedish Research Council [grant number 2018-00245 to Guillermo Montero-Melis].

Appendix A. Supplementary material

Supplementary data and scripts associated with this article can be downloaded at <https://doi.org/10.7910/DVN/HZNRJU>.

References

- Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2016). Enriching word vectors with subword information. ArXiv Preprint ArXiv:1607.04606.
- Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, 41(4), 977–990. <http://doi.org/10.3758/BRM.41.4.977>.
- Caramazza, A., & Brones, I. (1980). Semantic classification by bilinguals. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, 34(1), 77–81. <http://doi.org/10.1037/h0081016>.
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309–319. <http://doi.org/10.1037/1040-3590.7.3.309>.
- Costa, A., Foucart, A., Hayakawa, S., Aparici, M., Apesteguía, J., Heafner, J., & Keysar, B. (2014). Your morals depend on language. *PLoS One*, 9(4), e94842. <http://doi.org/10.1371/journal.pone.0094842>.
- Dijkstra, N., Bosch, S. E., & van Gerven, M. A. J. (2017). Vividness of visual imagery depends on the neural overlap with perception in visual areas. *Journal of Neuroscience*, 37(5), 1367–1373. <http://doi.org/10.1523/JNEUROSCI.3022-16.2016>.
- Dijkstra, N., Bosch, S. E., & van Gerven, M. A. J. (2019). Shared neural mechanisms of visual perception and imagery. *Trends in Cognitive Sciences*, 23(5), 423–434. <http://doi.org/10.1016/j.tics.2019.02.004>.
- Douven, I. (2018). A Bayesian perspective on Likert scales and central tendency. *Psychonomic Bulletin & Review*, 25(3), 1203–1211. <http://doi.org/10.3758/s13423-017-1344-2>.
- Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24(02), 143–188. <http://doi.org/10.1017/S0272263102002024>.
- Firth, J. R. (1957). A synopsis of linguistic theory 1930–1955. *Studies in linguistic analysis* (pp. 1–32). Oxford: Blackwell.
- Francis, W. S. (1999). Cognitive integration of language and memory in bilinguals: Semantic representation. *Psychological Bulletin*, 125(2), 193–222. <http://doi.org/10.1037/0033-2909.125.2.193>.
- Gayet, S., Paffen, C. L. E., & Van der Stigchel, S. (2013). Information matching the content of visual working memory is prioritized for conscious access. *Psychological Science*, 24(12), 2472–2480. <http://doi.org/10.1177/0956797613495882>.
- Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review*, 114(2), 211–244. <http://doi.org/10.1037/0033-295X.114.2.211>.
- Harris, K. D., & Mrsic-Flogel, T. D. (2013). Cortical connectivity and sensory coding. *Nature*, 503(7474), 51–58. <http://doi.org/10.1038/nature12654>.
- Harris, Z. (1954). Distributional structure. *Word*, 10(2/3), 146–162.
- Hayakawa, S., & Keysar, B. (2018). Using a foreign language reduces mental imagery. *Cognition*, 173, 8–15. <http://doi.org/10.1016/j.cognition.2017.12.010>.
- Hayakawa, S., Tannenbaum, D., Costa, A., Corey, J. D., & Keysar, B. (2017). Thinking more or feeling less? Explaining the foreign-language effect on moral judgment. *Psychological Science*, 28(10), 1387–1397. <http://doi.org/10.1177/0956797617720944>.
- Illes, J., Francis, W. S., Desmond, J. E., Gabrieli, J. D. E., Glover, G. H., Poldrack, R., ... Wagner, A. D. (1999). Convergent cortical representation of semantic processing in bilinguals. *Brain and Language*, 70(3), 347–363. <http://doi.org/10.1006/brln.1999.2186>.
- Kok, P., Bains, L. J., van Mourik, T., Norris, D. G., & de Lange, F. P. (2016). Selective activation of the deep layers of the human primary visual cortex by top-down feedback. *Current Biology*, 26(3), 371–376.
- Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience*, 2(9), 635.
- Kosslyn, S. M., Thompson, W. L., Klm, I. J., & Alpert, N. M. (1995). Topographical representations of mental images in primary visual cortex. *Nature*, 378(6556), 496.
- Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33(2), 149–174. <http://doi.org/10.1006/jmla.1994.1008>.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104(2), 211–240. <http://doi.org/10.1037/0033-295X.104.2.211>.
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, 25(2–3), 259–284. <http://doi.org/10.1080/01638539809545028>.
- Lawrence, S. J. D., van Mourik, T., Kok, P., Koopmans, P. J., Norris, D. G., & de Lange, F. P. (2018). Laminar organization of working memory signals in human visual cortex. *Current Biology*, 28(21), 3435–3440.e4. <https://doi.org/10.1016/j.cub.2018.08.043>
- Lison, P., & Tiedemann, J. (2016). *OpenSubtitles2016: Extracting large parallel corpora from movie and TV subtitles*. (LREC).
- Lupyan, G., & Ward, E. J. (2013). Language can boost otherwise unseen objects into visual awareness. *Proceedings of the National Academy of Sciences*, 110(35), 14196–14201. <http://doi.org/10.1073/pnas.1303312110>.
- Marian, V., & Fausey, C. M. (2006). Language-dependent memory in bilingual learning. *Applied Cognitive Psychology*, 20(8), 1025–1047. <http://doi.org/10.1002/acp.1242>.
- Marian, V., & Neisser, U. (2000). Language-dependent recall of autobiographical memories. *Journal of Experimental Psychology: General*, 129(3), 361–368. <http://doi.org/10.1037/0096-3445.129.3.361>.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality. *Advances in Neural Information Processing Systems 26 (NIPS 2013)*. Retrieved from <http://arxiv.org/abs/1310.4546>
- Ostarek, M., & Huettig, F. (2017). Spoken words can make the invisible visible—Testing the involvement of low-level visual representations in spoken word processing. *Journal of Experimental Psychology: Human Perception and Performance*, 43(3), 499–508. <http://doi.org/10.1037/xhp0000313>.
- Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Vol. 18. Amsterdam: John Benjamins.
- Pearson, J., & Kosslyn, S. M. (2015). The heterogeneity of mental representation: Ending the imagery debate. *Proceedings of the National Academy of Sciences*, 112(33), 10089–10092. <http://doi.org/10.1073/pnas.1504933112>.
- 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. <http://doi.org/10.1080/02643294.2016.1176907>.
- Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: The prospective brain. *Nature Reviews Neuroscience*, 8(9), 657–661. <http://doi.org/10.1038/nrn2213>.
- Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative after-images. *Nature Neuroscience*, 8(8), 1096. <http://doi.org/10.1038/nn1500>.
- van Paridon, J., & Thompson, B. (2019). subs2vec: Word embeddings from subtitles in 55 languages. <https://doi.org/10.31234/osf.io/fcrmy>
- Yuval-Greenberg, S., & Heeger, D. J. (2013). Continuous flash suppression modulates cortical activity in early visual cortex. *Journal of Neuroscience*, 33(23), 9635–9643. <http://doi.org/10.1523/JNEUROSCI.4612-12.2013>.

⁴ We thank an anonymous reviewer for this suggestion.