

Contents lists available at [ScienceDirect](http://ScienceDirect.com)

# Journal of Memory and Language

journal homepage: [www.elsevier.com/locate/jml](http://www.elsevier.com/locate/jml)

## The impact of recent and long-term experience on access to word meanings: Evidence from large-scale internet-based experiments



Jennifer M. Rodd\*, Zhenguang G. Cai, Hannah N. Betts, Betsy Hanby, Catherine Hutchinson, Aviva Adler

Department of Experimental Psychology, University College London, UK

### ARTICLE INFO

#### Article history:

Received 12 December 2013  
revision received 16 October 2015  
Available online 6 November 2015

#### Keywords:

Semantic ambiguity  
Lexical ambiguity  
Perceptual learning  
Priming  
Comprehension  
Web-based experiment

### ABSTRACT

Many word forms map onto multiple meanings (e.g., “ace”). The current experiments explore the extent to which adults reshape the lexical–semantic representations of such words on the basis of experience, to increase the availability of more recently accessed meanings. A naturalistic web-based experiment in which primes were presented within a radio programme (Experiment 1;  $N = 1800$ ) and a lab-based experiment (Experiment 2) show that when listeners have encountered one or two disambiguated instances of an ambiguous word, they then retrieve this primed meaning more often (compared with an unprimed control condition). This word–meaning priming lasts up to 40 min after exposure, but decays very rapidly during this interval. Experiments 3 and 4 explore longer-term word–meaning priming by measuring the impact of more extended, naturalistic encounters with ambiguous words: recreational rowers ( $N = 213$ ) retrieved rowing-related meanings for words (e.g., “feather”) more often if they had rowed that day, despite a median delay of 8 hours. The rate of rowing-related interpretations also increased with additional years’ rowing experience. Taken together these experiments show that individuals’ overall meaning preferences reflect experience across a wide range of timescales from minutes to years. In addition, priming was not reduced by a change in speaker identity (Experiment 1), suggesting that the phenomenon occurs at a relatively abstract lexical–semantic level. The impact of experience was reduced for older adults (Experiments 1, 3, 4) suggesting that the lexical–semantic representations of younger listeners may be more malleable to current linguistic experience.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

### Introduction

The ability to rapidly and accurately retrieve word meanings is critical for successful language comprehension, but is particularly challenging for words with multiple meanings. For example, when reading “the boy heard

the loud BARK”, the reader must determine that the final word most likely refers to the sound made by a dog and not the outer covering of a tree. Given that over 80% of common English words have more than one dictionary definition (Rodd, Gaskell, & Marslen-Wilson, 2002), the processes that enable readers/listeners to select appropriate word meanings and reject contextually inappropriate meanings form a core (and much studied) component of the language comprehension system (Twilley & Dixon, 2000; Vitello & Rodd, 2015). The current experiments

\* Corresponding author at: Department of Experimental Psychology, University College London, 26 Bedford Way, London WC1H 0AP, UK.

E-mail address: [j.rodd@ucl.ac.uk](mailto:j.rodd@ucl.ac.uk) (J.M. Rodd).

explore the hypothesis that learning mechanisms make an important, and previously underestimated, contribution to the efficiency with which ambiguous words are processed. We propose that lexical–semantic representations are reshaped throughout our adult lives on the basis of our linguistic experiences, such that information about how these words have been used across a wide range of timescales, from minutes to years, allows individuals to make better predictions about which meanings are more likely to be encountered in the future.

That the language system continues to be shaped by linguistic input throughout adulthood is now well established. Adult speakers can learn new word forms that enter the language (e.g., “blog”; Gaskell & Dumay, 2003) as well as new meanings for words they already know (e.g., “twitter”; Rodd et al., 2012). In addition to these abilities to learn *new* linguistic representations, an increasing body of evidence has shown that adults are remarkably skilled at ‘fine-tuning’ their *existing* linguistic representations to improve future comprehension. For example, adult listeners can rapidly adapt to unfamiliar speech accents (Adank, Evans, Stuart-Smith, & Scott, 2009; Bradlow & Bent, 2008; Clarke & Garrett, 2004; Cristia et al., 2012) and to the idiosyncratic pronunciation habits of unfamiliar speakers (Mullennix, Pisoni, & Martin, 1989; Nygaard & Pisoni, 1998). Similar retuning has also been observed at the syntactic level, where past syntactic experience helps listeners to predict upcoming words/phrases in sentence comprehension (Arai, Van Gompel, & Scheepers, 2007; Fine & Jaeger, 2013; Traxler, 2008). These adaptations at different linguistic levels and at different timescales, which retune listeners’ representations on the basis of experience, are now well established as making a critical contribution to language comprehension. For instance, adaptations to speech helps listeners’ ability to deal with between-talker variability, allowing them to accurately identify speech sounds that might otherwise have been ambiguous (see Samuel & Kraljic, 2009 for a review) and syntactic adaptations facilitate syntactic parsing and meaning interpretation, hence improving comprehension fluency (see Pickering & Ferreira, 2008, for a review). However, it remains relatively unexplored whether people also retune their lexical–semantic representations to linguistic input in adulthood. The current experiments explore the extent to which an adult’s lexical–semantic representations can be reshaped by recent exposure, and the time-course of these effects in different adult age groups.

The ability to adapt to interlocutors’ linguistic representations is argued to result from automatic alignment at different linguistic levels during language communication: interlocutors need to align their language systems in order to build up a common ground, against which they can accurately interpret words from each other’s speech and meanings from each other’s words and sentences (e.g., Pickering & Garrod, 2004). For instance, it has been observed that people assimilate each other’s accent, speech rate and articulation during dialogue (Giles, Coupland, & Coupland, 1991; Pardo, 2006); they gradually converge on the same terms (e.g., “couch” or “sofa”) to refer to an object (Clark & Wilkes-Gibbs, 1986); and they tend to

repeat each other’s syntactic structure in their utterances (Branigan, Pickering, & Cleland, 2000; Cai, Pickering, & Branigan, 2012). Though explicit attention may help interlocutors to align, most often, speakers and listeners can align by implicitly learning from each other’s linguistic input (Pickering & Garrod, 2004). Thus, it may follow that listeners may as well adapt their lexical–semantic representations with their conversational partners or according to their most recent experience. One typical example where such lexical–semantic alignment may occur is the interpretation of ambiguous words such as “gas” that have multiple meanings: successful communication would require a listener, for instance, to understand “gas” as referring to fuel when it is used by an American English speaker but as referring to an air-like fluid if it is used by a British English speaker (see Cai et al., 2015, for a demonstration). Similarly, if a word such as “pitch” has been previously used to mean the playing field (e.g., by a footballer/soccer player), or the throwing of a ball (e.g., by a baseball player) or as a musical-acoustic property (e.g., by a musician), a competent listener might take such recent experience into account when later interpreting the word in the same conversation in order to be able to process such word more accurately and fluently.

We therefore propose that the lexical–semantic representations of adults remain sufficiently malleable that our linguistic experiences can influence our interpretations of such words at a range of different time-scales. First, we suggest that our interpretations of such words are strongly influenced by our most recent experience, within the last few minutes. For example, a subordinate (low-frequency) meaning (e.g., the animal enclosure meaning of “pen”) might be relatively difficult to process the first time it was encountered during a conversation, but that this initial encounter would boost its subsequent availability making it easier to process later in the conversation. On the assumption that word-meanings are likely to be repeated within natural conversations, this would facilitate access of appropriate word meanings, relative to the situation where these meaning preferences were highly stable and only reflect the overall frequency of the meanings across the listener’s whole lifetime. Specifically, listeners/readers would be (i) faster to select the correct meaning on the basis of sentence context and (ii) less likely to assign an incorrect meaning. Given the ubiquity of lexical/semantic ambiguity in language, such a mechanism could potentially make a considerable contribution to the fluency of natural language comprehension. Such an effect would be akin to the processing enhancement seen for word recognition as a result of repetition priming, where words that have been most recently encountered are more easily recognised (See Wagenmakers, Zeelenberg, & Raaijmakers, 2000, for discussion of relationship between word frequency effects and repetition priming).

In addition, we propose that our linguistic experience can shape our lexical–semantic representations at longer time-scales, such that following repeated encounters with a particular word meaning over a period of days/weeks/years would increase the accessibility of that meaning relative to the word’s other meanings. This form of implicit learning would allow us to adapt to the changes in

linguistic environments that would occur when, for example, we move geographical location or take up a new hobby. This would allow, for example, psychology researchers to change over time their representations of words like “significant” and “normal” as a consequence of their experience with statistics.

Despite the intuitive appeal of the claim that experiences at different timescales might shift our default interpretations of ambiguous words, there is currently little relevant evidence to support it. This is perhaps surprising given the very large body of work looking at how ambiguous words are processed and the importance given to these words by the field in terms both of the challenge they present to models of language comprehension and their utility for testing key theoretical claims (see [Twilley & Dixon, 2000](#); [Vitello & Rodd, 2015](#)).

Existing models of how such ambiguous words are processed have instead focused on specifying the nature of the processing mechanisms involved, and have converged on the view that the alternative meanings of a word are initially retrieved in parallel, with the meaning that is most compatible with the sentence context then being rapidly selected (see [Duffy, Morris, & Rayner, 1988](#), for the influential re-ordered access model of semantic disambiguation, and [Twilley & Dixon, 2000](#), for a comprehensive review). The field has also converged on the view that the relative frequencies, also known as dominance, of the alternative meanings play a key role in these processes, such that a word’s most frequent (dominant) meaning is relatively easy to process compared with any lower-frequency (subordinate) meanings. For example, when an ambiguous word such as “spade” occurs in a constraining sentence context (e.g., “The gambler/gardener picked up a SPADE”), reading times are only longer when a relatively low-frequency card-related meaning is used (e.g., [Duffy et al., 1988](#)). In addition, when such words are preceded by a neutral context, most readers will by default assign the dominant meaning such that the reader/listener will subsequently need to engage in a cognitively demanding reinterpretation process if the subordinate meanings turns out to be required (e.g., “He picked up the SPADE but was hoping for the ace of diamonds”; e.g., [Rayner & Duffy, 1986](#); [Rodd, Johnsrude, & Davis, 2010](#)). It is striking that the field has consistently assumed that dominance, often explicitly defined as ‘meaning frequency’, is the key determinant of accessibility, implying that it is the overall frequency with which a meaning has been encountered across our lifetime that is the key determiner of meaning accessibility. Although these models allow that the semantic context in which a word occurs can influence availability, they allow no specific role for recent experience with a particular word in determining its accessibility. Thus, according to current models, the card-related meaning of “spade” would be more accessible within a conversation about card-playing, but there would be no additional benefit in the case where the listener had encountered the specific word “spade” within this conversation.

Consistent with this emphasis on meaning frequency as the key determinant of meaning availability, until very recently there was no strong evidence to support the idea that experiences with ambiguous words at relatively short

timescales might also be important. Some evidence that processing of a word meaning is facilitated when it is encountered for the second time (relative to its other non-repeated meaning) comes from a set of recognition memory experiments ([Light & Carter-Sobell, 1970](#)), which show better memory for an ambiguous word (e.g., JAM), when it is presented with adjectives that disambiguate it towards the same meaning on both the study and the test trials (e.g., RASPBERRY JAM–STRAWBERRY JAM) compared with a condition where the meaning changes between study and test (e.g., RASPBERRY JAM–TRAFFIC JAM). However because participants were aware of the subsequent memory test at the time of study, it is uncertain whether such effects of exposure to a particular word meaning would be seen in the absence of a deliberate attempt to remember the ambiguous target words. Stronger evidence on this issue comes from [Binder and Morris \(1995\)](#), who found that the second encounter with an ambiguous word within a paragraph of text was easier to process when the same meaning was used in both occasions, compared to when there was a switch in meaning. In addition there have been several reports that lexical decisions to words that are preceded by an ambiguous word (e.g., “bank-money”) are faster when participants had previously encountered a word pair that used the same meaning of the ambiguous word compared to trials using a different meaning (e.g., responses were faster following “bank-save” than following “bank-stream”), indicating that on the second presentation they were biased to retrieve the previously primed meaning ([Copland, 2006](#); [Simpson & Kang, 1994](#); [Simpson & Kellas, 1989](#); see [Masson & Freedman, 1990](#); [Bainbridge, Lewandowsky, & Kirsner, 1993](#), for similar demonstrations). However, in all these experiments, priming was observed at a relatively short delay (i.e. within a paragraph of text or a short block of isolated words), and critically, it is possible that such priming simply reflects a form of semantic priming rather than a specific change in how a particular ambiguous word should be interpreted. In other words the increased availability of the primed meaning may result from a more general boost for *all* words that fit with the preceding semantic context, not a specific shift in how an individual word should be interpreted.

The most direct evidence in support of this hypothesis that recent experience has a significant impact on meaning preferences comes from experiments using a novel ‘word-meaning priming’ paradigm ([Rodd, Lopez Cutrin, Kirsch, Millar, & Davis, 2013](#)). These experiments began with an initial prime phase in which participants heard sentences that contained fully disambiguated ambiguous words (e.g., “The footballers were greeted warmly by the adoring FANS”). About 20 min later, following an unrelated filler task (digit span), participants then heard an ambiguous word (e.g., “FAN”) for a second time, without a sentence context, and made a word association response (e.g., SPORTS) to the ambiguous word, a task that indicates which meaning of the ambiguous word participants retrieved first. The results showed that, when an ambiguous word had been disambiguated towards a particular meaning about 20 min earlier, participants gave 30–40% more responses consistent with the primed meaning. For

example, in the primed condition they were more likely to respond to “FAN” with “sports” or “tennis” and less likely to respond with associates like “cooling” or “summer” that relate to its alternative meaning, indicating that the primed meaning had become more accessible.

