Journal of Experimental Psychology: Learning, Memory, and Cognition

Warm and Fuzzy: Perceptual Semantics Can Be Activated Even During Shallow Lexical Processing

Olesia Platonova and Alex Miklashevsky

Online First Publication, February 13, 2025. https://dx.doi.org/10.1037/xlm0001429

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: Learning, Memory, and Cognition

ISSN: 0278-7393

https://doi.org/10.1037/xlm0001429

Warm and Fuzzy: Perceptual Semantics Can Be Activated Even During Shallow Lexical Processing

Olesia Platonova[1, 2] and Alex Miklashevsky[3, 4, 5]

[1] Neural Coding and Neuroengineering of Human Speech Functions, Institut Pasteur, Université Paris Cité, Hearing Institute
[2] Laboratory of Linguistic Anthropology, Department of General, Computational, and Cognitive Linguistics, Tomsk State University
[3] Museum für Naturkunde Berlin, Leibniz Institute for Evolution and Biodiversity Science, Berlin, Germany
[4] Potsdam Embodied Cognition Group, Faculty of Human Sciences, University of Potsdam
[5] Brain Language Laboratory, Department of Philosophy and Humanities, Free University of Berlin

OPEN DATA OPEN MATERIALS

According to the embodied cognition view, activation of perceptual semantics (such as visual information for the words “white” or “red” or tactile information for the words “warm” or “fuzzy”) should occur even in a relatively shallow lexical decision task. While some studies found this activation, other studies did not. We argue that minimizing the time gap between the stimuli is crucial for detecting the activation of perceptual semantics in this task. Furthermore, we suggest that modalities should be analyzed separately due to their possible qualitative differences. We designed two experiments addressing these points in Russian (Experiment 1) and German (Experiment 2) languages. We selected visual, tactile, and auditory adjectives (e.g., “white,” “warm,” and “loud,” respectively) and assessed lexical decision times for two stimuli at once (e.g., “white + fuzzy”), thus eliminating the time gap between the two stimuli. Our analysis accounted for word length, frequency, and shallow lexical associations between presented words. Overall, the results of both experiments demonstrated that perceptual semantics is indeed activated even during shallow lexical processing, such as in the lexical decision task. Importantly, in line with our predictions, the effect of perceptual semantics was not identical across all modalities. More specifically, there was a consistent advantage for processing visual semantics and a consistent disadvantage for processing haptic semantics. Thus, the exact combination of semantic modalities modulates the activation of modality information. Our results strongly support the embodied view of language semantics.

Keywords: semantic modality, perceptual semantics, Modality Switch effect, lexical decision, embodied cognition

Supplemental materials: https://doi.org/10.1037/xlm0001429.supp

Heather Ferguson served as action editor.

Alex Miklashevsky https://orcid.org/0000-0002-2693-5344

Data sets generated for this study, stimuli materials, and the protocol of power analysis can be found on the Open Science Framework (Platonova & Miklashevsky, 2022) at https://osf.io/a85h7/. This study was not preregistered. The research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflicts of interest.

The work of Olesia Platonova was supported by the Tomsk State University Development Programme (Priority 2030) in 2020–2021. The funder had no role in the study design, data collection and analysis, decision to publish, or article preparation. The authors thank Katharina Kühne for her helpful consultations on programming the online experiments and the entire Potsdam Embodied Cognition Group for valuable discussions of the study design and results. They thank Zoya I. Rezanova for fruitful discussions during the study planning and Martin H. Fischer, Andriy Myachykov, and Sophie Bouton for their valuable feedback on the earlier version of this article. They also thank the Potsdam Embodied Cognition Group research assistants, Taisiia Ulianova and Daniel Janko, for their help in obtaining data on word co-occurrence.

Olesia Platonova played a lead role in conceptualization, funding acquisition, investigation, and validation and an equal role in data curation, formal analysis, methodology, project administration, software, visualization, writing–original draft, and writing–review and editing. Alex Miklashevsky played a lead role in supervision and an equal role in conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, validation, visualization, writing–original draft, and writing–review and editing.

The data are available at https://osf.io/a85h7/

The experimental materials are available at https://osf.io/a85h7/

Correspondence concerning this article should be addressed to Alex Miklashevsky, Museum für Naturkunde Berlin, Leibniz Institute for Evolution and Biodiversity Science, Invalidenstrasse 43, 10115 Berlin, Germany. Email: armanster31@gmail.com

The embodied cognition approach suggests that language is grounded in sensorimotor experiences. In other words, embodied experiences (actions, senses, and emotions) represent the basis of conceptual processing. According to a perceptual symbols theory first introduced by Barsalou, activation of a concept requires *a mental simulation* of that concept in its relevant modalities (e.g., visual, auditory, or haptic; Barsalou, 1999, 2008, 2020). Thus, from the embodied cognition perspective, processing language partially relies on the same sensorimotor brain areas that are responsible for perception of real objects or performing real actions (Bergen, 2015; Fischer & Zwaan, 2008; Hauk et al., 2004; Pecher et al., 2003; but see Bottini et al., 2022).

The embodied view on semantics underscores the activation of sensory modalities during lexical processing, that is, word semantics corresponding to a particular perceptual property, such as vision, hearing, or touch (e.g., Chen et al., 2019; Harpaintner et al., 2020; Kiefer et al., 2008; Lynott & Connell, 2009, 2013; Lynott et al., 2020; Miklashevsky, 2018; Speed & Majid, 2017; Ulrich et al., 2023; van Dantzig et al., 2011; Vergallito et al., 2020; Vermeulen et al., 2008). To illustrate, the word "white" denotes a quality perceived through sight, and thus, its understanding should involve visual brain areas, that is, its dominant semantic modality is visual. Similarly, the auditory semantic modality can be illustrated by the word "loud," which should engage auditory brain areas during its comprehension, that is, its dominant semantic modality is auditory.

When a semantic Modality Switch occurs (e.g., the word "white" immediately follows the word "loud"), participants' cognitive resources are spent on switching attention from one semantic modality to another, which might be reflected in slower responses to the latter stimulus. This phenomenon is known as the linguistic *Modality Switch effect* (MSE; Pecher et al., 2003). Pecher et al. (2003) used a property verification task—a decision on whether the property is true of a concept—to measure the MSE. In their study, participants were presented with short phrases (e.g., "blender *can be* loud") and then asked to verify whether a property is typical for the given concept. The authors found that verification of a property belonging to a specific modality (e.g., an auditory property: "blender *can be* loud") is slower after verifying a property of a different modality (e.g., a gustatory property: "cranberries *can be* tart") as compared to the condition when a property of the same modality was shown right before (e.g., an auditory property: "leaves *can be* rustling"). This MSE was successfully replicated in further studies (e.g., Hald et al., 2011; Lynott & Connell, 2009; Pecher et al., 2004; Scerrati et al., 2015).

The following studies ruled out alternative explanations of the MSE. First, the MSE does not seem to be due to shallow semantic associations: Pecher et al. (2003) controlled for semantic associations between target stimuli using Nelson et al.'s (2004) associative norms. Pecher et al. demonstrated that the MSE appears irrespectively of associative links between target stimuli. More recently, Scerrati et al. (2015) also included semantically associated stimuli (e.g., "sleepy" for "bed") as fillers alongside the usual unassociated stimuli (e.g., "unripe" for "bed"). Such manipulation of priming materials increased the depth of processing for target trials. Again, there was no impact on the MSE.

Second, mental imagery also seems not to account for the MSE: Pecher et al. (2009) did not find any systematic relationship between the MSE and individual imagery abilities (as measured with spatial and object-oriented visual imagery questionnaires; though see recent studies by Muraki et al., 2023; Muraki & Pexman, 2021, for opposite results in the motor domain). To summarize, research suggested that modality information is activated regardless of lexical associations and, probably, even mental imagery, and therefore, modality information might constitute a necessary component in word comprehension.

However, some researchers argue that the embodied effects found in previous studies are functionally irrelevant or secondary to language processing (Mahon & Caramazza, 2008; Tomasino & Rumiati, 2013; see Meteyard et al., 2012, for a review). Indeed, several studies demonstrate inconsistent embodied effects in language processing: Sometimes embodied information has facilitatory and sometimes inhibitory effects (see Shebani & Pulvermüller, 2018, for discussion) or no effect at all (e.g., no MSE in the lexical decision task by Scerrati et al., 2017; see also Petrova et al., 2018, for similar null findings in the spatial domain).

This critique has led to a further search for factors influencing embodied effects in language processing. Converging evidence points to the timing of processing and the role of context as possible missing ingredients modulating these embodied effects (see Estes & Barsalou, 2018, for the spatial domain; see also García & Ibáñez, 2016, for the motor domain). Below, we are discussing factors that might contribute to the variability in processing embodied semantics: (a) the depth of semantic processing required in an experimental task and (b) the time frames of processing perceptual semantics. We will also briefly discuss the varying association strength between semantic modalities and the possible methodological consequences of this variability. We will then present our novel paradigm for studying the MSE while taking into account the potential limitations of the previous research on this topic.

Experimental Task: Depth of Semantic Processing

Successful demonstrations of the MSE have primarily involved a property verification task which is characterized by two parameters: First, it is a deep semantic task, in the sense, that it might require voluntary access to sensorimotor representations, at least to some extent, which is not always necessary in language processing. Second, in this task, words are presented in a context (e.g., in the phrase "blender *can be* loud," the word "blender" becomes a context word for the target word "loud"). In the following paragraphs, we discuss the depth of processing and the role of context.

Deeper semantic tasks, such as the property verification task discussed above or the semantic judgment task, proved their ability to activate embodied mental representations (Hald et al., 2011; Kuhnke et al., 2020; Lynott & Connell, 2009; Pecher et al., 2004; Scerrati et al., 2015; see also Willems & Casasanto, 2011, for the motor domain). However, this evidence can easily be disregarded by the opponents of embodied semantic theories as irrelevant: The property verification task explicitly focuses participants' attention on sensory information, and thus, it might not be surprising that sensory properties become activated in this task. In amodal theories of cognition, language semantics is seen as generally independent from perception and action systems (Landauer & Dumais, 1997; see Meteyard et al., 2012, for a review), and sensorimotor information only becomes activated under certain circumstances, such as when the task explicitly requires this activation. Previous studies of the MSE with a classical lexical decision task—that is, discrimination between words and meaningless strings of letters reminding real words or so-called pseudowords—demonstrated that this shallow word processing might simply not suffice to activate sensorimotor representations (e.g., Scerrati et al., 2017). If this is the case, then activation of sensorimotor representations might indeed be a mere by-product of the property verification task used in previous studies.

On the other hand, psycholinguistic research has shown that even if the task does not necessarily require semantic access, as is the case for lexical decision, semantic properties still can impact word processing (e.g., Balota et al., 1991, 2004; Hino & Lupker, 1996; James, 1975; Jastrzembski, 1981; Kiefer et al., 2022; Locker et al., 2003; Papadopoulos et al., 2017; Snefjella & Kuperman, 2016). The same seems to be true for sensorimotor semantics. To illustrate, the more a word is related to perception and motor actions, the faster the lexical decision for this word is (see Lynott et al., 2020; see also Connell & Lynott, 2014; Kuhnke et al., 2020). Similarly, faster reaction times are associated with higher levels of graspability in lexical decision (Heard et al., 2019). Performing lexical decision for sound- versus action-related words leads to activation of corresponding auditory versus motor brain areas (see Kiefer et al., 2012, 2008). Harpaintner et al. (2020) found activity in frontal and parietal motor areas when lexical decision was performed for motor-related abstract concepts and higher activity in temporo-occipital visual areas for lexical decisions on vision-related abstract concepts (see also Ulrich et al., 2023). Moreover, transcranial magnetic stimulation applied over arm versus leg brain motor areas facilitated lexical decisions to arm- versus leg-related words (e.g., "pick" vs. "kick"; Pulvermüller et al., 2005). Thus, sensorimotor information might be activated even in shallow tasks, such as the lexical decision task, and it should then be possible to find the MSE in such tasks.

Experimental Task: The Role of Context

Another task-related parameter is the amount of contextual information given to participants. In previous successful replications of the MSE, target words were always presented in context, for example, "blender *can be* loud." Richer linguistic contexts (e.g., a sentence or a narrative compared to a word shown in isolation) might enhance embodied effects in language processing (see Glenberg & Kaschak, 2002; Zwaan, 2004; see also Meteyard et al., 2012). Indeed, space-related cue words (e.g., the word "hat" is an upper-space cue) presented immediately after context words (e.g., "cowboy") seem to cause stronger activation of spatial representations than the same cue words presented without context words (see Estes et al., 2008, Experiment 1 vs. Experiment 3). Similar mechanisms might be at play for perceptual semantics. Note that all previous studies of MSE included syntactic structures ("concept *can be* property") or presented properties in the context of corresponding concepts ("concept + property"). In other words, all previous MSE studies measured semantic processing in a linguistic context (see Bernabeu et al., 2017; Collins et al., 2011; Connell & Lynott, 2011; Hald et al., 2011; Lynott & Connell, 2009; Pecher et al., 2003, 2004; Scerrati et al., 2015, 2017).

Like for deep semantic tasks discussed above, the use of linguistic context in the MSE research also questions to what extent sensorimotor activation found in those studies is part of word semantics itself or is merely preactivated by the surrounding words, that is, by the context.

To conclude, the MSE would provide stronger evidence in favor of embodied theories of semantics if it would be found (a) in a task that does not require explicit semantic processing, such as the lexical decision task (see the Experimental Task: Depth of Semantic Processing section) and (b) for stimuli presented without a context, that is, single words presented in isolation. In this case, the MSE would be considered functionally relevant for language processing (see Pulvermüller, 2012, for a theoretical account). However, no evidence of this kind has been found in behavioral studies so far.

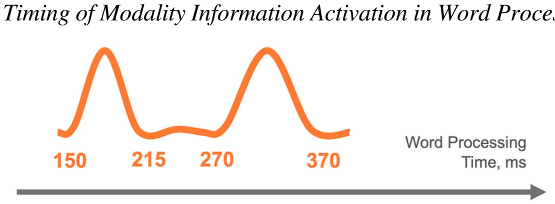
Timing of Sensorimotor Semantic Activation

If modality information is activated even in the lexical decision task, as discussed above, the question remains as to why no MSE has been found in previous research using this task (Scerrati et al., 2017). We suggest that the timing of related cognitive processes is the crucial component that should be considered to answer this question.

Activation of sensorimotor semantics occurs relatively early after stimulus presentation. Visually presented words related to auditory concepts (e.g., "loud") triggered activation of the auditory neural mechanisms as early as *150 ms* (Kiefer et al., 2008), with similar findings from Harpaintner et al. (2022) who observed differences in neural activity at 178 ms during lexical decisions involving motor- versus vision-related abstract words. Overall, activation of sensorimotor information was found in two time windows schematically illustrated in Figure 1—around 150–215 ms and around 270–370 ms across various studies involving lexical decision, though the exact timing slightly varied (Bernabeu et al., 2017; Hald et al., 2011; see also Louwerse & Hutchinson, 2012).

Notably, these findings from the literature on sensorimotor semantic processing broadly align with research on *semantic processing*, more specifically, with the notion of an early initial semantic access, occurring around 160–250 ms after word presentation (Hauk et al., 2006, 2012). In an experimental study implying single-word processing, Hauk et al. (2012) demonstrated retrieval of semantic information around 200 ms after stimulus presentation. The classical N400 signature of semantic processing (Kutas & Federmeier, 2000, 2011) might also overlap with the second MSE time window (270–370 ms). While this second window may not fully align with the N400 component, it likely reflects an earlier portion of it. Thus, both early (around 150–215 ms after stimulus onset) and later (around 270–370 ms) time windows

Figure 1
The Timing of Modality Information Activation in Word Processing



Note. The horizontal line represents time from stimulus onset (zero, not shown in this figure). The orange line schematically represents activation of semantic information. See the online article for the color version of this figure.

identified in the MSE research broadly correspond to the findings from the semantic processing literature (see also Kiefer et al., 2022, for similar findings in abstract concepts).

While previous research failed to find the MSE in lexical decision (Scerrati et al., 2017), it might be that the time interval between the stimuli in consecutive trials was simply too long for the modality information from prime stimulus to influence the processing of the following target stimulus. We suggest that presenting the stimuli without any interval might induce the MSE in the lexical decision task. In order to test this suggestion, in our study, we completely eliminated the time interval by presenting two words simultaneously.

Variability of the MSE Across Semantic Modalities

Importantly, in most previous studies, including the MSE lexical decision study (Scerrati et al., 2017), responses to target stimuli of varying modalities (i.e., visual, auditory, olfactory, etc.) were typically aggregated and thus reduced to congruent versus incongruent conditions, that is, without analyzing the MSE within each modality separately. However, rating studies demonstrate that some modalities in language correlate with each other stronger than others, and even the direction of these correlations can differ. For example, when data for large samples of words were analyzed, a positive correlation between visual and haptic modalities was found (see Lynott et al., 2020, Figure 5; Miklashevsky, 2018, Table 3). At the same time, a negative correlation between haptic and auditory modalities was shown (Lynott et al., 2020, Figure 5). In other words, at the level of language in general, if a given concept is related to vision, it is also likely to be related to touch. At the same time, if a given concept is related to touch, it is likely not to be related to hearing.