It is important to emphasise that word-meaning priming does *not* correspond to a form of semantic priming – at this relatively long delay (20 min) there was no impact on how “FAN” was interpreted in the word association task following a semantic priming control sentence (e.g., when the word “SUPPORTER” was used in place of “FAN” in the prime sentence). The absence of semantic priming indicates that this phenomenon of ‘word-meaning’ priming reflects the modulation of the accessibility of a specific word meaning, most likely by direct changes to the connections between a word’s input (phonological) representation and one of its semantic representations, such that each encounter with a particular word meaning strengthens the connection between its word form and this meaning, making the primed meaning for that specific word more readily available for subsequent processing even after a relatively long delay (Rodd et al., 2013).

This word-meaning priming effect was modulated by the baseline dominance of the meaning such that strongly subordinate meanings showed particularly strong changes. For example, very strongly subordinate meanings showed a fivefold increase in the likelihood of being retrieved after a single encounter during the prime phase (e.g., from 2% to 10%; see Rodd et al., 2013; Fig. 1b). Dominant word meanings, in contrast, showed little change in accessibility as a consequence of priming. Intermediate between these two extremes, Rodd et al. (2013) observed that priming can act to change listeners’ preferences: word meanings that were (on average) moderately subordinate (average dominance of 26%) can be boosted by a single previous exposure to such an extent that they come close to being equally preferred to the alternative meaning (average dominance of 40%). Although the consequence of this priming has not yet been measured for online measures of reading fluency, these priming effects are of a magnitude that are likely to have a marked impact on how ambiguous words are processed within sentence contexts. For example, according to current models of semantic disambiguation (e.g., the reordered access model; Duffy, Kambe, & Rayner, 2001; Duffy et al., 1988), readers have particular difficulty accessing the subordinate meaning of an ambiguous word even when the word occurs in a context that biases towards the subordinate meaning, presumably because readers have to engage in a time-consuming competition process with the contextually inappropriate dominant meaning. In contrast, readers have little difficulty in processing a word with relatively balanced alternative meanings regardless of which meaning the context biases towards because both meanings are somewhat equally accessible. As a boost in dominance via recent exposure can render a previously subordinate meaning to behave as a balanced meaning, therefore facilitating semantic interpretation when the same meaning is intended (as more often than not) in subsequent text/speech, word-meaning priming may serve a crucial function of improving comprehension fluency in reading/communication.

However, it should be noted that, while the word-meaning priming experiments reported in Rodd et al. (2013) provide clear evidence that recent experience has significant impact on the accessibility of word meanings, several key empirical questions remain.

First, unlike repetition priming effects in word recognition, which have been studied at a range of different delay intervals, and have been shown to persist on some word recognition tasks for delays of up to 8 days (see Woltz & Shute, 1995), all we currently know about the time-course of word-meaning priming is that it can be observed up to a delay of 20 min (Rodd et al., 2013). In addition, while the observed priming effect was numerically larger after a 3-min delay compared with a 20-min delay, this difference in priming magnitude was not significant (Rodd et al., 2013). Thus it is unclear how the priming effect changes over time. One key aim of the present research is to map out the time-course of word meaning priming. One possibility is that this phenomenon is relatively short-lived and does not significantly endure beyond the 20-min time window that has been observed to date (Rodd et al., 2013). This would suggest that the effect is being driven by short-lived changes to the accessibility of an individual meaning, and that these short-term changes have little or no consequence for the more stable meaning preferences (i.e. dominance) that endure beyond this time window. An alternative outcome is that word-meaning priming effects may persist significantly beyond 20 min, with the impact of encountering an individual word meaning having impact on preferences several hours, or perhaps even days/weeks/months, later. Such a finding would have important consequences for our understanding of the relative contributions of different types of experiences to our ability to understand these words under natural listening conditions and would impact on theoretical claims about the degree of plasticity within the adult lexicon. This finding would also place important theoretical constraints on the underlying psychological mechanism: if word-meaning priming persists for hours or days, then it is unlikely to reflect transient changes to accessibility due to residual activation of a lexical–semantic representation of the type thought to underlie semantic priming, but would more likely reflect substantive long-term changes to the connectivity within the lexical semantic network (see Rodd et al., 2013; see also Bock & Griffin, 2000, for a similar argument concerning structural priming). Finally, according to the most extreme version of this account, our overall preferences for words meanings may be *entirely* governed by our relatively recent experiences, with longer-term experience having *no* impact on current preferences. Under this extreme view, a dominant meaning (e.g., the writing implement meaning of “PEN”) would (on average) be more accessible than its subordinate counterpart, not because participants have all encountered it more often over their lifetime, but because, on average across participants, it is more likely to have been encountered within the preceding hours/days. Distinguishing between these possible accounts of how word meanings preferences arise, requires us to map out the impact of experiences with word meanings over a range of different timescales ranging from minutes (Experiments 1 & 2) to years (Experiments 3 & 4).

A second empirical issue to be explored is whether word-meaning priming might be modulated by a participant's age. The general issue of how comprehension processes change with age has recently been of considerable interest to researchers (see [Burke & Shafto, 2008](#), for a review), but studies of aging have tended to focus on how the processing mechanisms themselves change with age (e.g., age-related slowing of processing), and not on the nature of the representations that are used by comprehenders of different ages and how these might be differently affected by experience (but see [Ramscar, Hendrix, Shaoul, Milin, & Baayen, 2014](#), for a notable exception). Although some previous work has suggested that older adults can become impaired in their ability to rapidly use sentence context to modulate the processing of ambiguous words ([Dagerman, MacDonald, & Harm, 2006](#)) and, more generally, can become impaired on the ability to make use of sentential context to guide semantic processing ([Federmeier & Kutas, 2005](#)), it is currently unknown whether there are age-related changes in the abilities of older adults to use experience to facilitate processing of ambiguous words.

If it is indeed the case that our current preferences for particular word meanings reflect the combination of recent and long-term experience, then we might expect effects of recent experience to be reduced in older participants, who have more extensive lexical experience with these words, and so might plausibly place less weight on more recent encounters. This prediction of reduced word-meaning priming in older individuals might also arise from a general claim that older adults have reduced sensitivity to linguistic experience. For example, perceptual learning studies have shown that although older adults can show perceptual learning on a range of different experimental paradigms (e.g., [Golomb, Peelle, & Wingfield, 2007](#)), they can differ to younger adults in terms of (i) the magnitude of perceptual learning ([Scharenborg & Janse, 2013](#)), (ii) the degree of benefit from extended training ([Adank & Janse, 2010](#); [Peelle & Wingfield, 2005](#)), and (iii) the level of transfer of this learning to novel stimuli ([Peelle & Wingfield, 2005](#)). There is also evidence of reduced repetition priming in older adults compared with young adults ([Davis et al., 1990](#); [Fleischman et al., 1999](#); [Meiran & Jelicic, 1995](#)). Thus a reduction in word-meaning priming in older listeners could be predicted on the basis that this group may have a relatively pervasive reduction in the efficiency of the learning/priming mechanisms that are more efficiently recruited by younger listeners as a way of optimising their comprehension on the basis of recent experiences.

A final area of uncertainty, with respect to the mechanism underlying word-meaning priming, is the role of episodic memory. Thus far, we have described word-meaning priming as reflecting changes to stored lexical-semantic representations. However, an alternative explanation is that the underlying lexical-semantic representations remain unchanged after an encounter with an ambiguous word, and that the previously observed priming effect reflects a (short-lived) influence of episodic representations of the prime sentences. Under this view, when participants hear an ambiguous word in the word association task they recall episodic information associated with their earlier encounter with the relevant prime sentence and

this biases them to retrieve information related to the previously encountered meaning. Under this account, the priming effect could potentially lie outside of the mental lexicon, allowing our model of lexical processing to retain the idea that the underlying meaning preferences as stored in the lexicon are highly stable and *not* susceptible to large changes on the basis of recent experience. This episodic account was, to some extent, ruled out by [Rodd et al. \(2013; Experiment 2\)](#). Rodd et al. argued that if episodic representations were driving priming then the strength of priming should be modulated by the degree of perceptual overlap between the training and test exemplars. In particular, priming should be maximal when the training sentences and test words were spoken by the same person, and be reduced when there is clear mismatch between the speakers (i.e. female vs. male; see [Luce & Lyons, 1998](#); [Jackson & Morton, 1984](#), for similar arguments with respect to repetition priming). [Rodd et al. \(2013\)](#) found no significant effect of voice change, with comparable priming effects in both same-voice and different-voice conditions, suggesting that episodic factors are *not* critical. This null effect of speaker identity was consistent with the view that word-meaning priming effect reflects a modulation of the connection between a relatively abstract phonological representation of its form to its meaning as, according to most models of speech comprehension, this type of perceptual detail is not preserved at these relatively abstract levels of representation ([Luce & Lyons, 1998](#); [Orfanidou, Davis, Ford, & Marslen-Wilson, 2011](#)).

However, this finding needs revisiting for two reasons. First, given the importance of the finding from a theoretical point of view, it is important to replicate what is essentially a null result. Second, we must rule out the possibility that this observed null result was a consequence of listeners not retaining information about the identity of these unfamiliar speakers within their episodic representations of the prime sentences. In Experiment 1, we address this issue by making use of familiar speakers, for which it is more likely that this information will be both salient at the time of exposure, and subsequently retained.

The current experiments address these three critical issues: the time-course of priming, and the possible effects of speaker identity and listener age. Importantly, these experiments are, where possible, conducted in a more naturalistic listening environment compared with the earlier word-meaning priming experiments ([Rodd et al., 2013](#)). In Experiment 1 we examine the time-course of word meaning priming in a situation where participants obtain minimal experience with the ambiguous words: participants heard just one or two instances of the ambiguous word primes within a radio programme and then completed a web-based experiment to measure their meaning preferences for these ambiguous words. In Experiment 2 we then replicate some key findings from Experiment 1 within a more tightly controlled lab-based experiment. In Experiments 3 and 4 we explore the impact of more extensive (i.e. days/months/years), real-world experience with ambiguous words, by measuring how participants' recent and long-term participation in a particular sporting hobby (rowing) influenced their preferences for words used within this sporting context.

## Experiment 1

The current experiment aims to replicate the word-meaning priming effect, first shown by Rodd et al. (2013) in a more naturalistic listening environment. In addition, it addresses three key questions about word-meaning priming. First, we explore the time-course of word-meaning priming. Previous experiments (Rodd et al., 2013) tested word-meaning priming only at delays of both 3 and 20 min and found a non-significant numerical decline in priming between these time points. Here we measure priming in a much larger set of participants ( $N = 1800$ ) using prime–target delays ranging from a few minutes to several days in order to map out the time-course of this priming effect more comprehensively. Second, we explore the specificity of this form of priming – do listeners generalise more readily from an individual encounter with an ambiguous word to subsequent instances of the target word spoken by the same familiar speaker compared to instances where a different familiar speaker is used, or is this form of lexical–semantic priming immune to effects of speaker identity? Finally we explore the extent to which word-meaning priming is modulated by age.

These requirements for (i) a naturalistic listening environment, (ii) speakers who were highly familiar to participants, and (iii) a relatively long delay between prime and test are hard to achieve within conventional lab-based experiments. We therefore conducted the experiment in collaboration with BBC Radio 4, a British radio station. Specifically we worked with “The Human Zoo”, a scientific programme which covers a range of psychological topics and aims to “discover the way we think, behave and make decisions” (<http://www.bbc.co.uk/programmes/b01r6j16>). A key advantage of this approach is that the large number of participants that can be recruited compensates for the high level of noise in data from the word association task in which responses are influenced by numerous idiosyncratic aspects of a person’s individual experiences.

The prime phase of the experiment was conducted on live radio: listeners heard well-known radio presenters read short scripts that contained multiple, fully disambiguated ambiguous words. We then collected word association responses from listeners by asking them to log on to our website and complete a web-based word association task. By allowing listeners to complete this task at any time within the following week we could explore how their responses varied as a function of the delay between listening to the broadcast programme and completing the online task. Importantly, we used two familiar speakers during the prime phase and then systematically varied the voice used during the word association task in order to assess whether priming was reduced when there was a mismatch in speaker identity between prime and test.

### Materials and method

#### Participants

Participants were recruited live on air on the Radio 4 programme “The Human Zoo”, which was broadcast

throughout the UK. Listeners were invited to go to the programme website (<http://www.bbc.co.uk/programmes/b01r6j16>) and to follow the link to the experiment website. To recruit additional participants who had not listened to the programme, the link was also circulated via social media using the first author’s twitter account (@JenniRodd). The link to the experiment was active for seven days after the programme was broadcast. A total of 4035 participants clicked on the link that initiated the experiment. Of these, 2560 completed the experiment. Of those that did not complete the experiment, the majority (68%) dropped out very early in the experiment before the first task had begun, most likely because they did not have the facility to listen to sound or were not in a quiet listening environment. Only those participants who completed the entire experiment were included in the analysis. One participant was removed because they indicated that they were under 17, and 34 participants were removed because they answered “no” to the question “Is English the language that you use most often?”. Of the 2525 remaining participants (69.6% female; age 17–83; mean age = 52.0) 96.5% said that English was the language that they had learned first as a child, 95.0% said they had lived most of their life in the UK and 8.2% said that they considered themselves to be bilingual. Participants who took longer than 30 min to complete the experiment were then excluded (421 participants; 16.7%) as they were likely to have taken breaks during the experiment or have not fully engaged with the task. The remaining participants took an average of 24 min to complete the experiment (range: 8–30 min). We then excluded all participants who indicated that they had not listened to Episode 5 when it was broadcast but instead had listened via the BBC’s iPlayer service ( $N = 259$ , 12.3%) as we were not confident that they would have accurately reported the time at which they had listened to the programme. An additional 45 participants (2.4%) were excluded because they made more than 25% “error” responses on the meaning verification task, indicating that they are either mishearing these words or were somewhat uncertain about how to respond. After all these exclusion criteria had been applied 1800 participants were included in the final analysis.

#### Materials

The four prime paragraphs (38–58 words in length) were short descriptions of fictional characters containing seven fully disambiguated target ambiguous words that had at least two clearly distinct meanings (Appendix A). Where possible the lower frequency meaning was used (based on dominance scores from Rodd et al., 2013 and other pretests that used the same procedure). Half the target ambiguous words only appeared once in the script, while the other half appeared twice. Each participant heard all 28 ambiguous words without sentence context in the word association and meaning verification tasks (see procedure) together with one initial filler word that was not included in the analysis. All the paragraphs and isolated words were recorded by two different well-known journalists/presenters on BBC Radio 4 (Edward Stourton (ES) and Jenni Murray (JM)).

### Procedure

The prime paragraphs were broadcast within Episode 5 of the 2013 series of BBC Radio 4's "The Human Zoo". This episode focused on the influence of various factors in our environment on human decision making. There was no discussion of semantic ambiguity or any other topic related to this experiment. At the very end of the episode, the presenter (Michael Blastland) told the audience that he wanted them to participate in a memory experiment. They were instructed to listen carefully to some short clips and then to listen to the programme the following week to find out something interesting about their memory. The audience then heard each of the four prime paragraphs in turn, each spoken only once by one of the two well-known presenters in alternating order. The names of the presenters were not mentioned. The audience were then requested by one of the authors (JMR) to go to the programme website to participate in an experiment. They were told that the results from this experiment would be presented the following week. The link between the prime paragraphs and the online experiment was not mentioned.

An online survey tool ([www.qualtrics.com](http://www.qualtrics.com)) was used to present the main experiment. After indicating their consent to take part in the experiment, participants were instructed to listen to a single word spoken by one of the authors (JMR) and were instructed to adjust the volume of their speakers/headphones until they could hear it clearly. If they indicated that they could not hear it clearly then the experiment ended. They then input their age and gender, and answered yes/no to a series of questions about their language background. They then took part in two tasks: (i) word association; (ii) meaning verification.

### Word association

Participants heard a list of words and were instructed after each word to type the first word that they associated with the word that they had just heard. The first word was a filler item, not included in the analysis, followed by each of the 28 primed ambiguous words presented in a different random order for each participant. Each participant heard all 29 words spoken in the same voice (i.e. either ES or JM, counterbalanced across participants such that the same number of participants took part in each version), such that half the words were presented in the same voice as in the prime, but each item appeared in both the same- and different-voice conditions, across participants. They were then asked to type in any thoughts that they had about the aim of the experiment, and were asked to reply "yes", "maybe" or "no" to the question "Did you recognise the speaker". For those participants who indicated "yes" or "maybe" they were asked to give any information that they could about the speaker and were asked to indicate "yes", "maybe" or "no" to the question "Do you know the name of this person?". Finally, all participants were asked to guess the name of the speaker. Before starting the next stage of the experiment, participants were asked whether they had listened to Episode 5 of the Human Zoo (i.e. the episode that contained the primes) (i) live as it was broadcast, (ii) at a later time via BBC's iPlayer service (which allows listeners to listen to programmes on demand), or (iii) not at all.

### Meaning verification

The aim of this task was for participants to code their own earlier word association responses according to which meaning had been retrieved. In conventional word association dominance tests (e.g., Twilley, Dixon, Taylor, & Clark, 1994) this coding phase is done by the experimenters, but this approach was impractical for the large number of participants tested here. For each of the 29 items participants heard the ambiguous word again and saw their own word association response together with two short definitions of the ambiguous word's meanings – the meaning used in the paragraph and its most frequent alternative meaning. They were instructed to select the meaning that they had in mind when they had made their association response. These definitions were presented in the same order for each item across participants, with the paragraph meaning being given first on half of the items. A third "other meaning" option was to be selected if they had been thinking of a different meaning. A final "error" option was to be selected if they had misheard the word or were unsure of how to respond.

### Results

#### Voice recognition

Each participant was asked if they recognised the speaker that they had heard during the test phase of the experiment. 33% of the 415 participants who listened to ES during the experiment indicated that they recognised his voice. Within this subset 29% correctly identified ES, 36% indicated that they knew it was a Radio 4 presenter from a news/current affairs programme but were either unable to recall his name or gave the name of a different Radio 4 presenter associated with similar current affairs programmes, 32% of this subset did not give a specific name but responded with a vague (but correct answer) such as "Radio 4 presenter", and the remaining 3% answered incorrectly, identifying the speaker as someone not associated with Radio 4 news programmes (e.g., the main presenter of the Human Zoo). 27% of the 344 of participants who listened to JM's voice indicated that they recognised her voice. Within this subset 72% correctly identified JM, 8% indicated that they knew it was a presenter from Women's Hour but were either unable to generate her name or gave the name of a different presenter from this programme, 11% of this subset did not give a specific name or responded with a vague (but correct answer) such as "Radio 4 presenter", and the remaining 8% answered incorrectly, identifying the speaker as someone not associated with Women's Hour (e.g., a Radio 4 newsreader or continuity announcer). These responses indicate that many (but not all) participants were sufficiently familiar with the talkers to be explicitly aware of their identity.

#### Word association/meaning verification

392 participants indicated that they had not listened to the programme and so served as an unprimed control group. The remaining 1408 indicated that they had listened to the programme live and were categorised in terms of which day they had completed the experiment. 5 am was used as the cut-off, such that anyone who started

the experiment before 5 am on the morning after the programme was broadcast was considered to have completed it on ‘Day 1’, whereas participants who started the experiment after 5 am were considered to have completed it on ‘Day 2’. The same 5 am cut-off was used for the boundaries between subsequent days. The majority of the participants who completed the experiment after listening to the programme, completed it on Day 1 ( $N = 974$ ; 69.2%). 289 participants completed it on Day 2 (20.5%), and 145 completed it on Days 3–8 (10.3%).

For each participant/item within each of these four groups we calculated the proportion of the total meaning verification responses on which they had indicated that they had retrieved the meaning of the ambiguous word that was used in the prime paragraphs (Fig. 1). The highest rate of primed responses was for the participants who completed the experiment on Day 1, whereas participants who performed the experiment on later days were similar to the unprimed control group (whose performance reflects the baseline (unprimed) dominance of these meanings). ANOVAs showed that the proportion of primed responses varied significantly across these four groups ( $F_1(3, 1792) = 8.7, p < .001, \eta_p^2 = .01$ ;  $F_2(3, 81) = 6.7, p < .001, \eta_p^2 = .20$ ).<sup>1</sup> Pairwise comparisons between each of these four groups (using a Bonferroni corrected significance level of  $p < .008$ ), confirmed that the unprimed condition only differed significantly from participants who listened to the programme on Day 1 ( $F_1(1, 1362) = 21.8, p < .001$ ;  $F_2(1, 27) = 18.0, p < .001$ ). The only difference between the different priming groups that was significant in both analyses was the contrast between the participants who listened on Day 1 and those that listened on Day 2 ( $F_1(1, 1259) = 7.5, p = .006$ ;  $F_2(1, 27) = 12.4, p = .002$ ). The difference between Days 1 and 3 only reached corrected significance only in the items analysis ( $F_1(1, 1115) = 5.1, p = .025$ ;  $F_2(1, 27) = 9.4, p = .005$ ). All other pairwise comparisons were not significant ( $p > .1$ ). In summary, these initial analyses show that there is significant priming on Day 1, but this priming effect is significantly reduced by Day 2 and is not significant on this (and later) days.

Given that significant priming effects were restricted to participants who completed the experiment on Day 1, we then explored the performance of this group ( $N = 973$ ) in more detail using a simultaneous multiple regression approach.<sup>2</sup> The two predictors of interest were (i) the delay, measured in minutes, between the prime paragraphs and the onset of the experiment and (ii) participant age. The prime-test delay was estimated for each participant as the difference between the midpoint of the prime paragraphs and the time at which they started the experiment (as measured automatically by the survey software). Given the mean duration of the online experiment (24 min) and the assumption (based on experience with these tasks within the lab) that participants spent approximately half this time

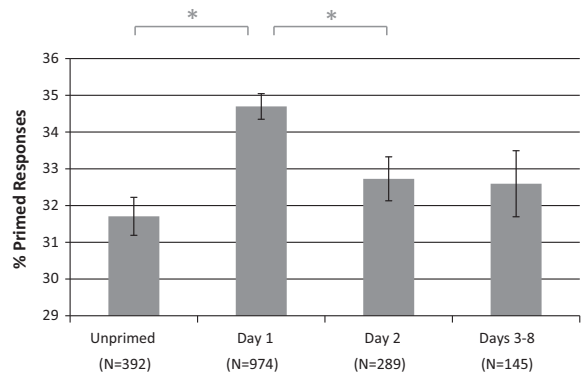


Fig. 1. Experiment 1: Proportion of trials on which participants generated a word association response that was related to the primed meaning.

on the word association task and half on the verification task, we estimated the average midpoint of the word association task relative to the start of the experiment as 6 min, which was then added to the delay estimates for all participants (see Rodd et al., 2013, for similar approach).<sup>3</sup>

To make the results easier to interpret, we subtracted from all responses the baseline rate of responding from the unprimed control group (31.7%), such that the dependent measure was the priming effect. A positive number reflected an increased probability of retrieving a primed meaning. As we did not have strong predictions about the precise nature of the relationships between delay ( $t$ ) and the proportion of primed meanings that were recalled ( $P$ ), we assessed several alternative models: (i) a linear relationship ( $P = b_0 + b_1 t$ ); a logarithmic function ( $P = b_0 + b_1 \ln(t)$ ); an inverse function ( $P = b_0 + b_1/t$ ); a power function ( $P = b_0 + t^{b_1}$ ); an exponential function ( $P = b_0 + e^{b_1 t}$ ). Although all of these functions were significantly correlated with the proportion of primed responses (all  $ps < .008$ ), the best model fit was provided by the logarithmic model ( $P = 5.25 + (0.83 * \ln(t))$ ;  $R^2 = .021, F = 20.8, p < .001$ ).<sup>4</sup> A stepwise regression approach confirmed that none of the alternative functions made a significant additional contribution above the variance explained by the logarithmic function ( $ps > .18$ ). We then repeated these steps for the predictor age ( $A$ ). In this case the best fit was provided by the linear model ( $P = 7.60 - (0.087 * A)$ ). The combined regression model including both delay (log-transformed) and age (linear) confirmed the significant negative effects of both age ( $\beta = -.07, t = -2.3; p = .021$ ) and delay ( $\beta = -.13, t = -4.1; p < .001$ ), such that there was greater priming for younger participants and for participants who completed the experiment at a shorter delay<sup>5</sup> ( $P = 8.5 - (0.75 * \ln(t)) - (0.065 * A)$ ). Although these two

<sup>1</sup> Version was included in these (and where appropriate in subsequent analyses), but main effects and interactions involving version are not (Pollatsek & Well, 1995).

<sup>2</sup> One participant was removed from this analysis as they entered an age of over 500. Given their other responses this was taken to be a one-off typo and they were retained in all other analyses.

<sup>3</sup> An alternative approach would have been to estimate delay based on both the start time and end time of the experiment, but this would introduce a correlation between delay and the time-taken to complete the experiment, which could introduce artefactual effects of delay on performance if those participants who were slower on the task showed less priming.

<sup>4</sup> A series of  $F$ -tests showed no significant difference in the level of fit given by these different models ( $p > .4$ ).

<sup>5</sup> The interaction term between these two variables was non-significant ( $p = .8$ ) and so was removed from the regression.



predictor variables were significantly correlated with each other ( $r = .187, p < .001$ ) reflecting the fact that older participants tended to wait longer before starting the experiment, but there was no evidence of serious collinearity in these data (variable inflation factor (VIF) = 1.5; O'Brien, 2007). The nature of these relationships is shown in Fig. 2.

Finally, to assess whether there was a significant effect of voice-congruency (i.e., more priming for the same-voice condition compared with the different-voice condition), we calculated for each participant the difference between these two conditions, which we refer to as the same-voice benefit. Across this large set of participants the mean same-voice benefit was small and negative (mean =  $-0.5\%$ ;  $SD = 18.4$ ), indicating that there was no priming increase when the voice was kept constant between prime and test. This absence of an effect was confirmed by a multiple regression (including both age and log-transformed delay) using the same-voice benefit as the dependent measure. In this analysis both the constant and the main effect of delay (log transformed) were not significant ( $ps > .15$ ), confirming that there was no main effect of voice-congruency, or interaction between voice-congruency and delay. To explore the possibility that a voice-congruency effect might be modulated by the listeners' familiarity with the speakers, we repeated the above analysis using only those participants who answered "yes" to the question "Do you recognise the speaker"<sup>6</sup> ( $N = 352$ ). Again the overall mean 'benefit' in this group was small and negative (mean =  $-0.03\%$ ;  $SD = 18.7$ ) and the constant and effect of delay in the regression analysis were non-significant ( $ps > .7$ ).