These asymmetric relationships between modalities might result from the topography of the corresponding brain regions. Different sensorimotor brain areas are activated depending on the semantic content of a word (e.g., visual, auditory, or tactile), and these areas may vary in their localization and levels of connectivity, resulting in a different association strength between semantic modalities. Indeed, processing visual semantics activates the primary visual cortex, while processing tactile semantics activates the parietal, temporal, and frontal cortices (Huth et al., 2016). Processing words with auditory semantics leads to the activation of posterior superior and middle temporal gyri, that is, the regions also activated in sound perception (Kiefer et al., 2008). Distinct sensorimotor semantic attributes, such as color, shape, visual motion, sound, or manipulation, are also processed by distinct brain areas (Fernandino et al., 2016).

To summarize, the rating studies and neuroscientific research suggest that the MSE might differ across modalities, and thus, the data from different semantic modalities should not be aggregated whenever possible.

The Present Study

The research question we aimed to answer is whether perceptual semantics is activated during shallow lexical processing; more specifically, we aimed to test whether it is possible to detect the behavioral signatures of the MSE in the lexical decision task. The present study is motivated by the following two considerations: (a) Activation of sensorimotor information during a task that does not require any context or explicit semantic processing, such as the lexical decision task, would provide stronger evidence in favor of embodied semantics compared to the previously used property verification task and (b) the timing of processing might be crucial to find the MSE in the lexical decision task. Specifically, in contrast to previous work, we will test simultaneous word presentation.

Furthermore, we suggest that the data from different semantic modalities should be analyzed separately and not aggregated by switch versus no-switch conditions whenever possible: There might be systematic associations between modalities in language, as correlational analysis demonstrates (see the Variability of the MSE Across Semantic Modalities section). For example, the switching cost between haptic and visual modalities (which often correlate) might be smaller than between visual and auditory modalities (which often do not correlate). On the other hand, an opposite prediction could be made: Haptic information might be closely related to motor features, and therefore, there might be a larger switch cost between haptic and visual modalities (cf. differences in brain activity found between motor and visual features, Harpaintner et al., 2020, 2022; see also strikingly different patterns of brain activity during processing motor-related tool names vs. vision-related animal names, Pulvermüller, 2001).

In the present study, we designed two experiments addressing the abovementioned points. In Experiment 1, we explored the MSE in the lexical decision task with no context. In this experiment, we simultaneously presented native speakers of Russian with two words of the same (e.g., "sparkling" + "white") versus different modalities (e.g., "loud" + "white"). We asked participants to perform a lexical decision for both words at once. In Experiment 2, we replicated the findings of Experiment 1 in German with German native speakers and an entirely new set of stimuli. The results of both our experiments suggest that (a) modality information can be activated even during shallow lexical processing, such as in the lexical decision task, when two stimuli are presented at once and (b) the exact combination of semantic modalities modulates the MSE.

Method

Experiment 1: Simultaneous Presentation (in Russian)

In this experiment, we explored the MSE in the lexical decision task with no context. Native speakers of Russian were simultaneously presented with two words of the same (e.g., "sparkling + white") versus different modalities (e.g., "loud + white"). The task was to perform a lexical decision for both words at once. This and the following two experiments were not preregistered.

Stimuli Selection

There are two main challenges in selecting stimuli materials for experiments on semantic modalities. First, different modalities are asymmetrically represented in language: An overwhelming number of words belong to the visual semantic modality as compared to olfactory or gustatory modalities (e.g., see Figure 3 in Lynott et al., 2020; or Table 2 in Miklashevsky, 2018). Second, semantic modalities correlate with each other because the majority of concepts can be perceived through several perceptual modalities (i.e., "tasty" is primarily related to taste, but also to smell or even vision in some contexts, such as "The food looks and smells tasty."). The notion of *modality exclusivity*, or the measure of how much a concept relates to one versus several modalities, was introduced to measure the unimodality of a particular property (Lynott & Connell, 2009).

For our study, we selected adjectives because Lynott and Connell (2013) found that adjectives are more distinct across modalities and exhibit higher modality exclusivity than nouns. We excluded the gustatory and olfactory modalities from our experiment: These two modalities usually highly correlated with each other as previous research has demonstrated (r coefficients vary from .66 to .81 in comparison to haptic and visual correlation, for which r coefficients vary from .38 to .54; see Lynott & Connell, 2009, 2013; Miklashevsky, 2018; see also Speed & Majid, 2020, for a review).

We used a psycholinguistic dictionary, *Sense Organs, Emotions and Adjectives of the Russian Language* (Kolbeneva & Alexandrov, 2010), to select stimuli in Russian. In this dictionary, 7,616 Russian adjectives were evaluated by native speakers as either related or unrelated to each of the five perceptual modalities (vision, audition, touch, olfaction, and taste). In the resulting database, statistics of "yes" versus "no" responses are provided for each word for each modality. For example, for the word "пластмассовый" ("plastic"), the ratio "11/1" is provided for the tactile modality, which means that out of 12 participants, 11 replied with "yes" to the statement "I assess this characteristic by touch" and one participant responded with "no."

First, for all adjectives, we calculated a modality score as a proportion of "yes" responses in the overall number of responses. In the example with "plastic," its association with the tactile modality would be 11 (yes responses)/12 (all responses) = .92. This modality score could range from 0 (*no association with that modality*) to 1 (*the strongest possible association with that modality*).

Next, we selected 24 adjectives from the visual, haptic, and auditory modalities with high modality exclusivity. In other words, we kept the link to the target modality (e.g., auditory) as high as possible while choosing the words from the dictionary so that the remaining two modalities have the lowest possible scores. The stimuli of each modality group were significantly different from the other two groups in the corresponding semantic modality (all p values <.001).

We also controlled for stimuli length (in letters) and frequency, trying to reduce the differences between the three groups (visual, audial, and haptic) as much as possible. Frequency was obtained from the frequency dictionary by Lyashevskaya and Sharov (2009) and log-transformed. Differences in length were not significant, $F(2, 69) = 0.672$, $p = .514$. In frequency, however, there was a significant difference, $F(2,69) = 4.406$, $p = .016$: The auditory modality had a lower frequency than the other two modalities, as Fisher's least significant difference post hoc test demonstrated (p = .021 for the comparison of visual and auditory modalities; p = .008 for the comparison of haptic and auditory modalities). Despite our attempts, it was not possible to completely eliminate these differences while keeping substantial samples of stimuli in each modality due to a limited number of stimuli with high modality exclusivity. Descriptive statistics for the stimuli in Experiment 1 are presented in Table 1.

Using a customized R script (see the Transparency and Openness section), we randomly selected stimuli pairs from the pools of stimuli for each modality. We did not control for whether each word appeared in the first versus second position (see below); instead, only word modality was relevant at this stage. Random pairs without a Modality Switch (e.g., "sparkling + white") versus with a switch (e.g., "loud + white") were formed. Every stimulus appeared in both positions, that is, a given stimulus appeared on the top and on the bottom of the fixation cross. From these stimuli, we created four lists with 360 pseudorandom stimuli pairs, where half included one word and one pseudoword and the other half consisted of two real words. No trials included two pseudowords. Pseudowords were created by rearranging syllables of our stimuli of interest. Thus, pseudowords matched the word stimuli in their length distribution and followed natural phonetic regularities of the Russian language. A typical ending "-ий/-ый" ("-ij/-yj"), a formal grammatical marker of adjectives in Russian, was used for all pseudowords. Thus, participants could not distinguish between words and pseudowords based on the mere processing of letter combinations. Instead, using pseudowords (and not random letter sequences) nudged participants to read the entire stimulus. All words and pseudowords were presented in their initial form (singular, masculine, nominative case for Russian).

To further control for possible shallow lexical associations between the stimuli, we extracted data about collocation frequency for each word pair from the Russian National Corpus (https://ruscorpora.ru, Plungian et al., 2005; includes 374 million words). In this corpus, it is only possible to search for co-occurrences of two words within a certain window (number of words). We selected the maximum possible window of 10 words on the left and 10 words on the right from a target word. That is, we searched for collocations within a window of 20 words, regardless of sentence borders.

The Russian National Corpus provides several collocation measures, the most accurate of which is the LogDice (Rychlý, 2008). This measure is calculated based on frequencies of each of the two words and the frequency of the co-occurrence of the two

Table 1
Descriptive Statistics of Stimuli in Experiment 1 (Russian Adjectives)

Semantic modality	Frequency	Length	Modality rating: Visual	Auditory	Haptic
Visual	1.35 (0.49)	7.04 (1.37)	0.92 (0.06)	0.04 (0.08)	0.04 (0.06)
Auditory	0.975 (0.51)	7.5 (1.53)	0.24 (0.19)	0.80 (0.15)	0.07 (0.14)
Haptic	1.41 (0.64)	7.17 (1.34)	0.49 (0.26)	0.12 (0.12)	0.79 (0.17)

Note. Mean values are provided in each cell, and corresponding standard deviation values are provided in parentheses.

words in a corpus; therefore, it is independent of the total size of a corpus. Usually, the LogDice value cannot exceed 14, meaning that all occurrences of Word A also co-occur with Word B and vice versa. The LogDice can also be negative; in that case, it means that there is less than one co-occurrence of Words A and B per 16,000 occurrences of A or B. The negative LogDice can also be interpreted as an absence of a statistically significant co-occurrence of two words (Rychlý, 2008).

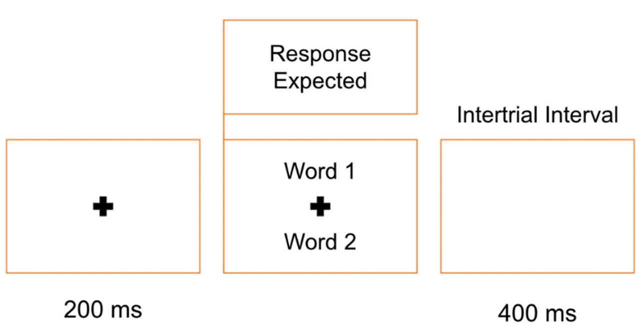
We found at least one co-occurrence for 32% of unique word pairs; the remaining word pairs never appeared together in the window of 10 + 10 words in the Russian National Corpus. Among those pairs we found in the corpus, the mean LogDice was equal to 7.45 (*SD* = 1.00). The minimum LogDice equaled to 5.42. We assigned the value of LogDice = 0 to those pairs that do not appear in the corpus. This new variable was used as a covariate in the main analysis below.

Design and Procedure

This and the following experiment were conducted online using the Gorilla Experiment Builder (Anwyl-Irvine et al., 2020). It was only possible to participate in the experiments using a PC or a laptop. A physical keyboard had to be used, and participants were instructed to keep their index fingers on response keys (P and Q, right and left index fingers correspondingly) during the entire experiment. In both experiments, participants were asked to keep a distance of around 50 cm from the screen. At the beginning of each experiment, we calibrated the size of the stimuli on the screen to keep it constant across participants. We asked our participants to minimize potential disturbances, to close all other browser tabs except the one with the experiment, and to maximize the browser window so that it occupies the entire screen. We did not strictly control the exact browser and the operation system used by the participants. The participants in Experiment 1 used the following browsers: *Google Chrome*, *Edge*, *Firefox*, *Opera*, *Safari*, or *Yandex*. Participants of Experiment 1 used the following operation systems: *macOS*, *Linux*, *Ubuntu*, or *Windows*.

At the beginning of the experiment, a participant was randomly assigned by Gorilla software to one of the four stimuli lists (see the Stimuli Selection section for details); the order of stimuli presentation within the list was also randomized for each participant. The experimental procedure of Experiment 1 is shown in Figure 2.

Figure 2
Experimental Procedure of Experiment 1



Note. See the online article for the color version of this figure.

Each trial started with a fixation cross appearing in the center of the screen. The cross ensured that participants always started visually inspecting the screen from the central point. After 200 ms, two stimuli appeared simultaneously: one on the top (Word 1) and one on the bottom of the cross (Word 2). These stimuli could be either two words (50% of all trials, i.e., 180 trials) or one word and one pseudoword (50% of all trials). A pseudoword appeared randomly either on the top (90 trials or 50% of trials with pseudowords) or on the bottom of the cross. Participants were asked to press the Q key if both stimuli were words and to press the P key if at least one of the stimuli was a pseudoword; keys were counterbalanced between participants. There was no time limit. After the response, a blank screen (intertrial interval) appeared for 400 ms, and the procedure was repeated. Both speed and accuracy were emphasized in the instruction.

We used a 3 (Word 1: Auditory/Haptic/Visual) × 3 (Word 2: Auditory/Haptic/Visual) design. Each possible combination of semantic modalities was presented 15 times. However, due to a programming mistake, combinations “auditory + visual,” “haptic + visual,” and “visual + visual” were repeated 30 times each. A practice including 10 trials preceded the experiment. The practice stimuli differed from the actual stimuli used in the experiment: They did not directly represent sensorimotor modalities but conveyed more abstract properties such as “kind” or “beautiful.” At the end of the experiment, participant-related data were collected: age, gender, and Handedness (using an abbreviated version of the Edinburgh Handedness Inventory, see Veale, 2014). The Handedness score could range from −100 (*purely left-handed*) to +100 (*purely right-handed*). The entire duration of the experiment was 30 min.

Sample Size Calculation

In this and the following experiment, we used unimodal stimuli selected based on a pretest, as suggested by Lynott and Connell (2009). Connell and Lynott argued that this procedure ensures a stronger effect size than in the original study by Pecher et al. (2003), in which no pretest was used to select stimuli. Therefore, since we selected stimuli in line with the study by Lynott and Connell (2009), we also calculated the effect size from their study, which resulted in Cohen’s *f* = .161. We used this value for further calculations in G*Power (Faul et al., 2007) and arrived at the minimum required sample size of *N* = 34 participants (α = .05 and power = .80). Because mixed-effects modeling used in the present study is even more powerful than separate analyses by participants and by items (Brysbaert & Stevens, 2018) often used in previous studies, using mixed-effects models further increased our ability to detect effects of interest. We included an additional 10% for possible dropouts and aimed to collect at least 37 participants in each experiment. A detailed protocol of power analysis is available in additional online material (see the Transparency and Openness section; https://osf.io/a85h7/).

Participants

A total number of 73 Russian native speakers participated in an online study. Five participants were excluded due to formal reasons (native language other than Russian, accuracy lower than 85%, average reaction time [RT] longer than 2,500 ms, or missing questionnaires). Data from 68 Russian native speakers (18 male, 47 female, three nonbinary; M_{age} = 27 years, SD = 11 years) were included in the analysis. Participants' mean Handedness score was 84 (SD = 40; range from −100 to +100), indicating a prevailing right-handedness of the sample. All participants submitted their informed consent at the beginning of the experiment by clicking on the corresponding checkbox and participated voluntarily with no compensation. Our research paradigm for the entire study was approved by the Ethics Committee of Tomsk State University (the protocol of the meeting of the ethical committee on May 18, 2021). The study was designed and conducted following the guidelines laid down in the Declaration of Helsinki.

Analysis and Results

Data preparation and analyses for this and the following experiment were performed using *Microsoft Excel* for *Microsoft 365* (Version 2401) and R (R Core Team, 2020, Version 4.2.3). We did not analyze trials with pseudowords (N = 12,240). From now on, the remaining 12,240 trials are referred to as 100%. The mean accuracy across participants was 95%, ranging from means of 87% to 100% across participants. Mean RT across participants was 1,234 ms (ranging from means of 745 to 2,045 ms across participants, SD = 1,166).

We discarded trials with incorrect responses (N = 441; 3.6% of the data), trials with RTs shorter than 400 ms or longer than 4,000 ms (indicating anticipation or procrastination, N = 91, 0.7%), and trials with RTs outside of ±3 *SDs* from individual means (N = 222; 1.8%). The remaining 11,486 trials (93.8%; mean RT = 1,159 ms) were submitted for further statistical analysis.

We performed mixed-effects linear regression using the *lme4* package (Bates et al., 2015, Version 1.1.32) in R. The initial model included Word 1 (semantic modality of Word 1: auditory/haptic/visual; converted into a factor with the auditory modality automatically taken as a baseline), Word 2 (semantic modality of Word 1: auditory/haptic/visual; converted into a factor with the auditory modality automatically taken as a baseline), and an interaction between the two factors. We also controlled for the following covariate terms: Length of Word 1, Length of Word 2, Frequency of Word 1, Frequency of Word 2, LogDice (co-occurrence measure for each word pair, see the Stimuli Selection section), Trial Number, participant's Handedness (ranging from −100 to +100), Response Mapping (Q for words and P for pseudowords/P for words and Q for pseudowords), and the interaction between Handedness and Response Mapping.