#### Awareness of experiment aims

We manually coded the responses of the 973 participants included in the regression analysis. 58.5% either gave no response or explicitly stated that they had no idea of the aim, 28.0% gave a response unrelated to the true aim (e.g. suggesting that we were exploring effects of age, gender, culture, experiences, political affiliations or social class on how people interpret words, or suggesting that this was a study into techniques used in advertising, which was the subject of the Human Zoo episode they had just listened to. Only 13.6% of participants indicated an awareness of the link between some of the words and the prime paragraphs. The main regression analysis was repeated excluding the 'aware' participants. The effects of both delay (log-transformed) and age (linear) remained significant for this subset ( $\beta = -.08, t = -2.4; p = .017$ ) and delay ( $\beta = -.13, t = -3.9; p < .001$ ).

#### Discussion

These results from this experiment provide clear answers to all our initial research questions.

<sup>6</sup> Note that caution must be taken when selecting participants on their knowledge of the specific speakers as this can bias sampling of participants differently for the two speakers. However, in this case it seemed unlikely that groups selected on the basis of their knowledge of these two speakers would differ in their baseline dominance in a way that would be consistent across this set of 28 words, which were not selected to be closely linked to the likely topics that would be associated with these two speakers.

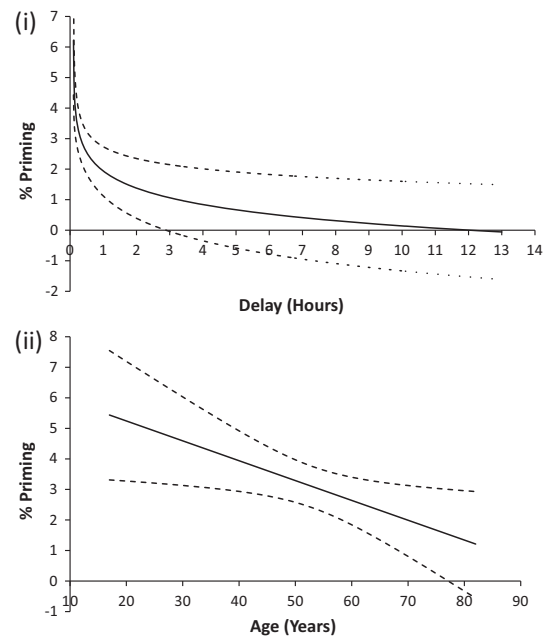


Fig. 2. Experiment 1: Predicted values from regression analysis (solid lines) as a function of variation in (i) delay between prime and test and (ii) participant age. In both cases these are plotted at median values of other predicted variables. Dotted lines represent the 95% confidence intervals for these predicted values.

Word-meaning priming was observed under relatively natural conditions and was significantly modulated by both time and age, such that it was enhanced at short durations and for younger participants. The effect of delay on priming was striking (see Fig. 2). For example, participants who completed the word association task after just 6 min had an estimated priming magnitude of 6.2% (see Fig. 2), reflecting an increase in the proportion of responses that were related to the primed meaning from a baseline rate of 31.7% to 37%. It is worth emphasizing that this absolute increase in primed responses of 6.2% reflects a *proportional* increase in the likelihood that a primed meaning is generated of 20%. This priming effect was strongly modulated by the delay between the prime and test, such that priming estimates were reduced to an absolute change of just 1% after a delay of about 3 h. The modulatory effect of age on priming magnitude was equally striking, such that at the median delay of 13 min, the estimated effect for participants aged 20 was 5.3%, but this was reduced to an estimated effect of just 2.0% for participants aged 70. In contrast to these clear effects of delay and age, word-meaning priming was *not* influenced by a change in speaker identity between prime and test: the mean priming for the same-voice condition was in fact numerically (though non-significantly) *less* than that seen in the different-voice condition.

These findings confirm that short-term priming effects are likely to contribute to listeners' ability to understand ambiguous words in context in the situation where they have already encountered that word within a conversation. The wider theoretical implications of these three key findings for current theories of language comprehension will be explored in detail in the General discussion.

However, while this experiment indicates that word-meaning priming decays relatively rapidly with time during the first hour after an encounter with an ambiguous word, one weakness of this experiment is that it used a between-participant design in which the experimenters had no control over the delay with which participants took part in the experiment. Thus it is possible that the individuals who chose to wait longer before completing the word association task differed systematically from those who participated more immediately. For example, the ‘long-delay’ participants might have been generally less interested in the experiment and their lower levels of priming could potentially reflect a less attentive listening attitude during the prime sentences. Given the non-significant effect of delay seen in the earlier lab-based study of word-meaning priming (Rodd et al., 2013), it is important to replicate this finding that priming decays with time during the first hour after exposure using a within-participant design in a more conventional lab-based experiment.

## Experiment 2

This experiment uses a modified version of the lab-based word-meaning priming paradigm introduced by Rodd et al. (2013) in which participants hear individual prime sentences that contain an ambiguous word that is disambiguated towards its subordinate meaning by the context (e.g., “The pig PEN was muddier than ever.”). The impact of these primes on the target ambiguous word (e.g., PEN) is measured on a subsequent word association task. In the experiments reported by Rodd et al. (2013) the prime sentences were presented in a separate block from the test words, resulting in a relatively long minimum time between each individual prime sentences and its corresponding word association trial (3 min). In addition, individual randomisations of both the prime sentences and the target words within these blocks resulted in variable delays between each individual prime sentence and its corresponding word association trial, depending on whether these occurred towards the beginning/end of these two blocks. In order to reduce the between-item variation in delay and to measure priming at a wider range of delays, Experiment 2 comprised three blocks of trials, which all contained both primes and targets, presented in alternation. In this modified version of the paradigm, participants only made responses to the isolated words, but were instructed to listen carefully to all the stimuli that they heard. Participants heard all prime sentences in Block 1, and the subsequent word association trials were positioned in either Block 1, 2 or 3 so that they appeared exactly either 1 min, 20 min or 40 min after their corresponding prime sentence. Between the blocks participants completed a non-linguistic filler task in order to allow for the appropriate delay between prime and test.

### Materials and methods

#### Participants

51 native speakers of British English were recruited from the University College London online recruiting

system and were each paid £6 for their participation. Nine participants were excluded due to technical problems (software crash or problems recording verbal responses). One participant was excluded for failing to adequately follow instructions and one was excluded for a high number of null responses on the word association task (49%).

#### Materials

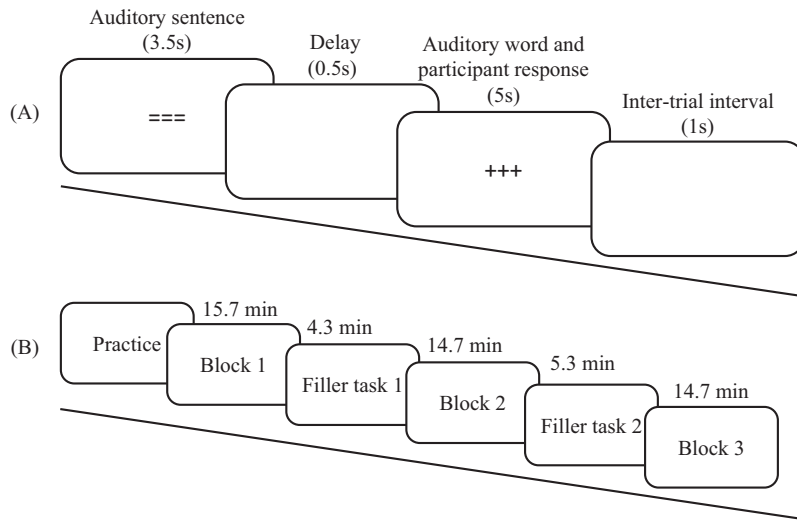
274 target words were selected for use in the word association task, of which 88 were ambiguous words (Appendix B) and the rest unambiguous filler words (e.g., bread, thief). The ambiguous words were all chosen to have a subordinate meaning (dominance range: 0–0.48; mean = 0.24) that was semantically distinct from the word’s dominant meaning (dominance range: 0.42–1; mean = 0.70). These dominance scores were taken from a variety of pretests that used the standard word association method (Twilley et al., 1994) and the same participant population as the main experiment. Most of the ambiguous words were homonyms, which share both spelling and pronunciation (e.g. “BARK”), but eight were non-homographic homophones, which share only pronunciation (e.g., “FLOUR/FLOWER”).

88 experimental prime sentences were constructed in which the initial part of the sentence strongly disambiguated the ambiguous word towards its subordinate meaning (e.g., “The author put his memos in the APPENDIX of the book”). An additional 166 filler sentences were constructed, which contained none of the ambiguous experimental words. Another four pairs of sentences and words were constructed as practice items. All these sentences and words were spoken by a female speaker with a Southern English accent (HB). Four lists of materials were created for use in constructing the four versions of the experiment (see Design).

#### Design

The experiment contained four conditions: three primed conditions that varied in the delay between the prime and target (1, 20 and 40 min) and an unprimed baseline. Four versions of the experiment were created with the 88 experimental items being assigned to the four priming conditions using a Latin-square design such that each word was assigned to each of the four conditions across the four versions. This ensured that each item occurred in every condition (across participants), that all participants contributed to all conditions, but that no participants encountered any item more than once.

The stimuli were presented in three blocks of trials. The first block contained all of the 88 experimental prime sentences in the same random order for all participants, except for the 22 sentences that were assigned to the no-prime condition for each version, which were replaced by filler sentences. The decision to keep the sentences for all primed conditions within the first block ensured that these prime sentences were equally well attended across conditions. These 88 sentences were followed by an additional 6 filler sentences at the end of the block, which were required to maintain the alternation of sentences and word association trials until all of the word association trials had been completed. The second and third blocks



**Fig. 3.** Experiment 2: Illustration of the trial structure and experiment composition. The upper panel presents the trial structure and the lower panel the experiment composition.

each contained 88 filler sentences in the same random order for all participants.

The word association target words were then positioned in such a way that a word always appeared exactly 1 min (i.e., 6 trials, one trial lasting for 10 s; see Procedure) later than its corresponding prime sentence if the word was assigned to the 1-min condition, 20 min (a block) later if it was assigned to the 20-min condition, and 40 min (2 blocks) later if it was assigned to the 40-min condition (see Fig. 3 for timing information). In other words, a target word occurred in Block 1 under 1-min condition, in Block 2 in the 20-min condition and in Block 3 in the 40-min condition. The target words in the no-prime condition were spread across the three blocks (7 in Blocks 1 and 2; 8 in Block 3). To ensure a strict alternation of sentence and word trials throughout the whole experiment, filler words were inserted into every slot within each block in which an experimental target had not already been positioned. Note that a target word never immediately followed its corresponding prime sentence.

### Procedure

Participants were tested individually in a cubicle and the experiment was run on E-Prime 2.0. After giving their informed consent, a participant began the experiment with a practice session. Each trial began with a screen with the symbol “===” lasting for 3.5 s, during which a sentence was played via headphones (Fig. 3a). The onset of the sound file was aligned to the start of the 3.5 s period. After a short delay (0.5 s) the symbol “+++” appeared for 5 s, during which participants heard a word to which they verbally gave an associate. The onset of the sound file was aligned to the start of the 5 s period. Participants were told to give the response as quickly as possible within the given time window and that responses beyond that time window would not be registered. The participants’ verbal responses were recorded as individual sound files. The screen was

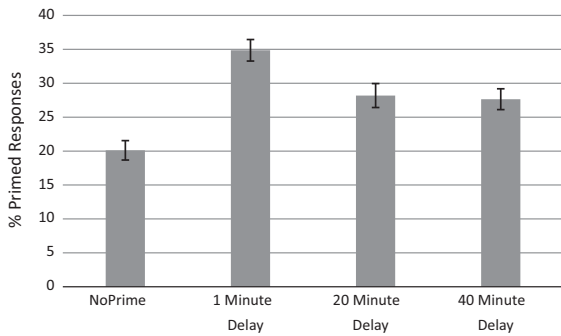
then replaced with a blank screen lasting for 1 s as an inter-trial interval. Thus, each trial lasted for 10 s in total.

As shown in Fig. 3b, the whole experiment consisted of a practice session, Block 1 lasting for 15.7 min (94 trials), Filler task 1 lasting for 4.3 min, Block 2 lasting for 14.7 min (88 trials), Filler task 2 lasting for 5.3 min, and finally Block 3 lasting for 14.7 min (88 trials). After the practice, all trials were presented according to a fixed time schedule that did not vary across participants. The filler tasks involved colouring pictures. Right before the end of a filler task, participants heard beep sounds via their headphones, at which point they were required to stop colouring and get ready for the next block of the experiment.

### Results

The word association responses of each participant were transcribed and coded, with each response coded as referring to either (i) the dominant meaning of the homophone, (ii) the subordinate meaning of the homophone that was used in the prime sentence, (iii) another meaning, or (iv) an error (e.g., the participant had misheard the word, their response could not be clearly heard, or it was unclear which meaning they had retrieved). The data were initially divided into two sets, each coded by a single experimenter (AA, HB). All these codes were then checked by a third experimenter (JMR). The coders were all blind to the prime condition that the word was assigned to.