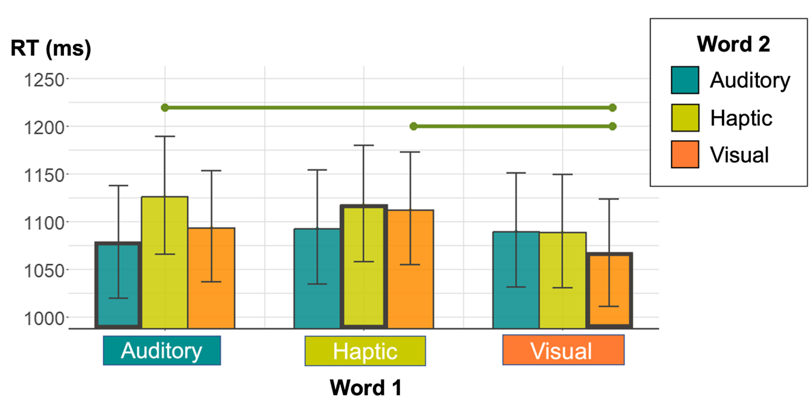
All continuous predictors except Handedness were mean-centered. The binary variable Response Mapping was assigned sum-coded contrasts (−0.5 and 0.5; Brehm & Alday, 2022; Schad et al., 2020). The dependent variable RT exhibited a high positive skewness of 1.853 and was therefore log-transformed (Note that Figure 3 shows back-transformed RTs). Participant, Word 1, and Word 2 were included as random intercepts. The models with random slopes for participants by the modality of Word 1 and/or Word 2 did not converge.

We employed backward elimination to identify the best-fitting models using the *drop1* function from the *lme4* package. Effects that were not significant or marginally significant (p < .100) and thus did not improve model fit were successively eliminated. Results from the final model are presented in Table A1. This and all other model-output tables include the effects of primary interest, no matter their level of significance (i.e., Word 1, Word 2, and their interaction), as well as each covariate term only where these were significant (p < .050) or marginally significant (p < .100).

LogDice, Handedness, Response Mapping, and the interaction between Handedness and Response Mapping turned nonsignificant and were removed from the model. All other covariates were significant or marginally significant. Analysis of variance (ANOVA) indicated that the effects of Word 1 and Word 2 were not significant, Word 1: $F(2, 67.2) = 2.145$, $p = .125$; Word 2: $F(2, 68.2) = 1.938$, $p = .152$. Importantly, the key interaction Word 1 × Word 2 turned out

Figure 3

Mean Reaction Times in Nine Conditions (Experiment 1)



Note. Combinations of semantic modalities: Word 1 × Word 2. Bars represent predicted mean RTs, that is, RTs predicted by a linear mixed-effects model accounting for covariates (see the main text for details). Whiskers represent ±1 *SE*. Thicker inner frames indicate no-switch conditions. Green horizontal lines represent significant differences between conditions (p values <.050). RTs = reaction times; *SE* = standard error. See the online article for the color version of this figure.

significant, $F(4, 8778.1) = 3.031$, $p = .017$. Note that the table output (Table A1) does not provide all possible comparisons between all levels of Word 1 and Word 2. Therefore, we further examined this interaction using a Tukey post hoc test from the *emmeans* R package (Lenth et al., 2024, Version 1.8.5) and found two significant comparisons across levels of Word 1 and Word 2 (see Table A2 and Figure 3). Participants were significantly faster in the "visual + visual" condition than in the "auditory + haptic" ($p = .041$) or "haptic + visual" condition ($p = .026$). All other comparisons were not significant.

Next, we separately examined variables Word 1 and Word 2 using a Tukey post hoc test (which can also be used unprotected, i.e., regardless of the significance of the original effect; see Barnette & McLean, 2005) from the *emmeans* package (Version 1.8.5) and found no significant RT differences between the levels of Word 1 or Word 2 ($p = .106$ or larger).

To additionally examine the pattern of results, we conducted a sensitivity analysis. We aggregated the data by the combinations of Word 1 and Word 2 levels, which resulted in the new binary variable Modality Switch (switch/no switch). This analysis is more comparable to those conducted in previous studies, which did not separate modalities (see the Variability of the MSE Across Semantic Modalities section). We sum-coded the contrasts of Modality Switch (−0.5 and 0.5) and submitted this variable to a mixed-effects linear model with the same initial covariates as in the main analysis described above. As before, we followed the same backward elimination procedure. Modality Switch turned significant, with longer RTs in the switch condition as compared to the no-switch condition ($b = 0.0145$, $t = 2.7059$, $p = .007$; see Table A3 for further details). In Figure 3, one can clearly see that the no-switch conditions for the visual and the auditory modalities are indeed the fastest, even though pairwise comparisons do not reach significance in most cases.

To summarize, we found a classical MSE for one combination of modalities (the combination "visual + visual" was faster than "haptic + visual"), though not others (e.g., there was no significant difference in RTs to "haptic + haptic" as compared to "auditory + haptic" pairs). Thus, our Experiment 1 demonstrated that the MSE (i.e., activation of a semantic modality) emerges even in the shallow lexical decision task with no context—as we predicted—though not for all possible combinations of modalities. While not explicitly related to MSE, yet still in line with this effect, participants responded to no-switch pairs "visual + visual" faster than to switch pairs "auditory + haptic."

However, the interpretation of the results of Experiment 1 is limited due to two suboptimal technical solutions in the procedure of Experiment 1: a differing number of stimuli per condition and an absent timeout for responses. Even more importantly, we aimed to increase the generalizability of our results by replicating the study in another language and with new stimuli. All these factors led us to design Experiment 2 (see below).

Experiment 2: Simultaneous Presentation (in German)

Experiment 1 has demonstrated that simultaneous presentation of stimuli pairs results in the activation of semantic modality even in lexical decision. In Experiment 2, our goal was to increase the generalizability and test the reliability of these results by replicating the study in German. We also accounted for minor issues with the experimental design, procedure, and stimuli mentioned above.

Stimuli Selection

To create a stimuli set, we translated sensory-related adjectives from previous studies in English and Italian (Lynott & Connell, 2013; Lynott et al., 2020; Pecher et al., 2003; Scerrati et al., 2015) into German, as well as all Russian adjectives from our Experiment 1. A resulting set of 233 adjectives was presented for a pretest to German native speakers ($N = 50$; $M_{age} = 25$ years, $SD = 8$ years; 45 female, four male, and one nonbinary participant). All participants were university students. Participants were asked to rate the relatedness of word stimuli to each of the five modalities (vision, hearing, touch, taste, and smell) by answering the question "Inwieweit nehmen Sie diese Eigenschaften durch das SEHEN wahr?" ("To what extent do you perceive these qualities through SEEING?"). The question was modified accordingly for other modalities. Rating answers ranged from 1 (*not at all*) to 5 (*very much*). There was a possibility of indicating a word as unknown.

All words unfamiliar to at least one participant were removed (four to five words from each list, mostly overlapping across participants). Cronbach's α for each of the five modalities (vision, audition, touch, olfaction, and taste) was high (>.862). All individual responses to adjectives were aggregated, and mean values for each modality were calculated. As before, we selected adjectives from the visual, haptic, and auditory modalities with high modality exclusivity, that is, we kept the link to the target modality (e.g., auditory) as high as possible while choosing the words so that the remaining two modalities have the lowest possible scores. As a result, three sets of stimuli of each modality of interest (visual, auditory, and haptic) were created, containing 25 adjectives each. The stimuli of each modality group were significantly different from those of the other two groups in the corresponding semantic modality (all p values <.001). Note that this operationalization of semantic modality on a scale from 1 to 5 used here is different from the one available for Russian in Experiment 1, where the proportion of "yes" responses in the overall number of responses was calculated, that is, the score in Experiment 1 ranged on a scale from 0 to 1 (see the Stimuli Selection section). Nevertheless, both scales estimate the same construct.

We calculated word length in letters and obtained log10 frequency values from the SUBTLEX-DE database (Brysbaert et al., 2011). We used ANOVAs to compare the three modality groups with each other on these two parameters. There were no significant differences between modality groups in length, $F(2, 72) = 1.010$, $p = .369$, or frequency, $F(2, 72) = 1.063$, $p = .351$. Stimuli characteristics are presented in Table 2.

As before, pseudowords were created by rearranging syllables of our stimuli, and thus, pseudowords matched words in their length distribution and followed natural phonetic regularities of the German language (see also the Stimuli Selection section regarding the use of pseudowords). All words and pseudowords were presented in their initial form (i.e., without endings for the German language). As in Experiment 1, using a customized R script (see the Transparency and Openness section), we randomly selected stimuli pairs from the pools of stimuli for each modality. We did not control whether each word appeared in the first versus second position. Instead, only word modality was relevant at this stage. We formed

Table 2
Descriptive Statistics of Stimuli in Experiment 2 (German Adjectives)

			Modality rating		
Semantic modality	Frequency	Length	Visual	Auditory	Haptic
Auditory	1.29 (0.93)	7.60 (2.06)	1.46 (0.25)	4.72 (0.23)	1.14 (0.19)
Haptic	1.72 (1.00)	6.76 (1.81)	3.22 (0.35)	1.50 (0.22)	4.55 (0.25)
Visual	1.54 (1.19)	7.08 (2.41)	4.79 (0.12)	1.01 (0.04)	1.04 (0.10)

Note. Mean values are provided in each cell, and corresponding standard deviation values are provided in parentheses.

word pairs without a Modality Switch (e.g., "sparkling + white"), pairs with a Modality Switch (e.g., "loud + white"), and pairs including pseudowords. Every stimulus appeared in both positions, that is, a given stimulus appeared on the top and bottom of the fixation cross. From all stimuli (i.e., words and pseudowords), we created four variants of lists, each with 270 pseudorandom stimuli pairs. An equal number of trials (1) consisted of two words versus (2) included at least one pseudoword. In trials including pseudowords, either one or both stimuli could be pseudowords.

To further control for possible shallow lexical associations between the stimuli, we extracted data about collocation frequency for each word pair from the COSMAS II Corpus (cosmas2.ids-mannheim.de/cosmas2-web, Bodmer, 2005; includes 5.5 billion words). In this corpus, it is only possible to search for co-occurrences of two words within a certain window (number of sentences; cf. stimuli selection in Experiment 1, where the number of words defined such windows). We selected the window of one sentence on the left and one sentence on the right from a target word. That is, we searched for collocations within a window of three sentences (14.87 words per sentence on average in the W-öffentlich subcorpus used for our study), regardless of sentence length.

To keep the analysis as similar to Experiment 1 as possible, we again focused on the LogDice measure (see the Stimuli Selection section for details). First, we extracted the absolute frequencies of each word from the COSMAS II Corpus. Second, we extracted the absolute frequencies of each co-occurrence. Last, we calculated LogDice values using the formula suggested by Rychlý (2008).

We found at least one co-occurrence for 56% of unique word pairs used in the experiment; the remaining word pairs never appeared together in the window of three sentences in the COSMAS II Corpus. Among those pairs we found in the corpus, the mean LogDice was equal to −3.31 ($SD = 2.00$). The minimum LogDice equaled to −7.44. We assigned the value of LogDice = −8 to all those pairs that do not appear in the corpus so that their LogDice value remained under the minimal value of pairs present in the corpus. This new variable was used as a covariate in the main analysis below.

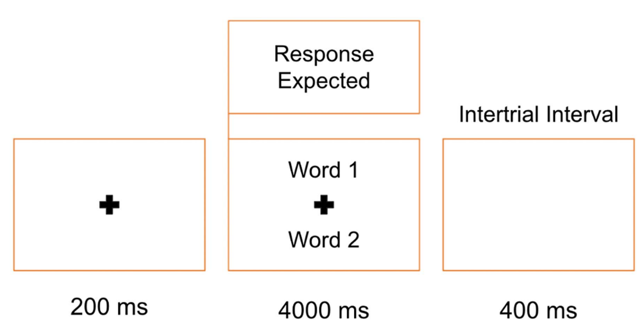
Design and Procedure

The design of Experiment 2 was identical to the design of Experiment 1: We used a 3 (Word 1: Auditory/Haptic/Visual) × 3 (Word 2: Auditory/Haptic/Visual) design. Each possible combination of semantic modalities was presented 15 times. Like in Experiment 1, participants discriminated words from nonwords (lexical decision task) for two stimuli presented simultaneously, one on the top and one on the bottom of a fixation cross. Participants were asked to press the Q key if both stimuli were words and to press the P key if at least one of the stimuli was a nonword; the keys were counterbalanced across participants. In half of all trials ($N = 270$), both stimuli were words, while in the other half, at least one was a pseudoword.

At the beginning of the experiment, a participant was randomly assigned by Gorilla software to one of the four stimuli lists (see the Stimuli Selection section for details); the order of stimuli presentation within the list was also randomized for each participant. The procedure looked as follows (Figure 4): The cross appeared on the screen, dragging participants' attention to the center. After 200 ms, two stimuli appeared on the top and bottom of the cross. The lexical decision was expected for 4,000 ms (timeout). The intertrial interval was set for 400 ms, and the procedure was repeated again. Both speed and accuracy were emphasized in the instruction.

A practice including 10 trials preceded the experiment. The practice consisted of the adjectives taken from the actual list of stimuli. However, the stimulus pairs introduced during practice were novel and were not repeated during the actual experiment. At the end of the experiment, participant-related data were collected: age, gender, and Handedness (using an abbreviated version of the Edinburgh Handedness Inventory, see Veale, 2014). The Handedness score could range from −100 (*purely left-handed*) to +100 (*purely right-handed*). The entire duration of the experiment was 30 min. All technical details were identical to those of Experiment 1 (see the Design and Procedure section). Again, we did not strictly control the exact browser and the operation system used by the participants. The participants in Experiment 2 used the following browsers: *Google Chrome*, *Edge*, *Firefox*, *Mozilla*, *Opera*, or *Safari*. The participants of Experiment 2 used the following operation systems: *macOS*, *Ubuntu*, or *Windows*.

Figure 4
Experimental Procedure for Experiment 2



Note. See the online article for the color version of this figure.

Participants

A total number of 62 participants were recruited via the university platform SONA of the University of Potsdam (https://uni-potsdam.sona-systems.com/) for an online experiment. All participants gave their informed consent before the experiment. We used the same exclusion criteria as in Experiment 1: One person was a nonnative speaker of German and thus was excluded. After preprocessing, we discarded the data of three participants whose mean accuracy was less than 85%. No participants had mean RT > 2,500 ms or missing questionnaires. The final set of participants included 58 people (47 female, 10 male, and one nonbinary participant; $M_{\text{age}} = 24$ years, $SD = 8$). Participants' mean Handedness score was 75 ($SD = 55$; range from −100 to +100), indicating a prevailing right-handedness of the sample.

Analysis and Results

The remaining 58 participants had a mean accuracy of 94% (ranging from 85% to 99%, $SD = 3\%$). Mean RT across participants was 1,161 ms (ranging from 715 to 1,816 ms, $SD = 224$ ms).

We did not analyze trials with pseudowords ($N = 7{,}830$). From now on, the remaining 7,830 trials are referred to as 100%. We removed trials with false responses ($N = 401$; 5.1% of the data) and timed out trials exceeding 4,000 ms ($N = 21$; 0.3% of the data). There were no trials with RTs shorter than 400 ms (distributed from 453 to 3,999 ms). We removed trials with RTs outside of ±3 *SD*s from individual mean ($N = 146$; 1.9%). The remaining 7,396 trials (94.5%; mean RT = 1,125 ms) were submitted for further analysis.

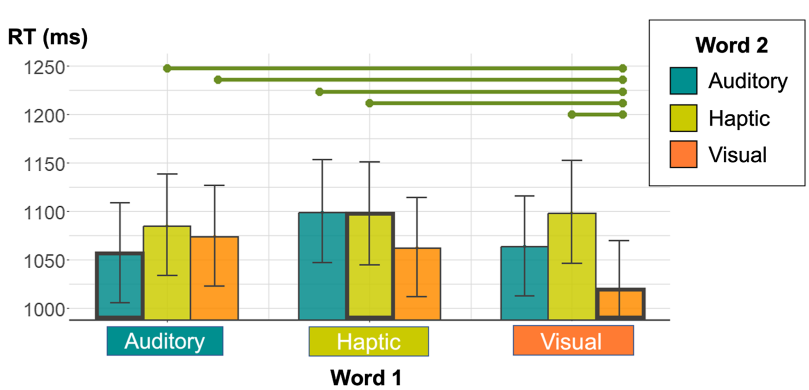
As in Experiment 1, we performed mixed-effects linear regression. The same approach was used: The initial model included Word 1 (semantic modality of Word 1: auditory/haptic/visual; converted into a factor with the auditory modality automatically taken as a baseline), Word 2 (semantic modality of Word 1: auditory/haptic/visual; converted into a factor with the auditory modality automatically taken as a baseline), and an interaction between the two factors. We also controlled for the following covariate terms: Length of Word 1, Length of Word 2, Frequency of Word 1, Frequency of Word 2, LogDice (co-occurrence measure for each word pair, see the Stimuli Selection section), Trial Number, participant's Handedness (ranging from −100 to +100), Response Mapping (Q for words and P for pseudowords/P for words and Q for pseudowords), and the interaction between Handedness and Response Mapping.

All continuous predictors except Handedness were mean-centered. The binary variable Response Mapping was assigned sum-coded contrasts (−0.5 and 0.5). The dependent variable RT exhibited a high positive skewness of 1.825 and was therefore log-transformed (Note that our figures show back-transformed RTs). Participant, Word 1, and Word 2 were included as random intercepts. The models with random slopes for participants by the modality of Word 1 and/or Word 2 did not converge.

We employed backward elimination to identify the best-fitting models using the *drop1* function from the *lme4* R package. Effects that were not significant or marginally significant ($p < .100$) and thus did not improve model fit were successively eliminated. Results from the final model are presented in Table A4.

Response Mapping and its interaction with Handedness turned nonsignificant; these terms were removed from the model. All other covariates were significant or marginally significant. ANOVA indicated that the effect of Word 1 was only marginally significant, $F(2, 68.4) = 2.778$, $p = .069$, while the effect of Word 2 was significant, $F(2, 71.1) = 8.254$, $p = .001$. Importantly, the key interaction Word 1 × Word 2 was significant, $F(4, 3660.6) = 4.624$, $p = .001$. We further examined this interaction using a Tukey post hoc test and found several significant comparisons across levels of Word 1 and Word 2 (see Table A5 and Figure 5). Participants were significantly faster in the "visual + visual" condition than in the "haptic + auditory" ($p < .001$), "auditory + haptic" ($p = .004$), "haptic + haptic" ($p < .001$), "visual + haptic" ($p < .001$), or

Figure 5
Mean Reaction Times in Nine Conditions (Experiment 2)



Note. Combinations of semantic modalities: Word 1 × Word 2. Bars represent predicted mean RTs, that is, RTs predicted by a linear mixed-effects model accounting for covariates (see the main text for details). Whiskers represent ±1 *SE*. Thicker inner frames indicate no-switch conditions. Green horizontal lines represent significant differences between conditions (all p values <.050). RTs = reaction times; *SE* = standard error. See the online article for the color version of this figure.

"auditory + visual" (p = .015) conditions. All other comparisons were not significant (p > .050).

Similar to Experiment 1, the "visual + visual" condition was the fastest. Indeed, this no-switch condition was faster than the switch condition "auditory + visual," thus again demonstrating a *classical MSE effect. Another example of the MSE* effect is the faster RTs in the "visual + visual" than in the "visual + haptic" condition.

Surprisingly, we found significantly slower RTs in the no-switch condition "haptic + haptic" as compared to the no-switch condition "visual + visual." There seems to be a general disadvantage in processing haptic semantics across experiments (cf. numerically longer RTs even in the no-switch condition "haptic + haptic" as compared to the other two no-switch conditions—i.e., "visual + visual" and "auditory + auditory"—in Experiments 1 and 2). Indeed, our following analysis confirmed this suggestion: As in Experiment 1, we separately examined variables Word 1 and Word 2 using a Tukey post hoc test and found a significant RT difference between the levels of Word 1: Participants were slower in responding to haptic than to visual words (p = .049). There was also a significant difference across the levels of Word 2, with the haptic modality resulting in slower responses than the visual modality (p < .001). Other comparisons across levels of Word 1 or Word 2 were not significant (p = .092 or larger).

To additionally examine the pattern of results, we conducted a sensitivity analysis. We aggregated the data by the combinations of Word 1 and Word 2 levels, which resulted in the new binary variable Modality Switch (switch/no switch). This analysis is more comparable to those conducted in previous studies, which did not separate modalities (see the Variability of the MSE Across Semantic Modalities section). We sum-coded the contrasts of Modality Switch (−0.5 and 0.5) and submitted this variable to a mixed-effects linear model with the same initial covariates as in the main analysis described above. As before, we followed the same backward elimination procedure. Modality Switch turned significant, with longer RTs in the switch condition as compared to the no-switch condition (b = 0.0163, t = 2.6400, p = .008; see Table A6 for further details). In Figure 5, one can clearly see that the two no-switch conditions for the visual and the auditory modality are indeed the fastest, even though pairwise comparisons do not reach significance in many cases.

Overall, Experiments 1 and 2 showed that processing simultaneously presented pairs of adjectives in Russian and German activates modality-specific information. The results from Experiments 1 and 2 consistently demonstrate the MSE across two languages (Russian and German) in the lexical decision task, as we predicted.

The original version of this article also included Experiment 3, in which we presented two adjectives of the same or different modalities sequentially, with a varying interval between the adjectives (170 ms vs. 270 ms). However, the results of this experiment were not significant and did not allow us to derive clear conclusions. Therefore, we decided to exclude it from the main text. Interested readers can find the description of this experiment and result tables in Appendix B.

Transparency and Openness

Data sets generated for this study, stimuli materials, the protocol of power analysis, and data processing R scripts can be found on the Open Science Framework (Platonova & Miklashevsky, 2022) at https://osf.io/a85h7/.

General Discussion

Summary of Results and Interpretation

The present study aimed to demonstrate the MSE in shallow lexical processing, namely, during a lexical decision task without context. We argue that such a demonstration provides stronger evidence in favor of embodied semantics. To achieve this goal, we conducted two experiments using a novel paradigm—lexical decision for two words at once.

In Experiment 1, we simultaneously presented native speakers of Russian with two words of the same (e.g., "sparkling + white") versus different modalities (e.g., "loud + white") and asked participants to perform lexical decisions for both words at once. In Experiment 2, we applied an identical design and paradigm and replicated the findings of Experiment 1 with German native speakers and a new set of stimuli.

We applied linear mixed-effects models and included word frequency, length, Trial Number, participants' Handedness, Response Mapping, and the interaction between participants' Handedness and Response Mapping as covariates. We also included participants and items as random terms, thus accounting for interindividual variability of reaction times across participants and words. Most importantly, we controlled for word co-occurrences, thus excluding the possibility that our effects emerge due to shallow lexical associations between two words rather than due to the semantic factor of modality.

As predicted, we found classical MSEs in Experiments 1 and 2. For example, the no-switch condition "visual + visual" was faster than the switch condition "haptic + visual" in Experiment 1, and it was faster than switch conditions "auditory + visual" or "visual + haptic" in Experiment 2. Once we aggregated data by Modality Switch (thus, turning all possible modality combinations into a binary variable, as it was mostly the case in previous studies), we observed a significant effect of Modality Switch in the predicted direction. This latter result provides additional validation to our findings.

However, our approach of analyzing semantic modalities separately led to important insights that would have been not possible otherwise. It seems that the visual modality, once all other factors are accounted for, has a clear processing advantage: In Experiment 2, the "visual + visual" condition was significantly faster than five other conditions, including the other no-switch condition "haptic + haptic." Furthermore, in Experiment 2, responses were faster when Word 1 or Word 2 belonged to the visual modality as compared to the haptic modality. This visual advantage could be compared to the well-known imageability effect: Words with higher imageability are generally processed faster (Connell & Lynott, 2012; Cortese & Schock, 2013; Schwanenflugel et al., 1988), and imageability and vision relatedness usually positively correlate (Connell & Lynott, 2012; Miklashevsky, 2018). However, we would like to underline that imageability and vision relatedness are not entirely identical: Connell and Lynott (2012) demonstrated that the two variables have unique contributions to RT distribution. Therefore, even though imageability might be partially related to the effect of visual modality, it is not necessarily identical to it. Future research employing our novel paradigm should explicitly address this issue, for example, by collecting imageability ratings and adding them as a covariate to the initial model. In addition to imageability, other semantic factors, such as valence, could introduce potential confounds and should also be considered in future research.

There also seems to be a general disadvantage in processing haptic-related words. Thus, participants in Experiment 2 were slower when Word 1 or Word 2 was haptic-related than vision-related. A striking result in Experiment 2 also supports this idea: The no-switch condition “haptic + haptic” was significantly slower than another no-switch condition, “visual + visual.” Numerically, the “haptic + haptic” condition was the slowest among the three no-switch conditions (see Figures 3 and 5). These observations align with the study by Connell and Lynott (2010) who also found a tactile disadvantage in semantic processing. Connell and Lynott argued that this processing disadvantage emerges due to the evolutionary adaptation of endogenous attention to incoming sensory stimuli. Note that we did not test motor relatedness in the present study; therefore, we cannot explore the exact relationship between haptic and motor information in our sample stimuli and tease apart the effects of these two types of information.

Overall, the results of both our experiments demonstrated that perceptual semantics is activated even during shallow lexical processing, such as in the lexical decision task. The MSE in lexical decision demonstrated by our study was a missing piece of evidence for the embodied view on semantics.

We believe that our study exhibited a significant effect of semantic modality information in lexical decision, unlike previous research (Scerrati et al., 2017), because of our novel paradigm in which we presented participants with two words at once. Therefore, the timing of processing semantic modality information might be the key factor influencing the probability of finding the MSE in lexical decisions (see also the Timing of Sensorimotor Semantic Activation section in Introduction). Interestingly, our results align with other studies on lexical decisions in Russian and German. For example, lexical decisions for Russian adjectives presented in isolation require approximately 802 ms (Jouravlev & Lupker, 2014; aggregated across all experimental conditions for adjectives; see Table 5 in that article), whereas the mean lexical decision RT for a German word is approximately 705 ms (Von Studnitz & Green, 1997; aggregated for trials with German words in same-language sequences; see Table 2). Indeed, we also found longer RTs for Russian (Experiment 1, mean RT = 1,159 ms for both words at once) than for German (Experiment 2, mean RT = 1,125 ms for both words at once; Note that due to simultaneous presentation, the processing of the two words in our study was not consecutive but instead overlapped, thus allowing for additional time savings and shorter RTs as one would get by simply doubling RTs for one word). Longer RTs in Russian in both our study and in the study by Jouravlev and Lupker (2014) as compared to German RTs might be the reason for partially different patterns of results across the two languages in our study: Processing of modality information for Word 1 and Word 2 could overlap more in Experiment 2 (German, see Figure 5) than in Experiment 1 (Russian, see Figure 3).

In this study, we used adjectives because they have higher unimodality than, for example, nouns (Lynott & Connell, 2009), which are often multimodal. Consider the word “loud”: Auditory information is central to this word’s semantics. On the other hand, for the word “cat,” the visual, haptic, and auditory modalities all have high ratings (Lynott et al., 2020). Thus, using adjectives could maximize the MSE in our study. Still, we suppose that the same underlying mechanisms—that is, activation of modality information—should occur in language processing regardless of the part of speech, even though it might be more challenging to detect such activation in the case of more multimodal concepts experimentally. Indeed, our results are generally in line with several other studies that found activation of embodied (sensorimotor) information in the lexical decision task (Harpaintner et al., 2020, 2022; e.g., Kiefer et al., 2008; Lynott et al., 2020; Pulvermüller et al., 2005; Ulrich et al., 2023; see also a study with bilinguals by Zhao et al., 2020).

It is essential to underline that our results cannot be accounted for by the “classical” psycholinguistic properties, such as length or frequency, and also not by the shallow lexical associations since we controlled for all these factors across our analyses. It cannot be excluded that some other psycholinguistic factors we did not explicitly control for, such as valence or arousal, could influence our results. However, note that we included items as random intercepts and therefore accounted to some extent for possible stimulus-related RT variability due to unknown parameters. At this point, the most convincing explanation for our results remains the activation of modality information associated with each concept.

Limitations and Future Directions

Our study has two potential limitations. First, we have no data on participants’ reading behavior. We do not know whether participants read the upper or the lower word first, whether there were systematic individual differences in this, and how long their dwell time was. Our paradigm, in which we simultaneously presented participants with two words, could be supplemented with eye tracking in future studies. This combination of methods would allow researchers to track the allocation of participants’ attention continuously and provide further insights into the timing of processing modality information.

The second potential limitation is related to the online testing used in both experiments: One can argue that some of our effects are unreliable due to additional data noise since the experiments were conducted online. However, a large study explicitly testing this issue found that online studies generally provide data of a quality comparable to that of lab-based studies (Bridges et al., 2020; see also Hartshorne et al., 2019), especially for visual stimuli like in our case. Furthermore, we followed several procedural steps to ensure the data quality (see, e.g., the Design and Procedure section), such as screen calibration, asking participants to keep a standardized distance from the screen, and minimizing disturbances. At the end of each experiment, our participants were also asked whether they had any technical problems during the testing session—in line with the recommendations provided in the literature on online testing (e.g., Rodd, 2024). Only 2%–6% of our participants across the experiments indicated technical problems, such as the need to restart the experiment or a cursor appearing on the screen. Therefore, we do not think the online testing and associated noise could significantly distort our results.

Implications

Our study has important implications for MSE research and embodied language processing research. First, and most importantly, we showed that the MSE occurs in the lexical decision task, that is, a task with no linguistic context and no explicit semantic processing required.

Second, we suggest that future studies carefully consider the exact timing of processing sensorimotor language to understand the underlying activation of perceptual semantics more deeply.

Third, our study suggests that semantic modalities should be analyzed separately in the MSE research since combinations of different modalities exhibit qualitatively different patterns of results (Connell & Lynott, 2010; see also Lam et al., 2015).

Fourth, we replicated the MSE in two new languages—Russian and German—which suggests the universal nature of this effect. Previous studies on semantic perceptual modalities were done in English (Connell & Lynott, 2011; Hauk et al., 2004), Italian (Scerrati et al., 2017), and Dutch (Pecher et al., 2004). However, no studies have examined the MSE in Russian and German before. Therefore, our study contributes to the cross-linguistic examination of the MSE. Moreover, preselected stimuli with controlled lexical characteristics are available for future studies in these two languages (see the Transparency and Openness section).

Last, our results provide further support for the embodied view of language processing (Barsalou, 1999, 2008, 2020; Bergen, 2015; Fischer & Zwaan, 2008; Hauk et al., 2004; Pecher et al., 2003): Our study demonstrated spontaneous activation of task-irrelevant sensory semantic information in a shallow lexical decision task.

Conclusion

In two experiments across different languages (Russian and German), we demonstrated that modality information is activated even during shallow lexical processing, such as in the lexical decision task. Our results suggest that the exact combination of semantic modalities modulates the activation of modality information. Thus, our study successfully demonstrated the MSE in lexical decisions. Our results strongly support the embodied view of language processing.

References

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133*(2), 283–316. https://doi.org/10.1037/0096-3445.133.2.283

Balota, D. A., Ferraro, F. R., & Connor, L. T. (1991). On the early influence of meaning in word recognition: A review of the literature. In P. J. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 187–222). Lawrence Erlbaum.

Barnette, J. J., & McLean, J. E. (2005). Type I error of four pairwise mean comparison procedures conducted as protected and unprotected tests. *Journal of Modern Applied Statistical Methods*, *4*(2), 446–459. https://doi.org/10.22237/jmasm/1130803740

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 577–660. https://doi.org/10.1017/S0140525X99002149

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*(1), 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Barsalou, L. W. (2020). Challenges and opportunities for grounding cognition. *Journal of Cognition*, *3*(1), Article 31. https://doi.org/10.5334/joc.116

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bergen, B. (2015). Embodiment, simulation and meaning. In P. H. Matthews (Ed.), *The Routledge handbook of semantics* (pp. 142–157). Routledge.

Bernabeu, P., Willems, R., & Louwerse, M. (2017). *Modality switch effects emerge early and increase throughout conceptual processing: Evidence from ERPs*. https://doi.org/10.31234/osf.io/a5pcz

Bodmer, F. M. (2005). COSMAS II. Recherchieren in den Korpora des IDS [COSMAS II. Researching in the IDS corpora]. *Sprachreport: Informationen und Meinungen zur deutschen Sprache*, *3*(21), 2–5.