For each participant/item within each of these four prime conditions, we calculated the proportion of trials on which they had retrieved the meaning of the ambiguous word that was used in the prime paragraphs as a proportion of the total number of non-error trials (Fig. 4). The highest rate of primed responses was for the 1-min condition. The 20-min and 40-min conditions showed a somewhat lower rate of primed responses, although this was still higher than that seen in the unprimed baseline. ANOVAs showed that the proportion of primed responses varied significantly across



**Fig. 4.** Experiment 2: The mean proportion of responses in each condition that matched the primed meaning for each participant. Error bars are  $\pm 1$  standard error away from the mean (subjects' analysis).

these four groups ( $F_1(3,108) = 16.1$ ,  $p < .001$ ,  $\eta_p^2 = 0.55$ ;  $F_2(3,252) = 14.5$ ,  $p < .001$ ,  $\eta_p^2 = 0.35$ ).<sup>7</sup> Pairwise comparisons among the four conditions (using a Bonferroni corrected significance level of  $p < .008$ ) confirmed that priming was present in all three prime conditions: compared to the unprimed condition, all the three priming conditions led to significantly more retrievals of the primed meanings (1-min condition:  $F_1(1,84) = 40.0$ ,  $p < .001$ ,  $\eta_p^2 = 0.53$ ;  $F_2(1,36) = 39.9$ ,  $p < .001$ ,  $\eta_p^2 = 0.32$ ; 20 min condition:  $F_1(1,84) = 18.8$ ,  $p < .001$ ,  $\eta_p^2 = 0.34$ ;  $F_2(1,36) = 13.7$ ,  $p < .001$ ,  $\eta_p^2 = 0.14$ ; 40-min condition:  $F_1(1,84) = 16.7$ ,  $p < .001$ ,  $\eta_p^2 = 0.32$ ;  $F_2(1,36) = 12.5$ ,  $p < .001$ ,  $\eta_p^2 = 0.14$ ). In addition, the reduction in priming between the 1-min condition and the 40-min condition was significant: ( $F_1(1,84) = 11.8$ ,  $p < .001$ ,  $\eta_p^2 = 0.25$ ;  $F_2(1,36) = 7.9$ ,  $p < .006$ ,  $\eta_p^2 = 0.09$ ). The reduction in priming between the 1-min condition and the 20 min condition was significant at corrected level of significance in the subjects analysis and at an uncorrected level in the items analysis ( $F_1(1,84) = 7.8$ ,  $p < .008$ ,  $\eta_p^2 = 0.32$ ;  $F_2(1,36) = 4.8$ ,  $p < .03$ ,  $\eta_p^2 = 0.06$ ). The difference between the 20- and 40-min conditions was not significant ( $F_1(1,84) = 0.1$ ,  $p = .73$ ,  $\eta_p^2 = 0.003$ ;  $F_2(1,36) = 0.42$ ,  $p = .52$ ,  $\eta_p^2 = 0.005$ ).

### Discussion

The results of this experiment replicate the reduction in word-meaning priming that was observed during the first hour after exposure in Experiment 1, under more tightly controlled lab-based conditions using a within-participants design: there were significantly fewer primed responses given after a 40 min delay (28%), compared to the 1-min delay condition (35%; Fig. 4). This suggests that the reduction in priming in Experiment 1 was not an artefact of the between-participant design and that word-meaning priming does decay within the first hour. However, despite this reduction, even after 40 min the primed condition was significantly different to the unprimed condition (20%) and the absence of a difference between the 20-min and 40-min conditions, which were numerically very similar (28.2%, 27.8%), indicates that priming is

relatively stable within this time window. This pattern of priming – a relatively rapid decay that occurs at some point within the first 20 min, followed by a more stable component – is broadly consistent with the logarithmic decay function that was found in Experiment 1. We return to the theoretical implications of this time-course in the General Discussion.

Despite the clarity of the results from these two experiments, several unanswered questions remain. In particular, while Experiments 1 and 2 have clearly shown that relatively recent experience can have a marked impact on listeners' preferences for a particular word meaning, Experiment 2 only studied the effect at a maximum delay of 40 min and Experiment 1 suggests that the impact of just one or two encounters with a word are relatively small beyond the first hour after exposure. Thus we have no evidence concerning how or when listeners make use of their own experiences with ambiguous words to alter their long-term preferences for their different meanings. We cannot yet be certain about how (and indeed if) these relatively short-term priming effects relate to longer-term changes in meaning preferences that endure for days, months, or even years. These longer-term changes in meaning preference will be explored in Experiments 3 and 4.

### Experiment 3

One experimental approach to the study of how preferences for word meanings change in the longer-term (i.e. over days/months/years) within a relatively naturalistic setting is to explore the preferences of participants with atypical linguistic experience, whose preferences for particular meanings are likely to differ from the population as a whole. Studies that explore the consequence of participants' specific types of exposure to language outside of the lab are relatively rare. One example of this approach is the study by Coane and Balota (2009) that shows that lexical decision latencies to words like "leprechaun" that are associated with a particular holiday are faster and more accurate when participants are tested around the time of year that the relevant holiday occurs. In addition, Coane and Balota (2011) have shown that semantic priming can be observed for pairs of items that will likely have co-occurred within participants recent linguistic experience (e.g., for words that occur within movie titles). Taken together these previous experiments indicate that particular words, and the connections between them, may become more accessible on the basis of real life experiences.

Here we explore whether particular word meanings also become more accessible. Specifically, we focus on recreational rowers who have learned new rowing-specific meanings for relatively common words. For example, a novice rower will quickly learn that the words "square" and "feather" refer to a position of the oar, while the word "crab" refers to an (often catastrophic) error in which their blade is pulled down into the water. We assume that exposure to the non-rowing meanings will be relatively constant across the group of rowers, and so any differences between the rowers' preferences will largely be driven by

<sup>7</sup> Version was included in these, but main effects and interactions involving version are not (Pollatsek & Well, 1995).

the nature of their rowing experience. We use a regression approach to determine which specific aspects of their rowing experience predict their overall preferences for these rowing-related meanings. In particular we assess whether these preferences are driven by (i) relatively recent rowing experience within the past week, (ii) medium-term experience over the previous weeks/months, or (iii) their longer-term experience that builds up over years. Finally we will, as in Experiment 1, assess whether age plays a critical role in modulating the effects of experience on lexical preferences.

The use of rowers as the target population has several advantages compared with other groups with specialised vocabulary. First, rowing has a relatively large set of these critical words which have both a relatively common non-rowing meaning and a meaning that is specific to rowing (e.g., “square”, “finish”, and “gate”). Also, unlike other more common sports (e.g., football), these rowing terms have not become part of everyday vocabulary. Coverage of this sport on television in the UK is minimal and so the majority of these meanings are unlikely to be encountered by either rowers or non-rowers outside of rowing training sessions. Therefore, by obtaining detailed information about the time spent rowing we can obtain a relatively accurate estimate of an individual’s exposure to these meanings. Finally, recreational rowers within any club vary substantially in their current level of training including individuals who row every day as well as those who only row once a week (or less). This provides the necessary variability to assess, using regression analyses, the impact of these different types of linguistic experience.

## Materials and method

### Participants

An email was sent to the captains of six rowing clubs in Cambridge, England with which the first author (JMR) had a previous connection, offering to pay the club £100 if 30 of their members completed the web-based experiment, with an additional bonus of £25 if 40 members took part. Three of these clubs responded and were sent a web link to distribute to their members. One club was only for university students, the other two were not associated with the university. Of the 119 rowers who took part, 18 were excluded from the analyses as they did not complete the experiment; a further 14 were excluded because they indicated that they were not fluent in English or were not native English speakers. Of the remaining 87 rowers (61% female; aged 17–55; mean age = 28.3 years,  $SD = 8.9$  years), 12.6% considered themselves to be bilingual and 85.1% were born in the UK.

Twenty-seven control participants with no rowing experience were recruited (56% female; aged 19–48; mean age = 30.1 years,  $SD = 9.2$  years) via social networks. They did not receive reward for their participation. 14.8% of the controls considered themselves to be bilingual and 96.3% were born in the UK.

### Materials

The 21 target ambiguous words (Appendix C) were selected by one of the authors (JMR) who has 10 years of rowing experience. The words have one dominant

non-rowing meaning that would be familiar to all participants plus an additional meaning that would be frequently encountered within a recreational rowing environment. Nineteen of the rowing-related words would be unlikely to be encountered outside of a rowing context, but “bow” and “stern”, which refer more generally to a part of a boat, are also used outside of a rowing context. These were included to maximise the number of stimuli with the intention to remove them if the boat-related meanings were very frequently generated by the non-rowing control participants. An additional 80 filler words were selected to have no association with rowing and to distract from the primary aim of the experiment.

### Procedure

An online survey tool ([www.surveymonkey.co.uk](http://www.surveymonkey.co.uk)) was used to present the experiment. Participants were told that the aim of the research was to discover the impact of exercise on memory. This provided a plausible cover story for the specific recruitment of rowers without making them aware that they were being targeted specifically because of their rowing experience. After indicating their consent to participation, they provided basic demographic information (age and gender), information about their language background and about how often they exercised (to maintain the cover story). They then completed the main experimental task: word association.

Participants made word-associations to 101 words: 21 homographs with a rowing-related meaning and 80 fillers (see Materials). Each word was visually presented, one at a time, and they were instructed to type in the first word that was related in meaning to the target word that came to their mind. Due to the relatively large set of words to be included in this experiment, and the corresponding increase in experiment length, we did not include the meaning verification task used in Experiment 1.

To maintain the cover story, participants were given a short recognition memory test which consisted of 24 words; twelve of which had been in the previous word-association task. Participants were then asked to state what their main sport was and to answer a series of questions about their participation in this sport. It was assumed that the majority of participants would respond that their main sport was rowing and so this question aimed to gain information about their rowing experience without revealing that that we were specifically recruiting rowers. They were asked (i) how long they had participated in this sport (years); (ii) how many days since they had last participated in this sport; (iii) on average how many times per week they had participated in this sport over the last month. Participants were then asked what they thought the aim of the study was. Finally, for each of the 21 rowing-related words, the rowers (but not the non-rowing controls) were asked to answer YES/NO to indicate whether they had heard the word in their sport and knew its meaning.

## Results

### Knowledge of rowing words

These data were looked at prior to the analysis of the word association responses in order to potentially exclude

**Table 1**  
Descriptive statistics for predictor variables in multiple regression analyses.

	Age	Long-term Number of years rowing experience	Medium-term Average weekly frequency (last month)	Medium-term Average weekly frequency (last 9 months)	Recent Number of days since last rowed	Recent Rowing frequency in previous week
<i>Experiment 3</i>						
Min	17	0	0.5		0	
Max	55	42	20		21	
Mean	27.5	5.9	6.0		2.2	
SD	8.1	6.3	4.0		3.8	
<i>Experiment 4</i>						
Min	18	0	0	0	0	0
Max	76	60	10	7.8	n/a	8
Mean	34.8	8.9	3.2	2.6	3.9	1.6
SD	16.1	13.1	2.1	1.9	3.1	2.0

responses to words or participants with low rates of knowledge. On average the rowers indicated that they knew 93.5% of the words. Each rower knew more than 60% of the rowing words. Each rowing word was known by more than 78% of the participants. No items/rowers were excluded entirely from the word association analysis (see following section for removal of items unknown to individuals).

#### Word association

A semi-automated approach was used to determine whether each word association response should be classified as 'rowing' or 'non-rowing'. In the first instance, one of the authors (BH) manually checked each response and highlighted any responses that seemed to indicate the retrieval of the rowing meaning. For example, for the ambiguous word "square", the responses "oar" and "blade" were classified as a rowing response, whereas "rectangle" and "circle" were not. Every such rowing-related response was compiled into list of likely rowing response such that any instances of the word "oar" for any other participant/item would then automatically be coded as a rowing response. This semi-automated approach was used to increase the consistency of coding across items/participants. All responses that were automatically coded as being a 'rowing response' were then manually checked and in a few cases these decisions were overridden. For example, while the response "race" was considered to be a rowing response to the ambiguous target "bump" (because the "bumps" are a type of rowing race), it was not considered a rowing response in response to the ambiguous target "finish", because the rowing meaning of this word is to do with the part of the rowing stroke and not to do with racing. Where there was any uncertainty in the coding this was discussed with a second author (JMR) who had extensive rowing experience. In any cases where coding was uncertain the default was to code as 'non-rowing'. All coding was done blind to whether the data came from a rower or control.

For the rowers, these word association data were then combined with the data from the rowing knowledge test such that we removed any word association response to a rowing word where that rower had indicated that they did not know the rowing-related meaning. We then calculated (across both participants and items) the proportion of

these 'known meaning' trials on which a rowing related meaning was generated. As expected, the mean proportion of rowing responses was substantially higher in the rowers ( $M = 29.3\%$ ;  $SD = 17.3$ ) compared with the controls (mean = 0.7%;  $SD = 1.7$ ). The non-zero score in the controls reflects the fact four of the controls gave a boat-related meaning for "stern". Due to the differences in variability in responses in the two groups, non-parametric tests were used to confirm the significance of this difference ( $U_1(114) = 25.5$ ,  $Z_1 = 7.7$ ,  $p < .001$ ;  $Z_1(21) = 4.0$ ,  $p < .001$ ). These initial results confirm that the rowing meanings were sufficiently dominant within the rowers that on a significant proportion of trials (29.3%) the rowing meaning was the first meaning to be retrieved, despite the fact that these meanings are strongly subordinate or unknown for non-rowing participants.