Bottini, R., Morucci, P., D'Urso, A., Collignon, O., & Crepaldi, D. (2022). The concreteness advantage in lexical decision does not depend on perceptual simulations. *Journal of Experimental Psychology: General*, *151*(3), 731–738. https://doi.org/10.1037/xge0001090

Brehm, L., & Alday, P. M. (2022). Contrast coding choices in a decade of mixed models. *Journal of Memory and Language*, *125*, Article 104334. https://doi.org/10.1016/j.jml.2022.104334

Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: Comparing a range of experiment generators, both lab-based and online. *PeerJ*, *8*, Article e9414. https://doi.org/10.7717/peerj.9414

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology*, *58*(5), 412–424. https://doi.org/10.1027/1618-3169/a000123

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, *1*(1), Article 9. https://doi.org/10.5334/joc.10

Chen, I.-H., Zhao, Q., Long, Y., Lu, Q., & Huang, C.-R. (2019). Mandarin Chinese modality exclusivity norms. *PLOS ONE*, *14*(2), Article e0211336. https://doi.org/10.1371/journal.pone.0211336

Collins, J., Pecher, D., Zeelenberg, R., & Coulson, S. (2011). Modality switching in a property verification task: An ERP study of what happens when candles flicker after high heels click. *Frontiers in Psychology*, *2*, Article 10. https://doi.org/10.3389/fpsyg.2011.00010

Connell, L., & Lynott, D. (2010). Look but don't touch: Tactile disadvantage in processing modality-specific words. *Cognition*, *115*(1), 1–9. https://doi.org/10.1016/j.cognition.2009.10.005

Connell, L., & Lynott, D. (2011). Modality switching costs emerge in concept creation as well as retrieval. *Cognitive Science*, *35*(4), 763–778. https://doi.org/10.1111/j.1551-6709.2010.01168.x

Connell, L., & Lynott, D. (2012). Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition*, *125*(3), 452–465. https://doi.org/10.1016/j.cognition.2012.07.010

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*(2), 527–533. https://doi.org/10.1037/a0034626

Cortese, M. J., & Schock, J. (2013). Imageability and age of acquisition effects in disyllabic word recognition. *Quarterly Journal of Experimental Psychology*, *66*(5), 946–972. https://doi.org/10.1080/17470218.2012.722660

Estes, Z., & Barsalou, L. W. (2018). A comprehensive meta-analysis of spatial interference from linguistic cues: Beyond Petrova et al. (2018).

Psychological Science, *29*(9), 1558–1564. https://doi.org/10.1177/0956797618794131

Estes, Z., Verges, M., & Barsalou, L. W. (2008). Head up, foot down: Object words orient attention to the objects' typical location. *Psychological Science*, *19*(2), 93–97. https://doi.org/10.1111/j.1467-9280.2008.02051.x

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fernandino, L., Binder, J. R., Desai, R. H., Pendl, S. L., Humphries, C. J., Gross, W. L., Conant, L. L., & Seidenberg, M. S. (2016). Concept representation reflects multimodal abstraction: A framework for embodied semantics. *Cerebral Cortex*, *26*(5), 2018–2034. https://doi.org/10.1093/cercor/bhv020

Fischer, M. H., & Zwaan, R. A. (2008). Embodied language: A review of the role of the motor system in language comprehension. *Quarterly Journal of Experimental Psychology*, *61*(6), 825–850. https://doi.org/10.1080/17470210701623605

García, A. M., & Ibáñez, A. (2016). A touch with words: Dynamic synergies between manual actions and language. *Neuroscience and Biobehavioral Reviews*, *68*, 59–95. https://doi.org/10.1016/j.neubiorev.2016.04.022

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*(3), 558–565. https://doi.org/10.3758/BF03196313

Hald, L. A., Marshall, J.-A., Janssen, D. P., & Garnham, A. (2011). Switching modalities in a sentence verification task: ERP evidence for embodied language processing. *Frontiers in Psychology*, *2*, Article 45. https://doi.org/10.3389/fpsyg.2011.00045

Harpaintner, M., Sim, E.-J., Trumpp, N. M., Ulrich, M., & Kiefer, M. (2020). The grounding of abstract concepts in the motor and visual system: An fMRI study. *Cortex*, *124*, 1–22. https://doi.org/10.1016/j.cortex.2019.10.014

Harpaintner, M., Trumpp, N. M., & Kiefer, M. (2022). Time course of brain activity during the processing of motor- and vision-related abstract concepts: Flexibility and task dependency. *Psychological Research*, *86*, 2560–2582. https://doi.org/10.1007/s00426-020-01374-5

Hartshorne, J. K., de Leeuw, J. R., Goodman, N. D., Jennings, M., & O'Donnell, T. J. (2019). A thousand studies for the price of one: Accelerating psychological science with Pushkin. *Behavior Research Methods*, *51*(4), 1782–1803. https://doi.org/10.3758/s13428-018-1155-z

Hauk, O., Coutout, C., Holden, A., & Chen, Y. (2012). The time-course of single-word reading: Evidence from fast behavioral and brain responses. *NeuroImage*, *60*(2), 1462–1477. https://doi.org/10.1016/j.neuroimage.2012.01.061

Hauk, O., Davis, M. H., Ford, M., Pulvermüller, F., & Marslen-Wilson, W. D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *NeuroImage*, *30*(4), 1383–1400. https://doi.org/10.1016/j.neuroimage.2005.11.048

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*(2), 301–307. https://doi.org/10.1016/S0896-6273(03)00838-9

Heard, A., Madan, C. R., Protzner, A. B., & Pexman, P. M. (2019). Getting a grip on sensorimotor effects in lexical-semantic processing. *Behavior Research Methods*, *51*, 1–13. https://doi.org/10.3758/s13428-018-1072-1

Hino, Y., & Lupker, S. J. (1996). Effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(6), 1331–1356. https://doi.org/10.1037/0096-1523.22.6.1331

Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, *532*, 453–458. https://doi.org/10.1038/nature17637

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance*, *1*(2), 130–136. https://doi.org/10.1037/0096-1523.1.2.130

Jastrzembski, J. E. (1981). Multiple meanings, number of related meanings, frequency of occurrence, and the lexicon. *Cognitive Psychology*, *13*(2), 278–305. https://doi.org/10.1016/0010-0285(81)90011-6

Jouravlev, O., & Lupker, S. J. (2014). Stress consistency and stress regularity effects in Russian. *Language, Cognition and Neuroscience*, *29*(5), 605–619. https://doi.org/10.1080/01690965.2013.813562

Kiefer, M., Pielke, L., & Trumpp, N. M. (2022). Differential temporo-spatial pattern of electrical brain activity during the processing of abstract concepts related to mental states and verbal associations. *NeuroImage*, *252*, Article 119036. https://doi.org/10.1016/j.neuroimage.2022.119036

Kiefer, M., Sim, E.-J., Herrnberger, B., Grothe, J., & Hoenig, K. (2008). The sound of concepts: Four markers for a link between auditory and conceptual brain systems. *The Journal of Neuroscience*, *28*(47), 12224–12230. https://doi.org/10.1523/JNEUROSCI.3579-08.2008

Kiefer, M., Trumpp, N., Herrnberger, B., Sim, E.-J., Hoenig, K., & Pulvermüller, F. (2012). Dissociating the representation of action- and sound-related concepts in middle temporal cortex. *Brain and Language*, *122*(2), 120–125. https://doi.org/10.1016/j.bandl.2012.05.007

Kolbeneva, M. G., & Alexandrov, Y. I. (2010). *Sense organs, emotions and adjectives of the Russian language: Lingvo-psychological dictionary*. Institute of Psychology RAS.

Kuhnke, P., Kiefer, M., & Hartwigsen, G. (2020). Task-dependent recruitment of modality-specific and multimodal regions during conceptual processing. *Cerebral Cortex*, *30*(7), 3938–3959. https://doi.org/10.1093/cercor/bhaa010

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), 463–470. https://doi.org/10.1016/S1364-6613(00)01560-6

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*(1), 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Lam, K. J. Y., Dijkstra, T., & Rueschemeyer, S.-A. (2015). Feature activation during word recognition: Action, visual, and associative-semantic priming effects. *Frontiers in Psychology*, *6*, Article 659. https://doi.org/10.3389/fpsyg.2015.00659

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. https://doi.org/10.1037/0033-295X.104.2.211

Lenth, R. V., Bolker, B., Buerkner, P., Giné-Vázquez, I., Herve, M., Jung, M., Love, J., Miguez, F., Riebl, H., & Singmann, H. (2024). *emmeans: Estimated marginal means, aka least-squares means* (Version 1.10.0) [Computer software]. https://cran.r-project.org/web/packages/emmeans/index.html

Locker, L., Jr., Simpson, G. B., & Yates, M. (2003). Semantic neighborhood effects on the recognition of ambiguous words. *Memory & Cognition*, *31*(4), 505–515. https://doi.org/10.3758/BF03196092

Louwerse, M., & Hutchinson, S. (2012). Neurological evidence linguistic processes precede perceptual simulation in conceptual processing. *Frontiers in Psychology*, *3*, Article 385. https://doi.org/10.3389/fpsyg.2012.00385

Lyashevskaya, O. N., & Sharov, S. A. (2009). *Chastotnyjj slovar' sovremennogo russkogo jazyka na materialakh Nacional'nogo korpusa russkogo jazyka* [Frequency dictionary of modern Russian language (on materials of the Russian national corpus)]. Azbukovnik. https://dict.ruslang.ru/freq.php

Lynott, D., & Connell, L. (2009). Modality exclusivity norms for 423 object properties. *Behavior Research Methods*, *41*(2), 558–564. https://doi.org/10.3758/BRM.41.2.558

Lynott, D., & Connell, L. (2013). Modality exclusivity norms for 400 nouns: The relationship between perceptual experience and surface word form. *Behavior Research Methods*, *45*(2), 516–526. https://doi.org/10.3758/s13428-012-0267-0

Lynott, D., Connell, L., Brysbaert, M., Brand, J., & Carney, J. (2020). The Lancaster Sensorimotor Norms: Multidimensional measures of perceptual and action strength for 40,000 English words. *Behavior Research Methods*, *52*(3), 1271–1291. https://doi.org/10.3758/s13428-019-01316-z

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology, Paris*, *102*(1–3), 59–70. https://doi.org/10.1016/j.jphysparis.2008.03.004

Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex*, *48*(7), 788–804. https://doi.org/10.1016/j.cortex.2010.11.002

Miklashevsky, A. (2018). Perceptual experience norms for 506 Russian nouns: Modality rating, spatial localization, manipulability, imageability and other variables. *Journal of Psycholinguistic Research*, *47*(3), 641–661. https://doi.org/10.1007/s10936-017-9548-1

Muraki, E. J., Dahm, S. F., & Pexman, P. M. (2023). Meaning in hand: Investigating shared mechanisms of motor imagery and sensorimotor simulation in language processing. *Cognition*, *240*, Article 105589. https://doi.org/10.1016/j.cognition.2023.105589

Muraki, E. J., & Pexman, P. M. (2021). Simulating semantics: Are individual differences in motor imagery related to sensorimotor effects in language processing? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(12), 1939–1957. https://doi.org/10.1037/xlm0001039

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R 2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402–407. https://doi.org/10.3758/BF03195588

Papadopoulos, J., Domahs, F., & Kauschke, C. (2017). The role of sensory perception, emotionality and lifeworld in auditory word processing: Evidence from congenital blindness and synesthesia. *Journal of Psycholinguistic Research*, *46*(6), 1597–1623. https://doi.org/10.1007/s10936-017-9511-1

Pecher, D., van Dantzig, S., & Schifferstein, H. N. J. (2009). Concepts are not represented by conscious imagery. *Psychonomic Bulletin & Review*, *16*(5), 914–919. https://doi.org/10.3758/PBR.16.5.914

Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2003). Verifying different-modality properties for concepts produces switching costs. *Psychological Science*, *14*(2), 119–124. https://doi.org/10.1111/1467-9280.t01-1-01429

Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2004). Sensorimotor simulations underlie conceptual representations: Modality-specific effects of prior activation. *Psychonomic Bulletin & Review*, *11*(1), 164–167. https://doi.org/10.3758/BF03206477

Petrova, A., Navarrete, E., Suitner, C., Sulpizio, S., Reynolds, M., Job, R., & Peressotti, F. (2018). Spatial congruency effects exist, just not for words: Looking into Estes, Verges, and Barsalou (2008). *Psychological Science*, *29*(7), 1195–1199. https://doi.org/10.1177/0956797617728127

Platonova, O., & Miklashevsky, A. (2022). [Supplementary Materials] Warm and fuzzy: Perceptual semantics can be activated even during shallow lexical processing [Dataset]. OSF. https://doi.org/10.17605/OSF.IO/A85H7

Plungian, V., Reznikova, T., & Sichinava, D. (2005). Nacional'nyj korpus russkogo jazyka: Obshhaja harakteristika [National corpus of the Russian language: General description]. *Nauchno-Tehnicheskaja Informacija*, *3*, 9–13.

Pulvermüller, F. (2001). Brain reflections of words and their meaning. *Trends in Cognitive Sciences*, *5*(12), 517–524. https://doi.org/10.1016/S1364-6613(00)01803-9

Pulvermüller, F. (2012). Meaning and the brain: The neurosemantics of referential, interactive, and combinatorial knowledge. *Journal of Neurolinguistics*, *25*(5), 423–459. https://doi.org/10.1016/j.jneuroling.2011.03.004

Pulvermüller, F., Hauk, O., Nikulin, V. V., & Ilmoniemi, R. J. (2005). Functional links between motor and language systems. *The European Journal of Neuroscience*, *21*(3), 793–797. https://doi.org/10.1111/j.1460-9568.2005.03900.x

R Core Team. (2020). *R*: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/

Rodd, J. M. (2024). Moving experimental psychology online: How to obtain high quality data when we can't see our participants. *Journal of Memory and Language*, *134*, Article 104472. https://doi.org/10.1016/j.jml.2023.104472

Rychlý, P. (2008). *A lexicographer-friendly association score*. RASLAN, 6–9. https://is.muni.cz/publication/937193/en/A-Lexicographer-Friendly-Association-Score/Rychly

Scerrati, E., Baroni, G., Borghi, A. M., Galatolo, R., Lugli, L., & Nicoletti, R. (2015). The modality-switch effect: Visually and aurally presented prime sentences activate our senses. *Frontiers in Psychology*, *6*, Article 1668. https://doi.org/10.3389/fpsyg.2015.01668

Scerrati, E., Lugli, L., Nicoletti, R., & Borghi, A. M. (2017). The multilevel modality-switch effect: What happens when we see the bees buzzing and hear the diamonds glistening. *Psychonomic Bulletin & Review*, *24*(3), 798–803. https://doi.org/10.3758/s13423-016-1150-2

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, *110*, Article 104038. https://doi.org/10.1016/j.jml.2019.104038

Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language*, *27*(5), 499–520. https://doi.org/10.1016/0749-596X(88)90022-8

Shebani, Z., & Pulvermüller, F. (2018). Flexibility in language action interaction: The influence of movement type. *Frontiers in Human Neuroscience*, *12*, Article 252. https://doi.org/10.3389/fnhum.2018.00252

Snefjella, B., & Kuperman, V. (2016). It's all in the delivery: Effects of context valence, arousal, and concreteness on visual word processing. *Cognition*, *156*, 135–146. https://doi.org/10.1016/j.cognition.2016.07.010

Speed, L. J., & Majid, A. (2017). Dutch modality exclusivity norms: Simulating perceptual modality in space. *Behavior Research Methods*, *49*(6), 2204–2218. https://doi.org/10.3758/s13428-017-0852-3

Speed, L. J., & Majid, A. (2020). Grounding language in the neglected senses of touch, taste, and smell. *Cognitive Neuropsychology*, *37*(5–6), 363–392. https://doi.org/10.1080/02643294.2019.1623188

Tomasino, B., & Rumiati, R. I. (2013). At the mercy of strategies: The role of motor representations in language understanding. *Frontiers in Psychology*, *4*, Article 27. https://doi.org/10.3389/fpsyg.2013.00027

Ulrich, M., Harpaintner, M., Trumpp, N. M., Berger, A., & Kiefer, M. (2023). Academic training increases grounding of scientific concepts in experiential brain systems. *Cerebral Cortex*, *33*(9), 5646–5657. https://doi.org/10.1093/cercor/bhac449

van Dantzig, S., Cowell, R. A., Zeelenberg, R., & Pecher, D. (2011). A sharp image or a sharp knife: Norms for the modality-exclusivity of 774 concept-property items. *Behavior Research Methods*, *43*(1), 145–154. https://doi.org/10.3758/s13428-010-0038-8

Veale, J. F. (2014). Edinburgh Handedness Inventory—Short Form: A revised version based on confirmatory factor analysis. *Laterality: Asymmetries of Body, Brain, and Cognition*, *19*(2), 164–177. https://doi.org/10.1080/1357650X.2013.783045

Vergallito, A., Petilli, M. A., & Marelli, M. (2020). Perceptual modality norms for 1,121 Italian words: A comparison with concreteness and imageability scores and an analysis of their impact in word processing tasks. *Behavior Research Methods*, *52*(4), 1599–1616. https://doi.org/10.3758/s13428-019-01337-8

Vermeulen, N., Corneille, O., & Niedenthal, P. M. (2008). Sensory load incurs conceptual processing costs. *Cognition*, *109*(2), 287–294. https://doi.org/10.1016/j.cognition.2008.09.004

Von Studnitz, R. E., & Green, D. W. (1997). Lexical decision and language switching. *The International Journal of Bilingualism*, *1*(1), 3–24. https://doi.org/10.1177/136700699700100102

Willems, R. M., & Casasanto, D. (2011). Flexibility in embodied language understanding. *Frontiers in Psychology*, *2*, Article 116. https://doi.org/10.3389/fpsyg.2011.00116

Zhao, T., Huang, Y., Chen, D., Jiao, L., Marmolejo-Ramos, F., Wang, R., & Xie, J. (2020). The modality switching costs of Chinese–English bilinguals in the processing of L1 and L2. *Quarterly Journal of Experimental Psychology*, *73*(3), 396–412. https://doi.org/10.1177/1747021819878089

Zwaan, R. A. (2004). The Immersed Experiencer: Toward an embodied theory of language comprehension. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 44, pp. 35–62). Elsevier Science.

Appendix A

Results of Statistical Analyses

Table A1

Experiment 1: Output From the Final Model

Term	Name	Variance	*SD*	
Random effects				
Word 1	Intercept	0.0011	0.0333	
Word 2	Intercept	0.0010	0.0313	
Participant	Intercept	0.0413	0.2032	
Residual		0.0609	0.2468	
Term	***b***	***SE***	***t***	***p***
Fixed effects				
(Intercept)	6.9820	0.0280	249.2517	<.001
Word 1 (hap)	0.0146	0.0163	0.8933	.372
Word 1 (vis)	0.0117	0.0160	0.7330	.464
Word 2 (hap)	0.0443	0.0158	2.7960	.005
Word 2 (vis)	0.0151	0.0142	1.0602	.289
Word 1 (Hap) × Word 2 (Hap)	−0.0223	0.0172	−1.2934	.196
Word 1 (Vis) × Word 2 (Hap)	−0.0453	0.0173	−2.6142	.009
Word 1 (Hap) × Word 2 (Vis)	0.0025	0.0150	0.1692	.866
Word 1 (Vis) × Word 2 (Vis)	−0.0370	0.0149	−2.4752	.013
Covariate terms				
Frequency Word 1	−0.0493	0.0100	−4.9503	<.001
Frequency Word 2	−0.0608	0.0098	−6.2123	<.001
Length Word 1	0.0153	0.0038	3.9922	<.001
Length Word 2	0.0177	0.0038	4.7292	<.001
Trial Number	−0.0005	0.0000	−19.9628	<.001

Note. Marginal R^2 = .061, conditional R^2 = .451 (see Nakagawa & Schielzeth, 2013, for details). As indicated in the main text, in this and other output tables for linear mixed-effects models, covariate terms are included only where those were significant ($p < .050$) or marginally significant ($p < .100$). Key variables Word 1 and Word 2 and their interaction are always included regardless of significance. In this table and other tables representing main analysis outputs, the auditory modality of Word 1 and Word 2 was used as a baseline by default; therefore, terms for Word 1 and Word 2 shown in the table represent comparisons of the corresponding modalities to the baseline within the variable (main effects) or to the condition "Aud + Aud" (interactions). See the post hoc analysis in the next table for a full list of comparisons across all modality conditions. In this table, the following abbreviations are used for semantic modalities: Hap = haptic; Vis = visual; Aud = auditory; *SE* = standard error.

(*Appendices continue*)

Table A2
Pairwise Comparisons Across All Conditions in Experiment 1

Condition 1	Condition 2	Estimate	*SE*	*z* ratio	*p*
Aud + Aud	Hap + Aud	−0.0146	0.0163	−0.893	.993
Aud + Aud	Vis + Aud	−0.0117	0.0160	−0.733	.998
Aud + Aud	Aud + Hap	−0.0443	0.0158	−2.796	.116
Aud + Aud	Hap + Hap	−0.0366	0.0190	−1.930	.593
Aud + Aud	Vis + Hap	−0.0107	0.0188	−0.569	>.999
Aud + Aud	Aud + Vis	−0.0151	0.0142	−1.060	.980
Aud + Aud	Hap + Vis	−0.0322	0.0177	−1.816	.671
Aud + Aud	Vis + Vis	0.0101	0.0175	0.582	>.999
Hap + Aud	Vis + Aud	0.0028	0.0158	0.180	>.999
Hap + Aud	Aud + Hap	−0.0298	0.0191	−1.560	.826
Hap + Aud	Hap + Hap	−0.0220	0.0158	−1.397	.899
Hap + Aud	Vis + Hap	0.0039	0.0186	0.208	>.999
Hap + Aud	Aud + Vis	−0.0006	0.0179	−0.031	>.999
Hap + Aud	Hap + Vis	−0.0176	0.0143	−1.237	.948
Hap + Aud	Vis + Vis	0.0247	0.0173	1.432	.886
Vis + Aud	Aud + Hap	−0.0326	0.0188	−1.733	.726
Vis + Aud	Hap + Hap	−0.0249	0.0185	−1.341	.919
Vis + Aud	Vis + Hap	0.0010	0.0157	0.066	>.999
Vis + Aud	Aud + Vis	−0.0034	0.0176	−0.193	>.999
Vis + Aud	Hap + Vis	−0.0205	0.0173	−1.184	.960
Vis + Aud	Vis + Vis	0.0219	0.0143	1.532	.841
Aud + Hap	Hap + Hap	0.0077	0.0160	0.484	>.999
Aud + Hap	Vis + Hap	0.0336	0.0159	2.110	.466
Aud + Hap	Aud + Vis	0.0292	0.0140	2.083	.485
Aud + Hap	Hap + Vis	0.0121	0.0174	0.695	.999
Aud + Hap	**Vis + Vis**	**0.0545**	**0.0172**	**3.170**	**.041**
Hap + Hap	Vis + Hap	0.0259	0.0156	1.656	.773
Hap + Hap	Aud + Vis	0.0215	0.0176	1.223	.952
Hap + Hap	Hap + Vis	0.0044	0.0140	0.314	>.999
Hap + Hap	Vis + Vis	0.0467	0.0169	2.758	.128
Vis + Hap	Aud + Vis	−0.0044	0.0174	−0.254	>.999
Vis + Hap	Hap + Vis	−0.0215	0.0170	−1.264	.942
Vis + Hap	Vis + Vis	0.0208	0.0139	1.504	.854
Aud + Vis	Hap + Vis	−0.0171	0.0136	−1.252	.945
Aud + Vis	Vis + Vis	0.0253	0.0132	1.910	.607
Hap + Vis	**Vis + Vis**	**0.0424**	**0.0128**	**3.309**	**.026**

Note. This table shows post hoc comparisons across all conditions (combinations of levels of Word 1 and Word 2) performed using a Tukey multiple comparison test. An estimate is the difference between two conditions (on a log scale since the dependent variable was log-transformed; see the main text for details). Significant comparisons ($p < .050$) are shown in bold. In this table, the following abbreviations are used for semantic modalities: Aud = auditory; Hap = haptic; Vis = visual; *SE* = standard error.

(*Appendices continue*)

Table A3
Experiment 1: Output From the Final Model (Additional Analysis With the Data Aggregated by Modality Switch)

Term	Name	Variance	*SD*	
Random effects				
Word 1	Intercept	0.0012	0.0348	
Word 2	Intercept	0.0010	0.0321	
Participant	Intercept	0.0413	0.2032	
Residual		0.0610	0.2469	
Term	***b***	***SE***	***t***	***p***
Fixed effects				
(Intercept)	6.9966	0.0254	275.2557	<.001
Modality Switch	0.0145	0.0054	2.7059	.007
Covariate terms				
Frequency Word 1	−0.0482	0.0097	−4.9704	<.001
Frequency Word 2	−0.0551	0.0094	−5.8871	<.001
Length Word 1	0.0158	0.0039	4.0068	<.001
Length Word 2	0.0182	0.0038	4.7944	<.001
Trial Number	−0.0005	0.0000	−19.9640	<.001

Note. Marginal R^2 = .059, conditional R^2 = .451. See note under Table A1 for details.

Table A4
Experiment 2: Output From the Final Model

Term	Name	Variance	*SD*	
Random effects				
Word 1	Intercept	0.0007	0.0268	
Word 2	Intercept	0.0005	0.0227	
Participant	Intercept	0.0280	0.1673	
Residual		0.0527	0.2295	
Term	***b***	***SE***	***t***	***p***
Fixed effects				
(Intercept)	6.8787	0.0392	175.6037	<.001
Word 1 (hap)	0.0396	0.0151	2.6214	.009
Word 1 (vis)	0.0067	0.0150	0.4434	.657
Word 2 (hap)	0.0267	0.0144	1.8571	.063
Word 2 (vis)	0.0162	0.0145	1.1230	.261
Word 1 (Hap) × Word 2 (Hap)	−0.0285	0.0186	−1.5322	.126
Word 1 (Vis) × Word 2 (Hap)	0.0054	0.0178	0.3045	.761
Word 1 (Hap) × Word 2 (Vis)	−0.0506	0.0180	−2.8021	.005
Word 1 (Vis) × Word 2 (Vis)	−0.0588	0.0188	−3.1177	.002
Covariate terms				
Frequency Word 1	−0.0291	0.0056	−5.2008	<.001
Frequency Word 2	−0.0285	0.0053	−5.3771	<.001
Length Word 1	0.0118	0.0028	4.2379	<.001
Length Word 2	0.0086	0.0025	3.4653	.001
Trial Number	−0.0004	0.0000	−11.9686	<.001
Handedness	0.0011	0.0004	2.7684	.006
LogDice	0.0021	0.0012	1.6488	.099

Note. Marginal R^2 = .099, conditional R^2 = .421. See note under Table A1 for details. LogDice is the word co-occurrence measure. See the post hoc analysis in the next table for a full list of comparisons across all modality conditions.

(*Appendices continue*)

Table A5
Pairwise Comparisons Across All Conditions in Experiment 2

Condition 1	Condition 2	Estimate	*SE*	*z* ratio	*p*
Aud + Aud	Hap + Aud	−0.0396	0.0151	−2.621	.178
Aud + Aud	Vis + Aud	−0.0067	0.0150	−0.443	>.999
Aud + Aud	Aud + Hap	−0.0267	0.0144	−1.857	.644
Aud + Aud	Hap + Hap	−0.0378	0.0160	−2.359	.307
Aud + Aud	Vis + Hap	−0.0388	0.0163	−2.388	.290
Aud + Aud	Aud + Vis	−0.0162	0.0145	−1.123	.971
Aud + Aud	Hap + Vis	−0.0052	0.0163	−0.321	>.999
Aud + Aud	Vis + Vis	0.0359	0.0159	2.249	.374
Hap + Aud	Vis + Aud	0.0329	0.0143	2.308	.337
Hap + Aud	Aud + Hap	0.0128	0.0159	0.805	.997
Hap + Aud	Hap + Hap	0.0018	0.0143	0.125	>.999
Hap + Aud	Vis + Hap	0.0007	0.0159	0.046	>.999
Hap + Aud	Aud + Vis	0.0233	0.0159	1.466	.871
Hap + Aud	Hap + Vis	0.0343	0.0140	2.445	.259
Hap + Aud	**Vis + Vis**	**0.0754**	**0.0163**	**4.614**	**<.001**
Vis + Aud	Aud + Hap	−0.0201	0.0159	−1.263	.942
Vis + Aud	Hap + Hap	−0.0311	0.0162	−1.922	.599
Vis + Aud	Vis + Hap	−0.0322	0.0140	−2.299	.343
Vis + Aud	Aud + Vis	−0.0096	0.0159	−0.601	>.999
Vis + Aud	Hap + Vis	0.0014	0.0158	0.090	>.999
Vis + Aud	Vis + Vis	0.0425	0.0143	2.979	.071
Aud + Hap	Hap + Hap	−0.0110	0.0147	−0.750	.998
Aud + Hap	Vis + Hap	−0.0121	0.0143	−0.844	.996
Aud + Hap	Aud + Vis	0.0105	0.0137	0.768	.998
Aud + Hap	Hap + Vis	0.0215	0.0160	1.347	.917
Aud + Hap	**Vis + Vis**	**0.0626**	**0.0162**	**3.864**	**.004**
Hap + Hap	Vis + Hap	−0.0011	0.0144	−0.073	>.999
Hap + Hap	Aud + Vis	0.0216	0.0161	1.335	.921
Hap + Hap	Hap + Vis	0.0325	0.0142	2.297	.344
Hap + Hap	**Vis + Vis**	**0.0737**	**0.0159**	**4.633**	**<.001**
Vis + Hap	Aud + Vis	0.0226	0.0159	1.425	.889
Vis + Hap	Hap + Vis	0.0336	0.0158	2.120	.459
Vis + Hap	**Vis + Vis**	**0.0747**	**0.0143**	**5.221**	**<.001**
Aud + Vis	Hap + Vis	0.0110	0.0146	0.754	.998
Aud + Vis	**Vis + Vis**	**0.0521**	**0.0150**	**3.482**	**.015**
Hap + Vis	Vis + Vis	0.0411	0.0148	2.781	.121

Note. See note under Table A2 for details.

Table A6
Experiment 2: Output From the Final Model (Additional Analysis With the Data Aggregated by Modality Switch)

Term	Name	Variance	*SD*	
Random effects				
Word 1	Intercept	0.0008	0.0283	
Word 2	Intercept	0.0007	0.0274	
Participant	Intercept	0.0280	0.1672	
Residual		0.0528	0.2297	
Term	***b***	***SE***	***t***	***p***
Fixed effects				
(Intercept)	6.8908	0.0378	182.2979	<.001
Modality Switch	0.0163	0.0062	2.6400	.008
Covariate terms				
Frequency Word 1	−0.0266	0.0057	−4.6374	<.001
Frequency Word 2	−0.0263	0.0058	−4.5680	<.001
Length Word 1	0.0123	0.0029	4.2988	<.001
Length Word 2	0.0084	0.0027	3.0486	.002
Trial Number	−0.0004	0.0000	−11.7813	<.001
Handedness	0.0011	0.0004	2.7715	.006

Note. Marginal R^2 = .095, conditional R^2 = .420. See note under Table A1 for details.

(*Appendices continue*)

Appendix B

Experiment 3: Consequent Stimuli Presentation

Here, we report an experiment planned as an extension of the current series reported in the main text. Due to null and inconclusive results, we decided not to report this experiment in the main text. For transparency and openness, we still have included it in Appendix B.

Rationale: Timing of Sensorimotor Semantic Activation

As suggested in the article (the Timing of Sensorimotor Semantic Activation section), the timing of related cognitive processes is a crucial factor in the activation of sensorimotor semantics during shallow lexical processing. Experiments 1 and 2 demonstrated that the simultaneous presentation of stimuli pairs is enough to activate perceptual semantics and the emergence of the MSE in a shallow lexical decision task, at least for some modalities. This effect turned out to be consistent across Russian and German languages. However, Experiments 1 and 2 cannot provide much information about the exact timing of semantic processing. To explore the timing of the MSE more deeply, we designed Experiment 3.

As discussed in the introduction (see Figure 1 in the main text), modality activation might have two peaks: the earlier peak from 160 to 215 ms after stimulus onset and the later peak from 270 to 370 ms (Bernabeu et al., 2017; Hald et al., 2011; Kiefer et al., 2008; see also Louwerse & Hutchinson, 2012). In Experiment 3, we extended the method of Experiments 1 and 2 by manipulating the timing of the presentation of stimuli. We presented participants with two words of same versus different semantic modality sequentially, with stimulus-onset asynchrony (SOA) intervals of 170 versus 270 ms. Thus, in the condition with the SOA of 170 ms, the later activation of modality information for the first word (270–370 ms) should overlap with the early activation of modality information of the second word (160–215 ms, see Figure B1, Panel A), which might lead to interaction between perceptual modalities of Word 1 and Word 2. On the other hand, with the SOA of 270 ms, the activation of modality information for the two target words does not overlap (see Figure B1, Panel B), and thus, no interaction should be expected.

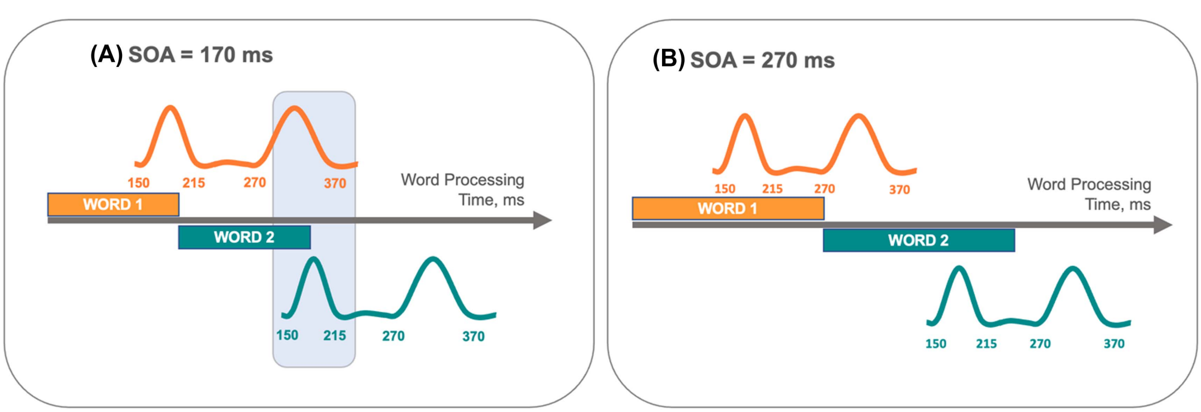
While previous research failed to find the MSE in lexical decisions (Scerrati et al., 2017), it might be that the time interval between the stimuli in consecutive trials was simply too long for the modality information from one stimulus (prime) to influence the processing of the next stimulus (target). We suggest that with a reduced time interval, or even without any interval between the two stimuli, it should be possible to induce the MSE in the lexical decision task. In contrast to previous work, in Experiments 1 and 2, we completely eliminated the time interval by presenting two words simultaneously, while in Experiment 3, we aimed to reduce the time interval between the presentation of two stimuli to 170 and 270 ms, therefore presenting two stimuli sequentially.