A simultaneous multiple regression analysis ( $N = 79$ )<sup>8</sup> then explored the influence of a participant's age and their rowing experience on the word association performance. We included a measure of long-term experience (number of years of rowing experience), a measure of medium-term experience (average weekly rowing frequency in previous month), and a measure of recent experience (number of days since they last rowed). Initial exploration of the data confirmed that the participants sampled varied substantially on all four predictor variables (see Table 1).

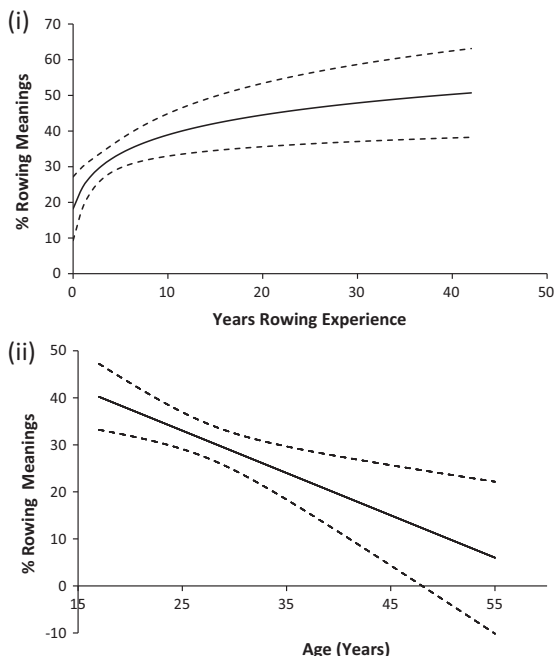
On the basis of Experiment 1, all three measures of experience were log-transformed. The inter-correlation matrix with all predictor variables showed a significant positive correlation between age and both an individual's long-term rowing experience ( $r = .464$ ,  $p < .001$ ) and a negative effect their medium-term rowing frequency ( $r = -.432$ ,  $p < .001$ ), indicating that older rowers tend to have rowed for longer but have rowed less over the past months. Medium-term rowing frequency was also significantly negatively correlated with how long since they had rowed ( $r = -.361$ ,  $p = .001$ ; more rowing by those who had rowed more recently). As with Experiment 1, there was no evidence of serious multi-collinearity in this data (maximum VIF = 1.61).

<sup>8</sup> Eight of the rowers were removed from this regression analysis because they gave a non-rowing sport as their main sport and so we had no information about their rowing experience.

The overall regression model was significant ( $F(4,74) = 3.7, p = .01, R^2 = 0.17$ ). Age was a significant negative predictor ( $\beta = -.42, t(74) = -3.1, p = .002$ ) such that older individuals tended to retrieve a lower proportion of rowing responses, whereas length of rowing experience had a significant positive influence on the proportion of rowing responses ( $\beta = .39, t(74) = 3.2, p = .002$ ) such that rowers with more overall rowing experience tended to provide a higher proportion of rowing responses. The measure of medium term rowing experience (rowing frequency over the past month) was only marginally significant ( $\beta = -.25, t(74) = -2.0, p = .053$ ), and surprisingly the measure of recent experience (days since they had rowed) was a non-significant predictor ( $\beta = -.14, t(74) = -1.2, p = .23$ ). The nature of the predicted effects of age and long-term rowing experience are shown in Fig. 5.

#### Awareness of experiment aims

Most participants indicated that they believed the cover story (46%) or gave either no answer or a very vague answer to the question about the experiment aim (39%). Only 12 of the 79 rowers included in the regression analysis (15%) indicated that they thought the aim of the experiment was related in any way to the presence of sports-related words. The multiple regression analysis reported above was repeated without these 12 participants. The significance levels were not changed by the exclusion of these participants except that the marginal effect of medium-term experience became significant ( $\beta = -.32, t(62) = -2.2, p = .035$ ).



**Fig. 5.** Experiment 3: Predicted values from regression analysis (solid lines) as a function of variation in (i) number of years rowing experience and (ii) participant age. In both cases these are plotted at median values of other predicted variables. Dotted lines represent the 95% confidence intervals for these predicted values.

#### Discussion

Several intriguing findings have emerged from this experiment, which suggests that age and long-term rowing experience have the most salient effects on how readily available the rowing meanings are to any given individual: rowing meanings were more likely to be generated by rowers who had rowed for a long time and by those who had gained this experience at a relatively young age. Importantly, although these two predictor variables are positively correlated (older rowers have typically rowed for longer), their effects are in the opposite directions – whilst older rowers tend (on average) to have rowed for longer, they give *fewer* rowing responses than younger rowers with similar amounts of experience. In contrast the effect of medium-term experience was only marginal, and the effect of recent experience was not significant.

One important issue is the extent to which these results could be explained in terms of demand characteristics – is it possible that the effects reflect the awareness of participants regarding the aim of the experiment and so have modified their behaviour in order to conform to the experimenter's expectations? There are two reasons that this is unlikely. First, the majority of participants reported no awareness of the experiment aims, suggesting that the cover story (that indicated they had been being recruited for a more general study on the effect of exercise on memory) had been successful. In addition, the high proportion of non-rowing fillers (80%) may have helped to distract from the true aim of the experiment. On average each rower only retrieved six rowing related meanings in response to a total list of 101 words, which is likely to explain their relatively low level of awareness. Second, the key findings remained significant even in a subset analysis that excluded those participants that indicated even partial awareness of the aims.

The most surprising aspect of these data, in the light of the strong influence of recent performance on meaning preferences that was seen in Experiments 1 and 2, is that this experiment showed no significant effect of recent experience. We therefore decided to pursue this recency effect (or the lack of it) in more detail in Experiment 4.

#### Experiment 4

The aim of this experiment was to follow up the non-significant effect of recent experience observed in Experiment 3. A similar web-based method was used, but we obtained additional information about their recent experience by asking them to tell us about every occasion on which they had rowed over the past week. This approach provides a far richer data set and allows us to look for effects not only of their most recent rowing episode, but also of rowing episodes earlier in the week.

#### Materials and method

##### Participants

35 rowing clubs throughout England, who did not participate in Experiment 3, were emailed using the same

incentive scheme as in Experiment 3. Data was obtained from 188 participants from nine different clubs. We excluded 39 participants who did not complete the whole experiment, 2 participants who were under 17, and 21 participants who indicated that English was not their first language or that they were not fluent in English. Of the remaining 126 participants who were included (54% female; aged 18–76; mean age = 34.5 years,  $SD = 15.9$ ) 92% were born and raised in Britain.

### Materials

The stimuli for the word association task were the same as in Experiment 3 (Appendix C), except that three of the suboptimal ambiguous words were excluded to reduce the length of the experiment: “stretcher” and “loom” had relatively low performance in the rowing knowledge test (<70%) indicating they were less frequently used by the rowers, and “stern” was sometimes generated by non-rowers indicating that it occurs relatively often in a non-rowing context. Three non-rowing fillers were also removed.

### Procedure

The procedure was similar to Experiment 3, except that we constructed seven different versions of the experiment (one for each day of the week), such that participants could select the version that was tailored to the day of the week on which they performed the experiment. This made it easier to ask questions about their recent rowing experience as we could refer to these days by their names. For example a participant who completed the experiment on a Sunday would be asked to tell us about whether they had been rowing on “Today (Sunday)”, “Yesterday (Saturday)”, “Friday”, etc. (see Rowing habits section below). The demographic questionnaire and the memory test were the same as in Experiment 3. The word association task and the rowing knowledge tasks were the same as in Experiment 3 except for the removal of six words (see Materials). The question about the experimental aims was removed (on the basis of the results of Experiment 3) to compensate for the additional length of the ‘rowing habits’ section.

Several important changes were made to the ‘rowing habits’ part of the experiment. To avoid loss of data for participants who did not consider rowing their main sport we explicitly instructed participants to answer these questions about rowing. This was done after the word association experiment so that it could not have biased these critical responses. As in Experiment 3, participants were then asked how long they had participated in this sport (years). They were then given a grid of two hour time slots for each of the preceding eight days (Before 6 am; 6–8 am ... 8–10 pm; after 10 pm) and were asked to tick all the time slots in which they had been rowing. There was a “no rowing” option for any participant who had not rowed in the past week. They were then asked to type in how often they had been rowing on average per week for each of the previous 10 months. All participants completed the experiment in November and they were asked to give this information for January through to October.

## Results

### Word association

Word association responses were coded using the same semi-automated procedure as for Experiment 3, except that the list of rowing responses from Experiment 3 was used as the starting list for automatically coding rowing responses with additional new rowing responses being added as required. As before, all individual responses that were coded as ‘rowing’ were manually checked and all coding was done blind to information from other variables.

A simultaneous multiple regression ( $N = 123$ )<sup>9</sup> explored the influence of the five potential predictor variables. Four of these variables were the same as in Experiment 3: age (years), and the measure of longer-term experience (number of years rowing experience), medium-term rowing experience (weekly rowing over the past month) and recent rowing experience (days since last rowed).<sup>10</sup> We also included a new measure of their medium-term rowing experience: their average weekly rowing frequency over the previous 9 months, and a new measure of the recent rowing experience: the number of times they had rowed in the previous week. As in Experiment 3, all the predictor variables related to rowing experience were log-transformed.

The two measures of medium-term frequency (frequency over the past month and over the past nine months) were very highly correlated ( $r = .57, p < .01$ ), and so to avoid problems of collinearity only the measure with the higher raw correlation with word association performance was included in the regression (frequency over the past nine months). Likewise the two measures of recent experience were very highly correlated ( $r = -.80, p < .001$ ) and so only rowing frequency over the past week was included. The final set of four predictor variables showed the following significant inter-correlations. Age was significantly positively correlated with long-term experience ( $r = .57, p < .001$ ), medium-term experience ( $r = .29, p = .001$ ) and recent experience ( $r = -.28, p = .001$ ), such that older rowers had rowed for longer and had rowed more often over the past nine months, but had rowed less often over the past week. (This counter-intuitive pattern of results may be explained by the fact that November is (on average) a relatively quiet period for the older non-university club rowers compared with the younger university club rowers). In addition, long-term rowing experience was positively correlated with medium-term experience ( $r = .39, p < .001$ ), and medium-term experience was positively correlated with recent experience ( $r = .44, p < .001$ ). Despite these intercorrelations there was no evidence of serious multicollinearity in this final set of predictor variables (maximum VIF = 2.00). A simultaneous multiple regression tested the effect of these variables on the proportion of

<sup>9</sup> Three rowers were excluded for suggesting that they had rowed 60 times per week on several months, this would be an extraordinary amount of rowing activity and likely reflects a misreading of the question.

<sup>10</sup> This data was only available for participants who had rowed within the last week, for participants who had not rowed within this week we assigned a value of 7 in this variable such that this measure of recent experience was not likely to be sensitive to differences in recent experience beyond the past seven days.



rowing responses. The overall regression produced a significant result ( $F(4, 118) = 6.3, p < .001, R^2 = 0.18$ ). As in Experiment 3, although age and overall rowing experience in years were positively correlated with each other, their influences on word associations were in opposite directions. Age had a non-significant negative effect ( $\beta = -.20, t = -1.6, p = .102$ ; more rowing responses from younger participants), whereas long-term experience positively predicted rowing responses ( $\beta = .28, t = 2.5, p = .01$ , more rowing response from rowers with longer experience). As in Experiment 3, the effect of medium-term rowing experience did not significantly predict rowing responses ( $p > .5$ ). In contrast to Experiment 3, the improved measure of recent experience had a significant positive effect ( $\beta = .25, t = 2.3, p = .021$ ).

An additional aim of this experiment was to explore in more detail the impact of very recent rowing experience. We therefore classified participants ( $N = 121$ )<sup>11</sup> on the basis of their recent experience into four categories: (i) people who had rowed on the same day as the experiment ( $N = 24$ ), (ii) people who had rowed on the day prior to the experiment (but not on the day of the experiment,  $N = 8$ ), (iii) people who had rowed both on the same day and the day prior ( $N = 11$ ), (iv) people who had not rowed in the last two days ( $N = 78$ ). The means from these conditions indicated that the rates of rowing responses were higher for participants who had rowed on the day of the experiment and participants who had rowed both that day and the day before, compared with those individuals who had only rowed the day before or had not rowed in the past two days (Fig. 6). An ANCOVA which included age and long-term rowing experience (log-transformed) as covariates, showed a main effect of recent experience ( $F(3, 115) = 3.8, p = .001$ ). (The covariate effect of long-term experience was also significant ( $F(1, 115) = 6.4, p = .012$ ), whereas the effect of age in this reduced subset was not significant ( $p > .1$ )). The complete set of pairwise comparisons between each of the four groups (using a Bonferroni corrected significance level of  $p < .008$ ) confirmed that people who had not rowed in the past two days differed significantly from those who had only rowed today ( $p < .001$ ), and from those who had rowed both today and yesterday ( $p = .002$ ), but not from those who had only rowed yesterday ( $p = .55$ ). The comparisons between the three groups of recent rowers showed that the group who only rowed yesterday was significantly different to the group who only rowed today ( $p = .002$ ), and was marginally different to the group who rowed both today and yesterday ( $p = .01$ ), but that these two latter groups did not differ from each other ( $p = .88$ ). In summary, these results indicate differences between those individuals who had rowed today and those who had not, but no additional effects of having rowed yesterday.