Stimuli Selection

The same German adjectives and pseudowords were used as in Experiment 2. As in Experiments 1 and 2, using a customized R script (see the Transparency and Openness section), we randomly selected stimuli pairs from the pools of stimuli for each modality. We did not control for whether each word appeared in the first versus second position. Instead, only word modality was relevant at this stage. Using this procedure, we formed four new lists, each including 540 pseudorandom stimuli pairs, where half of the pairs consisted of two real words and another half included at least one pseudoword. Pseudowords could appear at any position, that is, a pseudoword could be a first stimulus in a pair, a second stimulus in a pair, or both. In the next step, SOAs of 170 ms or 270 ms were randomly assigned to each pair of stimuli. On average, each combination of modalities (e.g., "visual + auditory") repeated with each SOA 60 times (min = 52 times, max = 68 times).

The same sources and approach were used for calculating word co-occurrences as in Experiment 2 (see the Stimuli Selection

Figure B1
Schematic Representation of Processing Modality Information in Experiment 3



Note. Panel A: Processing of sequentially presented pairs of stimuli with SOA = 170 ms. Panel B: Processing of pairs of stimuli with SOA = 270 ms. On both panels, colored rectangles denote presentation times of Word 1 and Word 2. SOA = stimulus-onset asynchrony. See the online article for the color version of this figure.

(*Appendices continue*)

section). In the COSMAS II Corpus, we found at least one co-occurrence for 56% of unique word pairs used in Experiment 3; the remaining word pairs never appeared together in the window of 1 + 1 sentences in the COSMAS II Corpus. Among those pairs we found in the corpus, the mean LogDice was equal to −3.32 (SD = 1.85). The minimum LogDice equaled to −7.98. We assigned the value of LogDice = −8 to all those pairs that do not appear in the corpus so that their LogDice value remained under the minimal value of pairs present in the corpus. This new variable was used as a covariate in the main analysis below.

Design and Procedure

At the beginning of the experiment, each participant was randomly assigned by Gorilla software to one of the four stimuli lists (see Design and Procedure sections of the main text for details); the order of stimuli presentation within the list was also randomized for each participant. Stimuli pairs were presented sequentially (i.e., "Word 1" → "Word 2"). Each stimulus appeared in the same position at the center of the screen. Participants were asked to perform a lexical decision task for both stimuli at once and respond by pressing a key after the presentation of the second stimulus. If both stimuli were existing German words, the participants were asked to press the "P" key on the keyboard; if at least one of the words was a pseudoword, participants were asked to press the "Q" key (counterbalanced across participants).

Depending on the SOA condition, the first word was presented for 170 ms (vs. 270 ms in the SOA = 270 condition) and immediately followed by the second word, which was also presented for 170 ms (vs. 270 ms in the SOA = 270 condition), followed by a blank screen. The response was registered starting from the onset of the second word. We set a timeout of 4,000 ms after the onset of the blank screen. Both speed and accuracy were emphasized in the instruction. See Figure B2 for a depiction of the procedure.

A practice including 10 trials preceded the experiment. The practice consisted of the adjectives taken from the actual list of stimuli. However, the stimulus pairs introduced during practice were novel and were not repeated during the actual experiment. At the end of the experiment, participant-related data were collected: age, gender, and Handedness (using an abbreviated version of the Edinburgh Handedness Inventory, see Veale, 2014). The Handedness score could range from −100 (*purely left-handed*) to +100 (*purely right-handed*). The entire duration of the experiment was 30 min. All technical details were identical to those of Experiment 1 (see the Design and Procedure section). Again, we did not strictly control the exact browser and the operation system used by the participants. The participants in Experiment 3 used the following browsers: *Google Chrome*, *Edge*, *Firefox*, *Opera*, or *Safari*. Participants in Experiment 3 used the following operating systems: *macOS*, *Ubuntu*, or *Windows*.

Participants

A total of 50 native German speakers participated in the online study. The same recruiting procedure was used for Experiment 2. In general, we followed the same exclusion criteria as for Experiments 1 and 2, with only one exception: We accepted all participants whose accuracy was 75% or higher because the task in Experiment 3 was more difficult (which is also indicated by the overall higher number of incorrect responses in Experiment 3: on average 6% more errors than in the other two experiments). No participants had mean RT > 2,500 ms or missing questionnaires. During preprocessing, data from four participants with *mean* accuracy of less than 75% were excluded. Data from 46 participants (seven male, 39 female; M_{age} = 23 years, SD = 7) were included in the final analysis. Participants' mean Handedness score was 71 (SD = 60; range from −100 to +100), indicating a prevailing right-handedness of the sample.

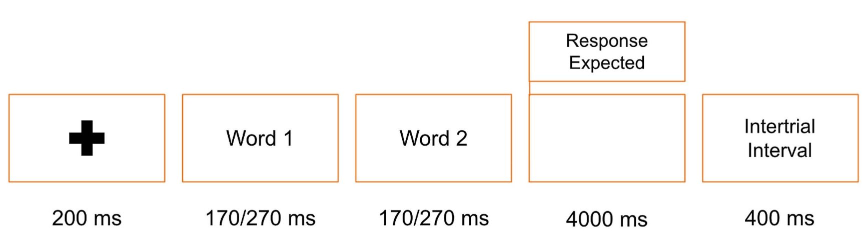
Analysis and Results

The average accuracy of the remaining set of participants was 89% (SD = 5%). Mean RT = 747, SD = 143. We did not analyze trials with pseudowords (N = 12,420). From now on, the remaining 12,420 trials are referred to as 100%. RT was measured from the onset of the second word, that is, from when participants received all information they needed to make a lexical decision. We removed trials with false responses (N = 995; 8% of the data) and timed out trials exceeding 3,000 ms (N = 12; less than 0.1% of the data). Note that we used different cutoff thresholds here because of the overall much faster responses in Experiment 3 than in Experiments 1 and 2—more than 400 ms faster on average—and based on visual examination of the RT distribution. Because there was a clear difference between RTs in the two SOA conditions, 41 ms longer RTs in the SOA = 170 ms; $F(1, 45) = 55.750$, $p < .001$, we decided to preprocess the two conditions separately. We removed trials with RTs outside of ±3 SDs from individual means by SOA condition (N = 254; 2% of the data). The remaining 11,159 trials (89.8%) were submitted for further analysis.

As in Experiment 1, we performed mixed-effects linear regressions. A similar approach was used as in Experiments 1 and 2: The initial model included Word 1 (semantic modality of

Figure B2

Experimental Procedure in Experiment 3



Note. See the online article for the color version of this figure.

(*Appendices continue*)

Word 1: auditory/haptic/visual), Word 2 (semantic modality of Word 1: auditory/haptic/visual), SOA as a binary variable (170 ms/270 ms), and its triple interaction with Word 1 and Word 2. We also controlled for the following covariate terms: Length of Word 1, Length of Word 2, Frequency of Word 1, Frequency of Word 2, LogDice (co-occurrence measure for each word pair, see Section 2.3.1), Trial Number, participant's Handedness (ranging from −100 to +100), Response Mapping (Q for words and P for pseudowords/P for words and Q for pseudowords), and the interaction between Handedness and Response Mapping.

All continuous predictors except Handedness were mean-centered. The binary variable Response Mapping was assigned sum-coded contrasts (−0.5 and 0.5). The dependent variable RT was log-transformed. Participants were included as random intercepts.

We employed backward elimination to identify the best-fitting models using the *drop1* function from the *lme4* package. Effects that were not significant or marginally significant ($p \geq .100$) and thus did not improve model fit were successively eliminated; the key terms of interest—that is, Word 1, Word 2, SOA, and the triple interaction between them—were preserved in the models regardless of significance.

Handedness and its interaction with Response Mapping turned nonsignificant; also, Length of Word 2 was not significant; all these terms were removed from the model. The key triple interaction between SOA, Word 1, and Word 2 was nonsignificant: The model with this interaction fits data no better than the model without it ($p =$ .297). All other factors were significant. Results from the final model are presented in Table B1.

We further examined the key interaction using a Tukey post hoc test and found no significant comparisons across levels of Word 1 and Word 2 within each of the SOAs (see Table B2 and Figure B3). In other words, there were no significant differences when modality combinations from the SOA = 170 ms conditions were compared to each other. The same were true when modality combinations from the SOA = 270 ms conditions were compared to each other. All significant differences emerged only between SOAs, with all conditions from the SOA = 170 ms being slower ($p < .001$) than all conditions from the SOA = 270 ms.

Table B1
Experiment 3: Output From the Final Model

Term	Name	Variance	*SD*	
Random effects				
Word 1	Intercept	0.0007	0.0274	
Word 2	Intercept	0.0014	0.0368	
Participant	Intercept	0.0581	0.2411	
Residual		0.1204	0.3470	
Term	***b***	***SE***	***t***	***p***
Fixed effects				
(Intercept)	6.0384	0.0401	150.6982	<.001
Word 1 (hap)	0.0324	0.0204	1.5910	.112
Word 1 (vis)	0.0238	0.0219	1.0894	.276
Word 2 (hap)	0.0241	0.0219	1.0993	.272
Word 2 (vis)	−0.0053	0.0231	−0.2284	.819
SOA	−0.4084	0.0276	−14.8221	<.001
Word 1 (Hap) × SOA	0.0519	0.0372	1.3930	.164
Word 1 (Vis) × SOA	0.0642	0.0396	1.6199	.105
Word 2 (Hap) × SOA	0.0510	0.0370	1.3778	.168
Word 2 (Vis) × SOA	0.0868	0.0388	2.2351	.025
Word 1 (Hap) × Word 2 (Hap)	−0.0260	0.0255	−1.0224	.307
Word 1 (Vis) × Word 2 (Hap)	−0.0296	0.0268	−1.1051	.269
Word 1 (Hap) × Word 2 (Vis)	−0.0390	0.0261	−1.4932	.135
Word 1 (Vis) × Word 2 (Vis)	−0.0234	0.0295	−0.7934	.428
Word 1 (Hap) × Word 2 (Hap) × SOA	−0.0506	0.0495	−1.0235	.306
Word 1 (Vis) × Word 2 (Hap) × SOA	−0.0403	0.0530	−0.7604	.447
Word 1 (Hap) × Word 2 (Vis) × SOA	−0.0556	0.0529	−1.0510	.293
Word 1 (Vis) × Word 2 (Vis) × SOA	−0.1145	0.0569	−2.0114	.044
Covariate terms				
Frequency Word 1	−0.0178	0.0081	−2.2123	.027
Frequency Word 2	−0.0314	0.0065	−4.8724	<.001
Length Word 1	0.0110	0.0038	2.8713	.004
Response Mapping	−0.1581	0.0728	−2.1725	.030
Trial Number	−0.0004	0.0000	−16.1382	<.001
LogDice	0.0043	0.0019	2.2302	.026

Note. Marginal $R^2 = .201$, conditional $R^2 = .467$. See note under Table A1 for details. LogDice is the word co-occurrence measure. See the post hoc analysis in the next table for a full list of comparisons across all modalities and SOA conditions. SOA = stimulus-onset asynchrony.

(*Appendices continue*)

Table B2
Pairwise Comparisons Across All Conditions in Experiment 3

Condition 1	Condition 2	Estimate	*SE*	*z* ratio	*p*
Comparisons within SOA 170 ms					
SOA_170: Aud + Aud	SOA_170: Hap + Aud	−0.0065	0.0259	−0.250	>.999
SOA_170: Aud + Aud	SOA_170: Vis + Aud	0.0082	0.0279	0.296	>.999
SOA_170: Aud + Aud	SOA_170: Aud + Hap	0.0014	0.0267	0.052	>.999
SOA_170: Aud + Aud	SOA_170: Hap + Hap	−0.0044	0.0267	−0.165	>.999
SOA_170: Aud + Aud	SOA_170: Vis + Hap	0.0191	0.0291	0.657	>.999
SOA_170: Aud + Aud	SOA_170: Aud + Vis	0.0487	0.0298	1.636	.977
SOA_170: Aud + Aud	SOA_170: Hap + Vis	0.0534	0.0281	1.899	.913
SOA_170: Aud + Aud	SOA_170: Vis + Vis	0.0230	0.0297	0.776	>.999
SOA_170: Hap + Aud	SOA_170: Vis + Aud	0.0148	0.0271	0.544	>.999
SOA_170: Hap + Aud	SOA_170: Aud + Hap	0.0079	0.0272	0.290	>.999
SOA_170: Hap + Aud	SOA_170: Hap + Hap	0.0021	0.0248	0.084	>.999
SOA_170: Hap + Aud	SOA_170: Vis + Hap	0.0256	0.0287	0.892	>.999
SOA_170: Hap + Aud	SOA_170: Aud + Vis	0.0553	0.0304	1.813	.941
SOA_170: Hap + Aud	SOA_170: Hap + Vis	0.0599	0.0261	2.296	.686
SOA_170: Hap + Aud	SOA_170: Vis + Vis	0.0296	0.0299	0.990	>.999
SOA_170: Vis + Aud	SOA_170: Aud + Hap	−0.0069	0.0291	−0.237	>.999
SOA_170: Vis + Aud	SOA_170: Hap + Hap	−0.0127	0.0286	−0.444	>.999
SOA_170: Vis + Aud	SOA_170: Vis + Hap	0.0108	0.0291	0.373	>.999
SOA_170: Vis + Aud	SOA_170: Aud + Vis	0.0404	0.0321	1.258	.999
SOA_170: Vis + Aud	SOA_170: Hap + Vis	0.0451	0.0294	1.534	.988
SOA_170: Vis + Aud	SOA_170: Vis + Vis	0.0148	0.0305	0.485	>.999
SOA_170: Aud + Hap	SOA_170: Hap + Hap	−0.0058	0.0240	−0.241	>.999
SOA_170: Aud + Hap	SOA_170: Vis + Hap	0.0177	0.0265	0.669	>.999
SOA_170: Aud + Hap	SOA_170: Aud + Vis	0.0473	0.0291	1.627	.978
SOA_170: Aud + Hap	SOA_170: Hap + Vis	0.0520	0.0272	1.913	.908
SOA_170: Aud + Hap	SOA_170: Vis + Vis	0.0217	0.0297	0.730	>.999
SOA_170: Hap + Hap	SOA_170: Vis + Hap	0.0235	0.0248	0.948	>.999
SOA_170: Hap + Hap	SOA_170: Aud + Vis	0.0531	0.0294	1.805	.943
SOA_170: Hap + Hap	SOA_170: Hap + Vis	0.0578	0.0249	2.321	.668
SOA_170: Hap + Hap	SOA_170: Vis + Vis	0.0275	0.0283	0.972	>.999
SOA_170: Vis + Hap	SOA_170: Aud + Vis	0.0296	0.0315	0.939	>.999
SOA_170: Vis + Hap	SOA_170: Hap + Vis	0.0343	0.0286	1.198	.999
SOA_170: Vis + Hap	SOA_170: Vis + Vis	0.0040	0.0293	0.136	>.999
SOA_170: Aud + Vis	SOA_170: Hap + Vis	0.0047	0.0279	0.168	>.999
SOA_170: Aud + Vis	SOA_170: Vis + Vis	−0.0256	0.0304	−0.843	>.999
SOA_170: Hap + Vis	SOA_170: Vis + Vis	−0.0303	0.0270	−1.124	>.999
Comparisons within SOA 270 ms					
SOA_270: Aud + Aud	SOA_270: Hap + Aud	−0.0584	0.0292	−1.999	.870
SOA_270: Aud + Aud	SOA_270: Vis + Aud	−0.0559	0.0310	−1.805	.943
SOA_270: Aud + Aud	SOA_270: Aud + Hap	−0.0496	0.0306	−1.622	.979
SOA_270: Aud + Aud	SOA_270: Hap + Hap	−0.0566	0.0303	−1.869	.923
SOA_270: Aud + Aud	SOA_270: Vis + Hap	−0.0558	0.0307	−1.819	.939
SOA_270: Aud + Aud	SOA_270: Aud + Vis	−0.0381	0.0306	−1.245	.999
SOA_270: Aud + Aud	SOA_270: Hap + Vis	−0.0297	0.0320	−0.928	>.999
SOA_270: Aud + Aud	SOA_270: Vis + Vis	−0.0134	0.0342	−0.391	>.999
SOA_270: Hap + Aud	SOA_270: Vis + Aud	0.0025	0.0281	0.088	>.999
SOA_270: Hap + Aud	SOA_270: Aud + Hap	0.0087	0.0293	0.298	>.999
SOA_270: Hap + Aud	SOA_270: Hap + Hap	0.0017	0.0265	0.065	>.999
SOA_270: Hap + Aud	SOA_270: Vis + Hap	0.0026	0.0281	0.093	>.999
SOA_270: Hap + Aud	SOA_270: Aud + Vis	0.0202	0.0293	0.691	>.999
SOA_270: Hap + Aud	SOA_270: Hap + Vis	0.0286	0.0283	1.010	>.999
SOA_270: Hap + Aud	SOA_270: Vis + Vis	0.0450	0.0324	1.391	.996
SOA_270: Vis + Aud	SOA_270: Aud + Hap	0.0063	0.0308	0.203	>.999
SOA_270: Vis + Aud	SOA_270: Hap + Hap	−0.0008	0.0294	−0.026	>.999
SOA_270: Vis + Aud	SOA_270: Vis + Hap	0.0001	0.0280	0.005	>.999
SOA_270: Vis + Aud	SOA_270: Aud + Vis	0.0178	0.0307	0.580	>.999
SOA_270: Vis + Aud	SOA_270: Hap + Vis	0.0262	0.0310	0.845	>.999
SOA_270: Vis + Aud	SOA_270: Vis + Vis	0.0425	0.0324	1.313	.998
SOA_270: Aud + Hap	SOA_270: Hap + Hap	−0.007012	0.0260	−0.269	>.999
SOA_270: Aud + Hap	SOA_270: Vis + Hap	−0.0061	0.0265	−0.231	>.999
SOA_270: Aud + Hap	SOA_270: Aud + Vis	0.0115	0.0283	0.407	>.999
SOA_270: Aud + Hap	SOA_270: Hap + Vis	0.0199	0.0299	0.666	>.999
SOA_270: Aud + Hap	SOA_270: Vis + Vis	0.0363	0.0329	1.102	>.999

(*table continues*)

(*Appendices continue*)

Table B2 (*continued*)

Condition 1	Condition 2	Estimate	*SE*	*z* ratio	*p*
SOA_270: Hap + Hap	SOA_270: Vis + Hap	0.0009	0.0246	0.036	>.999
SOA_270: Hap + Hap	SOA_270: Aud + Vis	0.0185	0.0283	0.654	>.999
SOA_270: Hap + Hap	SOA_270: Hap + Vis	0.0269	0.0272	0.989	>.999
SOA_270: Hap + Hap	SOA_270: Vis + Vis	0.0433	0.0312	1.389	.996
SOA_270: Vis + Hap	SOA_270: Aud + Vis	0.0176	0.0285	0.619	>.999
SOA_270: Vis + Hap	SOA_270: Hap + Vis	0.0260	0.0286	0.909	>.999
SOA_270: Vis + Hap	SOA_270: Vis + Vis	0.0424	0.0304	1.395	.996
SOA_270: Aud + Vis	SOA_270: Hap + Vis	0.0084	0.0274	0.306	>.999
SOA_270: Aud + Vis	SOA_270: Vis + Vis	0.0248	0.0303	0.818	>.999
SOA_270: Hap + Vis	SOA_270: Vis + Vis	0.0164	0.0305	0.537	>.999
		Comparisons across SOAs			
SOA_170: Aud + Aud	**SOA_270: Aud + Aud**	**0.4084**	**0.0276**	**14.822**	**< .001**
SOA_170: Aud + Aud	**SOA_270: Hap + Aud**	**0.3501**	**0.0263**	**13.294**	**<.001**
SOA_170: Aud + Aud	**SOA_270: Vis + Aud**	**0.3526**	**0.0277**	**12.714**	**<.001**
SOA_170: Aud + Aud	**SOA_270: Aud + Hap**	**0.3588**	**0.0280**	**12.804**	**<.001**
SOA_170: Aud + Aud	**SOA_270: Hap + Hap**	**0.3518**	**0.0274**	**12.822**	**<.001**
SOA_170: Aud + Aud	**SOA_270: Vis + Hap**	**0.3527**	**0.0280**	**12.582**	**<.001**
SOA_170: Aud + Aud	**SOA_270: Aud + Vis**	**0.3703**	**0.0278**	**13.325**	**<.001**
SOA_170: Aud + Aud	**SOA_270: Hap + Vis**	**0.3787**	**0.0296**	**12.782**	**<.001**
SOA_170: Aud + Aud	**SOA_270: Vis + Vis**	**0.3951**	**0.0318**	**12.424**	**<.001**
SOA_270: Aud + Aud	**SOA_170: Hap + Aud**	**−0.4149**	**0.0283**	**−14.665**	**<.001**
SOA_270: Aud + Aud	**SOA_170: Vis + Aud**	**−0.4002**	**0.0304**	**−13.155**	**<.001**
SOA_270: Aud + Aud	**SOA_170: Aud + Hap**	**−0.4070**	**0.0294**	**−13.845**	**<.001**
SOA_270: Aud + Aud	**SOA_170: Hap + Hap**	**−0.4129**	**0.0295**	**−14.003**	**<.001**
SOA_270: Aud + Aud	**SOA_170: Vis + Hap**	**−0.3893**	**0.0317**	**−12.269**	**<.001**
SOA_270: Aud + Aud	**SOA_170: Aud + Vis**	**−0.3598**	**0.0322**	**−11.177**	**<.001**
SOA_270: Aud + Aud	**SOA_170: Hap + Vis**	**−0.3551**	**0.0309**	**−11.499**	**<.001**
SOA_270: Aud + Aud	**SOA_170: Vis + Vis**	**−0.3854**	**0.0324**	**−11.885**	**<.001**
SOA_170: Hap + Aud	**SOA_270: Hap + Aud**	**0.3566**	**0.0244**	**14.590**	**<.001**
SOA_170: Hap + Aud	**SOA_270: Vis + Aud**	**0.3590**	**0.0273**	**13.165**	**<.001**
SOA_170: Hap + Aud	**SOA_270: Aud + Hap**	**0.3653**	**0.0286**	**12.781**	**<.001**
SOA_170: Hap + Aud	**SOA_270: Hap + Hap**	**0.3583**	**0.0257**	**13.932**	**<.001**
SOA_170: Hap + Aud	**SOA_270: Vis + Hap**	**0.3592**	**0.0272**	**13.206**	**<.001**
SOA_170: Hap + Aud	**SOA_270: Aud + Vis**	**0.3768**	**0.0286**	**13.170**	**<.001**
SOA_170: Hap + Aud	**SOA_270: Hap + Vis**	**0.3852**	**0.0275**	**13.999**	**<.001**
SOA_170: Hap + Aud	**SOA_270: Vis + Vis**	**0.4016**	**0.0318**	**12.622**	**<.001**
SOA_270: Hap + Aud	**SOA_170: Vis + Aud**	**−0.3418**	**0.0278**	**−12.285**	**<.001**
SOA_270: Hap + Aud	**SOA_170: Aud + Hap**	**−0.3487**	**0.0280**	**−12.435**	**<.001**
SOA_270: Hap + Aud	**SOA_170: Hap + Hap**	**−0.3545**	**0.0258**	**−13.718**	**<.001**
SOA_270: Hap + Aud	**SOA_170: Vis + Hap**	**−0.3310**	**0.0293**	**−11.280**	**<.001**
SOA_270: Hap + Aud	**SOA_170: Aud + Vis**	**−0.3014**	**0.0311**	**−9.682**	**<.001**
SOA_270: Hap + Aud	**SOA_170: Hap + Vis**	**−0.2967**	**0.0270**	**−11.008**	**<.001**
SOA_270: Hap + Aud	**SOA_170: Vis + Vis**	**−0.3270**	**0.0305**	**−10.737**	**<.001**
SOA_170: Vis + Aud	**SOA_270: Vis + Aud**	**0.3443**	**0.0276**	**12.454**	**<.001**
SOA_170: Vis + Aud	**SOA_270: Aud + Hap**	**0.3505**	**0.0305**	**11.498**	**<.001**
SOA_170: Vis + Aud	**SOA_270: Hap + Hap**	**0.3435**	**0.0294**	**11.703**	**<.001**
SOA_170: Vis + Aud	**SOA_270: Vis + Hap**	**0.3444**	**0.0279**	**12.338**	**<.001**
SOA_170: Vis + Aud	**SOA_270: Aud + Vis**	**0.3621**	**0.0304**	**11.900**	**<.001**
SOA_170: Vis + Aud	**SOA_270: Hap + Vis**	**0.3704**	**0.0309**	**12.006**	**<.001**
SOA_170: Vis + Aud	**SOA_270: Vis + Vis**	**0.3868**	**0.0322**	**12.016**	**<.001**
SOA_270: Vis + Aud	**SOA_170: Aud + Hap**	**−0.3512**	**0.0294**	**−11.941**	**<.001**
SOA_270: Vis + Aud	**SOA_170: Hap + Hap**	**−0.3570**	**0.0285**	**−12.512**	**<.001**
SOA_270: Vis + Aud	**SOA_170: Vis + Hap**	**−0.3335**	**0.0292**	**−11.411**	**<.001**
SOA_270: Vis + Aud	**SOA_170: Aud + Vis**	**−0.3039**	**0.0324**	**−9.371**	**<.001**
SOA_270: Vis + Aud	**SOA_170: Hap + Vis**	**−0.2992**	**0.0295**	**−10.158**	**<.001**
SOA_270: Vis + Aud	**SOA_170: Vis + Vis**	**−0.3295**	**0.0303**	**−10.870**	**<.001**
SOA_170: Aud + Hap	**SOA_270: Aud + Hap**	**0.3574**	**0.0248**	**14.438**	**<.001**
SOA_170: Aud + Hap	**SOA_270: Hap + Hap**	**0.3504**	**0.0248**	**14.131**	**<.001**
SOA_170: Aud + Hap	**SOA_270: Vis + Hap**	**0.3513**	**0.0249**	**14.105**	**<.001**
SOA_170: Aud + Hap	**SOA_270: Aud + Vis**	**0.3689**	**0.0271**	**13.611**	**<.001**
SOA_170: Aud + Hap	**SOA_270: Hap + Vis**	**0.3773**	**0.0286**	**13.187**	**<.001**
SOA_170: Aud + Hap	**SOA_270: Vis + Vis**	**0.3937**	**0.0317**	**12.431**	**<.001**
SOA_270: Aud + Hap	**SOA_170: Hap + Hap**	**−0.3632**	**0.0254**	**−14.313**	**<.001**
SOA_270: Aud + Hap	**SOA_170: Vis + Hap**	**−0.3397**	**0.0276**	**−12.295**	**<.001**
SOA_270: Aud + Hap	**SOA_170: Aud + Vis**	**−0.3101**	**0.0299**	**−10.355**	**<.001**
SOA_270: Aud + Hap	**SOA_170: Hap + Vis**	**−0.3054**	**0.0286**	**−10.674**	**<.001**

(*table continues*)

(*Appendices continue*)

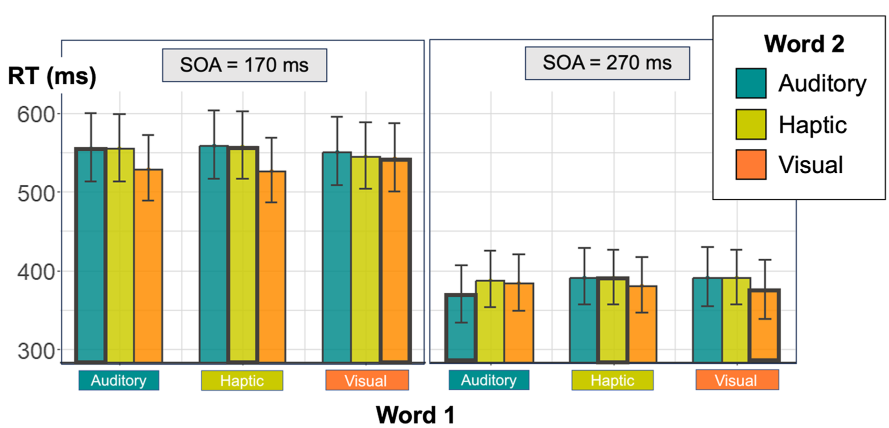
Table B2 (*continued*)

Condition 1	Condition 2	Estimate	*SE*	*z* ratio	*p*
SOA_270: Aud + Hap	**SOA_170: Vis + Vis**	**−0.3357**	**0.0311**	**−10.787**	**<.001**
SOA_170: Hap + Hap	**SOA_270: Hap + Hap**	**0.3562**	**0.0216**	**16.511**	**<.001**
SOA_170: Hap + Hap	**SOA_270: Vis + Hap**	**0.3571**	**0.0235**	**15.205**	**<.001**
SOA_170: Hap + Hap	**SOA_270: Aud + Vis**	**0.3747**	**0.0275**	**13.624**	**<.001**
SOA_170: Hap + Hap	**SOA_270: Hap + Vis**	**0.3831**	**0.0265**	**14.448**	**<.001**
SOA_170: Hap + Hap	**SOA_270: Vis + Vis**	**0.3995**	**0.0304**	**13.152**	**<.001**
SOA_270: Hap + Hap	**SOA_170: Vis + Hap**	**−0.3327**	**0.0260**	**−12.800**	**<.001**
SOA_270: Hap + Hap	**SOA_170: Aud + Vis**	**−0.3031**	**0.0302**	**−10.041**	**<.001**
SOA_270: Hap + Hap	**SOA_170: Hap + Vis**	**−0.2984**	**0.0257**	**−11.604**	**<.001**
SOA_270: Hap + Hap	**SOA_170: Vis + Vis**	**−0.3287**	**0.0289**	**−11.357**	**<.001**
SOA_170: Vis + Hap	**SOA_270: Vis + Hap**	**0.3336**	**0.0247**	**13.526**	**<.001**
SOA_170: Vis + Hap	**SOA_270: Aud + Vis**	**0.3512**	**0.0298**	**11.798**	**<.001**
SOA_170: Vis + Hap	**SOA_270: Hap + Vis**	**0.3596**	**0.0300**	**11.989**	**<.001**
SOA_170: Vis + Hap	**SOA_270: Vis + Vis**	**0.3760**	**0.0315**	**11.951**	**<.001**
SOA_270: Vis + Hap	**SOA_170: Aud + Vis**	**−0.3040**	**0.0303**	**−10.021**	**<.001**
SOA_270: Vis + Hap	**SOA_170: Hap + Vis**	**−0.2993**	**0.0273**	**−10.984**	**<.001**
SOA_270: Vis + Hap	**SOA_170: Vis + Vis**	**−0.3296**	**0.0282**	**−11.699**	**<.001**
SOA_170: Aud + Vis	**SOA_270: Aud + Vis**	**0.3216**	**0.0271**	**11.878**	**<.001**
SOA_170: Aud + Vis	**SOA_270: Hap + Vis**	**0.3300**	**0.0293**	**11.258**	**<.001**
SOA_170: Aud + Vis	**SOA_270: Vis + Vis**	**0.3464**	**0.0320**	**10.811**	**<.001**
SOA_270: Aud + Vis	**SOA_170: Hap + Vis**	**−0.3170**	**0.0257**	**−12.320**	**<.001**
SOA_270: Aud + Vis	**SOA_170: Vis + Vis**	**−0.3472**	**0.0283**	**−12.264**	**<.001**
SOA_170: Hap + Vis	**SOA_270: Hap + Vis**	**0.3253**	**0.0252**	**12.919**	**<.001**
SOA_170: Hap + Vis	**SOA_270: Vis + Vis**	**0.3417**	**0.0292**	**11.717**	**<.001**
SOA_270: Hap + Vis	**SOA_170: Vis + Vis**	**−0.3556**	**0.0283**	**−12.551**	**<.001**
SOA_170: Vis + Vis	**SOA_270: Vis + Vis**	**0.3720**	**0.0296**	**12.585**	**<.001**

Note. See note under Table A2 for details. SOA = stimulus-onset asynchrony.

Figure B3

Mean Reaction Times in 18 Conditions (Experiment 3)



Note. Combinations of semantic modalities: Word 1 × Word 2. Whiskers represent standard errors. Thicker inner frames indicate no-switch conditions. RT = reaction time; SOA = stimulus-onset asynchrony. See the online article for the color version of this figure.

(*Appendices continue*)

Discussion (Experiment 3)

Experiment 3 was meant to extend the method by considering the neuroimaging evidence on the timing of word processing discussed above (see Figure 1 of the main text). In Experiment 3, we sequentially presented native speakers of German with two words of the same versus different modalities with a varying time interval between the onset of the two words (170 vs. 270 ms). Similarly to Experiments 1 and 2, a lexical decision for both words at once was required. However, contrary to our expectations, no MSE was found within each of the two time intervals. It might be that Experiment 3 was simply too demanding to participants who only could see Word 1 for 170 or 270 ms, and therefore, this difficulty overshadowed any semantic effects that we aimed to find. Future research might explore other paradigms controlling for presentation time (such as self-paced reading) or examining processing time (such as eye tracking), which do not impose too high processing demands on participants.

Received June 20, 2023
Revision received September 19, 2024
Accepted September 25, 2024 ■