Finally, although the data concerning the time of day at which participants had rowed was not included in the analysis due to a lack of variance across participants, this data indicates that the two groups of rowers who had

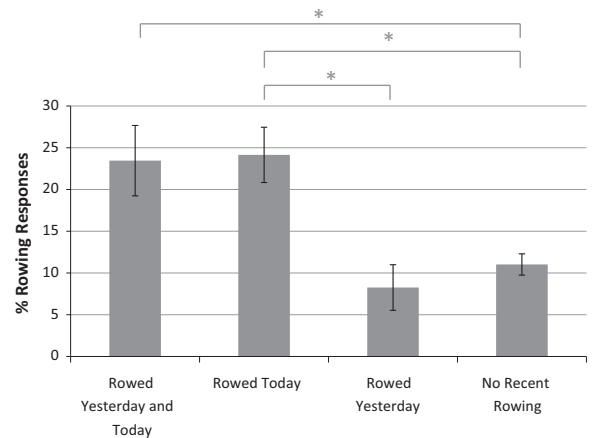


Fig. 6. Experiment 4: Proportion of rowing responses as a function of recent rowing experience.

rowed ‘today’ had median delay between their rowing training and them taking part in experiment of approximately 8 h (median rowing slot was 6–8 am; median time at which the experiment was started was 2.50 pm). Thus the effect of having rowed ‘today’ has had an effect on their performance several hours after this experience.

## General discussion

In this study, we investigated whether adult listeners align their lexical–semantic representations to previous experiences. To this aim, four experiments were conducted to determine how both recent and longer-term experiences with word meanings influence their accessibility. While the overall frequency with which a word meaning occurs in the language as a whole has long been considered to be a key factor that determines the accessibility of word meanings (Twilley & Dixon, 2000; Vitello & Rodd, 2015), the role of recent and medium-term experience was far from clear. Experiment 1 emphasised the importance of very recent experience on how words are interpreted. Listeners heard the critical ambiguous word primes (e.g., “ACE”) once or twice within fully disambiguating paragraph contexts as part of a radio programme. This brief and relatively naturalistic experience with the words influenced how listeners interpreted these ambiguous words when they were presented during a subsequent web-based word-association experiment (without disambiguating context). Those participants who had listened to the critical radio programme very recently were more likely to interpret these words in a way that was consistent with the word meaning that was used in the radio programme (e.g., “ACE-tennis” vs. “ACE-card”), compared to participants who either did not listen to the programme or who had waited several hours/days before taking part in the experiment. This effect of recent experience was replicated, within participants, in a lab-based setting in Experiment 2. We refer to this change in meaning preference on the basis of recent experience with a specific ambiguous word as ‘word-meaning priming’ (Rodd et al., 2013).

<sup>11</sup> Six participants were not categorised on this variable, because they had incorrectly selected the wrong version of the experiment, (i.e. their time of completion was inconsistent with the survey version) making this data unreliable.

On the assumption that speakers/writers tend to reuse word meanings within conversations/narratives, this priming effect is likely to have a strongly beneficial effect on listeners' ability to deal with ambiguous words in everyday life by boosting the availability of these meanings that they are more likely to encounter in the near future.

In both Experiments 1 and 2, this word-meaning priming effect was strongly modulated by the delay between prime and target. In Experiment 1, priming was largest for those participants who started the experiment very soon after hearing the prime paragraphs. This priming effect declined rapidly: after a 6 min delay the estimated priming magnitude was 6.2% (see Fig. 2), but this priming effect reduced to just 1% after a delay of about 3 h. Consistent with this, Experiment 2 found a significant decline of priming from 15% after 1 min to 8% after 20 min. Both experiments also indicate that this decay is non-linear such that the rate of decline reduces with time: in Experiment 1 a logarithmic decay function provided the best fit, while Experiment 2 found that in contrast with the significant decline seen in the first 20 min after the prime, there was no significant decline between 20 and 40 min, suggesting that the magnitude of priming becomes more stable during this period.

In contrast to these first two experiments, which investigated the effect of just one or two encounters with a word meaning over a very short delay, Experiments 3 and 4 studied the impact of more substantial exposure to rowing-related word meanings during the rowing activities of recreational rowers. Experiment 4 found that the tendency of recreational rowers to spontaneously generate rowing-related interpretations for words like “square” and “feather” was strongly influenced by their most recent rowing experience: those rowers who had not rowed on the same day as the experiment retrieved just 11% of rowing responses, but this increased to 24% in the rowers who had been rowing earlier that day and who had likely heard multiple instances of the target words. This proportional increase of 122% in the rate of rowing responses is particularly striking given that this group of ‘same day’ rowers had a median delay between their rowing experience and the experiment of 8 h. In contrast to this large effect of ‘same day’ rowing, their performance was not affected by whether or not they had rowed the previous day.

However, with respect to the rowing experiments (Experiments 3 and 4), one important caveat must be kept in mind: due to the naturalistic nature of the experiments, it remains somewhat uncertain exactly what aspects of their rowing experience are driving the observed changes in meaning preferences. Our preferred interpretation is that these changes in meaning preferences are a result of encountering the specific words used in this experiment earlier within the rowing environment. The majority of the words used in the experiment are words that we would expect to be encountered on the vast majority of rowing training episodes as they form a core part of the rowing vocabulary used both by coxes and coaches during training as well as by all participants discussing the training plan both before and after they get on the water. However, due to the fact that we could not monitor the language

use during training, we cannot rule out the possibility that (particularly for the same-day rowing effects) the changes in their interpretations of the words are (in part) being driven by some more generic effect of the rowing experience. For example, it may be that the experience of being in a rowing environment has activated a cluster of semantically related word meanings, including those that were not specifically heard during the rowing episode and that all rowing-related word meanings are boosted by this experience. The latter account, which is akin to a semantic priming account, would make the prediction that priming could be observed even for those rowing terms that were not encountered. We suggest that this explanation is unlikely. Semantic priming effects have not been reported in the literature at such long delays, and we have previously shown that word-meaning priming is more long-lived than semantic priming (Rodd et al., 2013). Future studies looking at this issue will need to use carefully designed novel paradigms that combine the naturalistic elements of this approach with a greater degree of experimental control over the priming phase of the experiment to explore whether experiences such as going rowing, which usually last for over an hour, can produce sufficiently long-lived semantic priming effects that can drive relatively long-term changes to meaning preferences.

In addition to these findings that recent (same-day) experience can influence the interpretation of ambiguous words, Experiments 3 and 4 also emphasised the importance of *longer-term* experience on how individuals interpret words. In both these experiments we found that the tendency to retrieve rowing-related meanings increased as a function of the number of years that they had been rowing, indicating that their experiences had had a cumulative effect on their overall preferences. Surprisingly, in both Experiments 3 and 4 there was no additional effect of medium-term rowing experience – there was no significant effect of the average rowing frequency over either the last month (Experiment 3) or the last 9 months (Experiment 4).

Taken together, these results suggest that our most recent (i.e. same day) linguistic experience can have a very large effect on how we interpret ambiguous words. In the case where this experience comprises just one or two instances of the ambiguous word (Experiments 1 and 2) the increases can be relatively modest and fast fading, but when participants repeatedly hear these words within their usual natural context (Experiment 4), the effects of this experience can be large and can last for several hours. However the absence of any effect of linguistic experience from the previous day or of average exposure over the preceding months (Experiments 3 and 4) suggests that the long-term changes that are produced by many years of linguistic experience develop very slowly and incrementally over time.

These changes in meaning preferences are important as they are predicted to directly affect the ease with which these meanings can be processed within sentence contexts. For example, in Experiment 4 we found that having rowed earlier that day produced an absolute increase of 12% in the retrieval of rowing related meanings, such that participants were more than twice as likely to generate the

rowing related meanings if they had rowed that day. As previously discussed, current models (e.g., the reordered access model, Duffy et al., 1988, 2001) predict that these changes can have clear impact on the ease with which these meanings can be processed within sentences contexts. In particular, the models predict a strong benefit for those words for which priming can cause a moderately subordinate meaning (e.g., “slide”; dominance = 25%) to become equally preferred to the alternative non-rowing meaning. It is less clear what impact word-meaning priming might have on comprehension of strongly subordinate word meanings (e.g., the rowing meaning of “square”, with a dominance of 5%). Even after priming such meanings are likely to remain subordinate, and therefore are predicted to remain relatively difficult to access due the presence of a dominant competitor meaning.

In addition to providing insights about the *time-course* with which experience can affect the interpretation of ambiguous words, Experiment 1 showed that word-meaning priming occurs with equal magnitude regardless of whether the speaker used in the test phase was the same as in the prime phase. For example, hearing the BBC radio presenter Jenni Murray use the word “ace” to refer to a tennis serve during the prime phase increased participants’ tendency to interpret the word “ace” in this way in the test phase regardless of whether they were hearing the word spoken by her or by a different radio presenter, Edward Stourton. This absence of priming same-speaker advantage was even seen for the subset of participants who indicated that they had recognised the speaker that they heard during the word association task. This finding is consistent with the results of Rodd et al. (2013; Experiment 3), who found a similar null effect of a voice-change manipulation in a lab-based priming experiment using *unfamiliar* speakers.

A final key finding from these experiments is that the degree to which participants change their preferences for individual word meanings is modulated by age. In Experiment 1, the short-term word-meaning priming effect was significantly larger for younger participants, and in Experiments 3 and 4 regression analyses, which took into account the individuals’ rowing experience, showed that this experience had a larger effect on performance in younger participants. (This effect was significant in Experiment 3 and marginal in Experiment 4.) There are a host of possible explanations for these age effects. One intriguing possibility is that this finding reflects a more general phenomenon whereby the plasticity of lexical-semantic representations declines with age such that these representations become increasingly stable over time. Such an explanation is closely linked to explanations of the age-of-acquisition effects seen on a range of verbal and non-verbal tasks (Ellis & Lambon Ralph, 2000; Lambon Ralph & Ehsan, 2006). However, the complex patterns of results concerning how age can impact on *perceptual* learning to speech stimuli (see Scharenborg & Janse, 2013, for a recent review) suggests that any effect of age on this form of lexical-semantic retuning is likely to be complex, with multiple contributory factors. In particular, it is important to assess the extent to which these age effects may arise indirectly as a consequence of other age-related changes.

While hearing loss in older adults may reduce the efficiency of linguistic input (and hence word-meaning priming), we reasoned that this is an unlikely contributor to the age-related priming reduction we observed as, for instance, participants in Experiment 3 (aged 17–55) were below the age at which age-related hearing loss usually has a substantial impact (Van Rooij, Plomp, & Orlebeke, 1989). More likely candidates that may underlie that age-related reduction in priming include attention and memory factors: as people age, their attention and memory functions decline, rendering them less efficient in comprehending linguistic input and thus leading to reduced word-meaning priming. In addition, in the case of Experiments 3 and 4 it is possible that the qualitative nature of the linguistic experience may be changing with age, such that younger rowers’ rowing environment is more verbally rich than the environment of older rowers. Therefore, while these data provide clear indications that there are age-related changes in how older listeners make use of their everyday linguistic input, they do not yet reveal the precise mechanism(s) by which these age-related changes arise. Future lab-based experiments are clearly needed to disentangle these potential contributory factors.

Taken together these data about the time-course of word-meaning priming, the absence of a same-speaker advantage, and the modulatory effect of age provide important insights into the underlying mechanism(s) of word-meaning priming. Any plausible account must accommodate the finding that (i) there is a very large short-lived component that fades rapidly within the first 40 min after an encounter with the ambiguous word but that is not affected by a change in speaker identity, (ii) the impact of these encounters can last several hours, and (iii) there is an incremental effect of exposure that builds up across our lifetimes. We suggest that this constellation of findings may prove relatively challenging for current models to accommodate. At first glance, distributed connectionist models of how ambiguous words are represented and processed (e.g., Joordens & Besner, 1994; Kawamoto, Farrar, & Kello, 1994; Rodd, Gaskell, & Marslen-Wilson, 2004) seem well placed to accommodate such priming effects due to their use of a learning algorithms. For example, in Rodd et al.’s (2004) model, whenever the form of an ambiguous word is encountered, this activation feeds forward to activate the semantic units that are associated with its meanings. Initially, this pattern of semantic activation corresponds to a blend (or mixture) of its two meanings, but the recurrent connections between the individual semantic units then ‘clean up’ this activation to ensure that the network eventually settles into a pattern of activation that corresponds to one of its known meanings. Within this framework, any experience with one of the meanings would strengthen the connections between its phonological/orthographic units and its semantic units such that when the model next encounters the word’s form there is an increased probability of it settling into the recently encountered meaning. In addition, equivalent changes to the connections within the semantic layer could potentially make the attractor basin for that meaning more stable, relative to the alternative unprimed meaning. However, while such a connectionist model can

almost certainly accommodate word-meaning priming in a very general sense, substantial modifications to the learning algorithm may well be needed to accommodate the time-course with which word-meaning priming decays: these models do not inherently contain any mechanism by which changes to connection strengths on the basis of recent experience vary as a function of time per se. One possibility is that the decay function could arise purely due to interference from intervening encounters with other unrelated words: each such encounter would result in weight changes, which could potentially influence even apparently unrelated words because these may share some connections within the highly interconnected distributed network. But in our view it is far from clear that interference effects of this type can necessarily accommodate the decay function we observed. Future experimental work looking at whether this decay is driven by time per se or by interference from other linguistic input, together with computational simulations, is necessary to determine the likely mechanism for the decay function seen in these data. Future work must also consider the possibility that multiple mechanisms are at work, such that the large by relatively short-lived priming seen in Experiments 1 and 2 and the longer-term learning observed in Experiments 3 and 4 are driven by qualitatively different mechanisms. Again such a finding cannot easily be accommodated by current models. To address this issue, future experiments should explore the different factors that might modulate word-meaning priming at different time-courses.

Finally, these experiments illustrate the utility of relatively large-scale experiments in which the variation in participants' experiences arises either due to natural variability in their linguistic experiences (Experiments 3 and 4) or is manipulated in a relatively naturalistic manner, such as via a radio programme (Experiment 1). These experiments were only possible due to the use of web-based data collection procedures, which do not require participants to come into the lab for testing. We are hopeful that future studies which extend this approach to include more sensitive reaction time measures may allow researchers to make rapid progress on these key theoretical issues.

In summary, the three large-scale web-based experiments and one lab-based experiment reported here provide a novel demonstration that adult listeners adapt their lexical-semantic representations to their most recent, as well as longer-term, experience such that recently encountered word meanings are subsequently more readily available. As ambiguous words tend to have the same meaning when occurring in the same context, such lexical-semantic alignment with past experience allows listeners and readers to more accurately and rapidly select appropriate word meanings, thus facilitating language comprehension fluency, as other forms of linguistic alignment have been suggested to do (e.g., Pickering & Garrod, 2004). More generally, these results require a fundamental shift in thinking away from the view that lexical-semantic representations are stable and fixed, towards a more dynamic, experience-driven account, in which they are viewed as highly fluid, flexible representations that are continually updated in order to optimise the efficiency of comprehension.

## Acknowledgments

This work was supported by the Economic and Social Research Council [grant numbers ES/K013351/1; ES/L010224/1; ES/J500185/1] and the Experimental Psychology Society. We thank Pier Productions Limited and BBC Radio 4, especially Toby Murcott, Michael Blastland, Edward Stourton and Jenni Murray for their help with Experiment 1. We thank the Radio 4 listeners and the members of the 12 rowing clubs who participated in these experiments. We thank Andy Fugard for his help creating the figures.

## Appendix A

Experiment 1 stimuli: Scripts used in prime stage in the order that they were presented. Target ambiguous words are underlined.

Mark hopes to become a professional tennis player. He practises on court at his club with his coach most days. His strongest weapon on court is his serve. He regularly gets out of trouble in a match with an ace on his serve. In his last match Mark saved three set points with an ace. (Spoken by ES.)

Neil is in prison for fraud. He'd like to appeal against his sentence, but does not have a strong case. Neil is finding life in his cell very hard, but his sentence is nearly over. His victims may decide to sue him, and he would find it hard to be in the dock again at a civil case. (Spoken by JM.)

Grace is a keen musician and plays the piano, organ and French horn. The organ that Grace plays in her local church has a beautiful tone, especially for the bass notes and for chords with a lower pitch. (Spoken by ES.)

Daisy is a weekly boarder at her school. The school principal has let the pupils organise a ball at the end of term. The principal has agreed to a medieval theme for the ball, so Daisy and the other pupils will be dressing as knights and princesses. Definitely no jeans allowed. (Spoken by JM.)

## Appendix B

Experiment 2: Ambiguous words.

Appendix, Ball, Band, Bar, Bark, Bat, Bed, Blue/Blew, Bonnet, Bow, Bowl, Box, Break, Bulb, Button, Cabinet, Calf, Cap, Card, Case, Change, Chest, China, Coach, Cold, Craft, Cricket, Cross, Cup, Deck, Draw/Drawer, Fan, Fence, Figure, Flower/Flour, Gear, Glasses, Gum, Hair/Hare, Hand, Interest, Iron, Issue, Jam, Joint, Key, Lace, Landing, Letter, Match, Mark, Mould, Mouse, Nail, Night/Knight, Note, Organ, Palm, Panel, Park, Pear/Pair, Pen, Pipe, Plug, Punch, Pupil, Racket, Record, Ring, Sea/See, Sign, Sink, Skip, Spade, Speaker, Spring, Staff, Step, Stitch, Straw, Strike, Sun/Son, Temple, Toast, Trailer, Trunk, Watch, Wave.

## Appendix C

Experiment 3 and 4 stimuli: Ambiguous words with a rowing-related meaning. (Words in brackets were only included in Experiment 3).

Gate, (Stretcher,) Head, Scratch, Finish, Crab, Bump, Stroke, Bow, Eight, (Stern,) Pot, Shell, Sweep, Square, Feather, Blade, (Loom,) Slide, Catch, Collar.

## References

- Adank, P., Evans, B. G., Stuart-Smith, J., & Scott, S. K. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 520–529.
- Adank, P., & Janse, E. (2010). Comprehension of a novel accent by young and older listeners. *Psychology and Aging*, 25, 736–740.
- Arai, M., Van Gompel, R. P., & Scheepers, C. (2007). Priming ditransitive structures in comprehension. *Cognitive Psychology*, 54, 218–250.
- Bainbridge, J. V., Lewandowsky, S., & Kirsner, K. (1993). Context effects in repetition priming are sense effects. *Memory & Cognition*, 21, 619–626.
- Binder, K. S., & Morris, R. K. (1995). Eye movements and lexical ambiguity resolution: Effects of prior encounter and discourse topic. *Journal of Experimental Psychology: Learning Memory and Cognition*, 21, 1186–1196.
- Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General*, 129, 177–192.
- Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, 106, 707–729.
- Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition*, 75, B13–B25.
- Burke, D. M., & Shafto, M. A. (2008). Language and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (3rd ed., pp. 373–443). New Jersey: Lawrence Erlbaum Associates.
- Cai, Z. G., Davis, M., Gaskell, G., Farrar, L., Adler, S., & Rodd, J. M. (2015). Speaker accent modulates access to word meanings. *Paper presented at the 21st architectures & mechanisms for language processing, Malta*.
- Cai, Z. G., Pickering, M. J., & Branigan, H. P. (2012). Mapping concepts to syntax: Evidence from structural priming in Mandarin Chinese. *Journal of Memory and Language*, 66, 833–849.
- Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22, 1–39.
- Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America*, 116, 3647–3658.
- Coane, J. H., & Balota, D. A. (2009). Priming the holiday spirit: Persistent activation due to extra-experimental experiences. *Psychonomic Bulletin & Review*, 16, 1124–1128.
- Coane, J. H., & Balota, D. A. (2011). Face (and nose) priming for book. *Experimental Psychology*, 58, 62–70.
- Copland, D. A. (2006). Meaning selection and the subcortex: Evidence of reduced lexical ambiguity repetition effects following subcortical lesions. *Journal of Psycholinguistic Research*, 35, 51–66.
- Cristia, A., Seidl, A., Vaughn, C., Schmale, R., Bradlow, A., & Floccia, C. (2012). Linguistic processing of accented speech across the lifespan. *Frontiers in Psychology*, 3.
- Dagerman, K. S., MacDonald, M. C., & Harm, M. W. (2006). Aging and the use of context in ambiguity resolution: Complex changes from simple slowing. *Cognitive Science*, 30, 311–345.
- Davis, H. P., Cohen, A., Gandy, M., Colombo, P., Van Dusseldorp, G., Simolke, N., & Roman, J. (1990). Lexical priming deficits as a function of age. *Behavioral Neuroscience*, 104, 288–297.
- Duffy, S. A., Kambe, G., & Rayner, K. (2001). The effect of prior disambiguating context on the comprehension of ambiguous words: Evidence from eye movements. In D. S. Gorfein (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity*. Washington, DC: American Psychological Association.
- Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, 27, 429–446.
- Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1103–1123.
- Federmeier, K. D., & Kutas, M. (2005). Aging in context: Age-related changes in context use during language comprehension. *Psychophysiology*, 42, 133–141.
- Fine, A. B., & Jaeger, T. F. (2013). Evidence for implicit learning in syntactic comprehension. *Cognitive Science*, 37, 578–591.
- Fleischman, D. A., Gabrieli, J. D. E., Gilley, D. W., Hauser, J. D., Lange, K. L., Dwornik, L. M., ... Wilson, R. S. (1999). Word-stem completion priming in healthy aging and Alzheimer's disease: The effects of age, cognitive status, and encoding. *Neuropsychology*, 13, 22–30.
- Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition*, 89, 105–132.
- Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context and consequences. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of accommodation: Developments in applied sociolinguistics* (pp. 1–68). Cambridge, England: Cambridge University Press.
- Golomb, J. D., Peelle, J. E., & Wingfield, A. (2007). Effects of stimulus variability and adult aging on adaptation to time-compressed speech. *Journal of the Acoustical Society of America*, 121, 1701–1708.
- Jackson, A., & Morton, J. (1984). Facilitation of auditory word recognition. *Memory & Cognition*, 12, 568–574.
- Joordens, S., & Besner, D. (1994). When banking on meaning is not (yet) money in the bank – Explorations in connectionist modeling. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 1051–1062.
- Kawamoto, A. H., Farrar, W. T., & Kello, C. T. (1994). When two meanings are better than one: Modeling the ambiguity advantage using a recurrent distributed network. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 1233–1247.
- Lambon Ralph, M. A., & Ehsan, S. (2006). Age of acquisition effects depend on the mapping between representations and the frequency of occurrence: Empirical and computational evidence. *Visual Cognition*, 13, 928–948.
- Light, L. L., & Carter-Sobell, L. (1970). Effects of changed semantic context on recognition memory. *Journal of Verbal Learning and Verbal Behavior*, 9, 1–11.
- Luce, P. A., & Lyons, E. A. (1998). Specificity of memory representations for spoken words. *Memory & Cognition*, 26, 708–715.
- Masson, M. E. J., & Freedman, L. (1990). Fluent identification of repeated words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 355–373.
- Meiran, N., & Jelicic, M. (1995). Implicit memory in Alzheimer's disease: A meta-analysis. *Neuropsychology*, 9, 291–303.
- Mullenix, J. W., Pisoni, D. B., & Martin, C. S. (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America*, 89, 365–378.
- Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception & Psychophysics*, 60, 355–376.
- O'Brien, R. M. (2007). A caution regarding rule of thumb for variance inflation factors. *Quality and Quantity*, 41, 673–690.
- Orfanidou, E., Davis, M. H., Ford, M. A., & Marslen-Wilson, W. D. (2011). Perceptual and response components in repetition priming of spoken words and pseudowords. *Quarterly Journal of Experimental Psychology*, 64, 96–121.
- Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America*, 119, 2382–2393.
- Peelle, J. E., & Wingfield, A. (2005). Dissociations in perceptual learning revealed by adult age differences in adaptation to time-compressed speech. *Journal of Experimental Psychology: Human Perception and Performance*, 31, 1315–1330.
- Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, 134, 427–459.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 169–190.
- Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 785–794.
- Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., & Baayen, H. (2014). The myth of cognitive decline: Non-linear dynamics of lifelong learning. *Topics in Cognitive Science*, 6, 5–42.
- Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading – Effects of word-frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, 14, 191–201.
- Rodd, J. M., Berriman, R., Landau, M., Lee, T., Ho, C., Gaskell, M. G., & Davis, M. H. (2012). Learning new meanings for old words: Effects of semantic relatedness. *Memory & Cognition*, 40, 1095–1108.
- Rodd, J. M., Gaskell, G., & Marslen-Wilson, W. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language*, 46, 245–266.
- Rodd, J. M., Gaskell, M. G., & Marslen-Wilson, W. D. (2004). Modelling the effects of semantic ambiguity in word recognition. *Cognitive Science*, 28, 89–104.
- Rodd, J. M., Johnsrude, I. S., & Davis, M. H. (2010). The role of domain-general frontal systems in language comprehension: Evidence from

- dual-task interference and semantic ambiguity. *Brain and Language*, 115, 182–188.
- Rodd, J. M., Lopez Cutrin, B., Kirsch, H., Millar, A., & Davis, M. H. (2013). Long-term priming of the meanings of ambiguous words. *Journal of Memory and Language*, 68, 180–198.
- Samuel, A. G., & Kraljic, T. (2009). Perceptual learning for speech. *Attention, Perception, and Psychophysics*, 71, 1207–1218.
- Scharenborg, O., & Janse, E. (2013). Comparing lexically guided perceptual learning in younger and older listeners. *Attention, Perception, and Psychophysics*, 75, 525–536.
- Simpson, G. B., & Kang, H. (1994). Inhibitory processes in the recognition of homograph meanings. In *Inhibitory processes in attention, memory, and language* (pp. 359–381). New York: Academic Press.
- Traxler, M. J. (2008). Lexically independent priming in online sentence comprehension. *Psychonomic Bulletin & Review*, 15, 149–155.
- Twilley, L. C., & Dixon, P. (2000). Meaning resolution processes for words: A parallel independent model. *Psychonomic Bulletin & Review*, 7, 49–82.
- Twilley, L. C., Dixon, P., Taylor, D., & Clark, K. (1994). University of Alberta norms of relative meaning frequency for 566 homographs. *Memory & Cognition*, 22, 111–126.
- Van Rooij, J. C. G. M., Plomp, R., & Orlebeke, J. F. (1989). Auditive and cognitive factors in speech perception by elderly listeners. I: Development of test battery. *Journal of the Acoustical Society of America*, 86, 1294–1309.
- Vitello, S., & Rodd, J. M. (2015). Brain mechanisms of ambiguity resolution in spoken sentences. *Language and Linguistic Compass*, 9, 391–405.
- Wagenmakers, E. J., Zeelenberg, R., & Raaijmakers, J. G. (2000). Testing the counter model for perceptual identification: Effects of repetition priming and word frequency. *Psychonomic Bulletin and Review*, 7, 662–667.
- Woltz, D. J., & Shute, V. J. (1995). Time course of forgetting exhibited in repetition priming of semantic comparisons. *American Journal of Psychology*, 108, 499–525.