

Context effects on word association production: a semantic warping account

Stanley H. West

Department of Psychology

Eileen K. Haebig

Department of Communication Sciences and Disorders

Christopher R. Cox

Department of Psychology

Louisiana State University
Baton Rouge, LA 70803 USA

Abstract

An important aspect of human cognition is our ability to adapt our behavior to changing situations and contexts. Semantic control is generally broken into two different modes acting at varying levels of domain specificity: general rule-based selection or contextually-altered semantic space. The current study examines how context shifts influence associative behavior across three context domains. We instructed participants to make word associations as if they were interacting with a toddler (i.e. child condition), interacting with a peer (i.e. peer), or to just produce short words. We found that participants in the child condition produced more child-directed speech than the other conditions. Specifically, these responses were shorter, acquired earlier, and higher frequency and contextual diversity. Additionally, the child condition resulted in different representational similarity structure than the other two conditions, providing evidence for a context-effect that is less rule based and more akin to a flexible shifting of semantic space.

Keywords: controlled semantics, semantic cognition, context-effects

Introduction

Imagine walking out of a theater after enjoying the latest superhero movie with someone you know well. Having shared that experience, you are excited to discuss with them. Now, consider how you might initiate that conversation if your companion is a young child, or a peer, or someone a generation your senior. Despite experiencing the same audio-visual stimulation, this other person will have experienced the movie differently and encoded different semantic structure. Certain themes, events, or characters are more important to them than they were to you; your appreciation of the motives and morality of the protagonist differs, and perhaps you identified different heroes and villains within the story. This is all possible, despite being able to name all the same characters and recount the same overarching narrative as your companion. Because you know them well, you may be able to infer aspects of their experience and engage in a conversation that is sensitive to that understanding.

What are the cognitive mechanisms that underlie such context sensitivity? Generally, there are two modes of control over the activation of semantic knowledge. The first involves targeted selection and inhibition of concepts or conceptual dimensions via an explicit rule or instruction. This is a general kind of control that relies on the ventral lateral prefrontal cortex (vlPFC) along with other areas involved with executive control across all cognitive domains. For example, if discussing the movie with a young child, you may selectively

inhibit discussing themes of mortality, legacy, and regret that may be more salient to someone who has lived more life and selectively target themes of compassion, selflessness, and bravery. Such inhibition may take the form of rejecting candidate topics that come to mind if they are appraised to be inappropriate for a child, or by selectively and explicitly inhibiting dimensions the semantic representation prior to activation/retrieval (e.g., Jackson et al., 2021).

The second mode of control involves mechanisms specific to the semantic domain and is associated with posterior lateral temporal lobe, perhaps especially posterior middle temporal gyrus (pMTG). While the nature of this domain-specific semantic control is less well understood, by studying which semantic control processes load onto the pMTG and how the pMTG is functionally and anatomically situated with respect to multiple distributed neural networks one can motivate a plausible hypothesis. The pMTG is more engaged by demanding semantic control tasks, not context-agnostic auto-associative priming tasks (Gennari et al., 2007; Noonan et al., 2013). It is positioned at the intersection of the default mode (DMN) and multiple demand (MDN) networks (Davey et al., 2016), which engage in automatic and controlled processing, respectively. The pMTG is more strongly connected with the anterior vlPFC (Snijders et al., 2010), which is more engaged in semantic retrieval, than posterior vlPFC, which is more engaged when selecting among active alternatives (i.e., stimuli have been presented or which have already been retrieved).

We hypothesize that domain-specific semantic control “warps” semantic space such that context appropriate knowledge is more accessible and context appropriate associations are strengthened. This hypothesis can be experimentally tested in paradigms designed to assess association strengths following different context manipulations, such as contextualized lexical decision tasks (LDT; including semantic primes) and word association tasks. The LDT requires prior expectations about which pairs of words will show differential priming effects between contexts. Prior work with LDT has clearly demonstrated pre-task priming can moderate the interpretation of polysemous cues.

The current study examines how context shifts influence associative behavior across three context domains. We instructed our participants to make word associations as if they were interacting with a toddler (i.e., child condition), as if interacting with a peer (i.e., peer), or to just produce short

4021

Table 1: Participant characteristics

N(male)	Age				Is a parent		
	Mean	SD	Min	Max	Child	Peer	Short
300(63)	19.39	2.18	18	44	.01	.01	.01

Note. “Is a parent” indicates the proportion of participants from each condition that are the parent of a toddler.

words. We hypothesized that if participants were not using any rule-based selection strategy (e.g., picking only short words), the child condition would elicit a different profile of responses than the short-response condition. This would provide evidence for a context-effect that is less rule-based and more akin to a flexible shifting of semantic space.

Methods

Participants

All participants were native English speakers, 18 years of age or older, enrolled as undergraduates at Louisiana State University, and recruited through the SONA Systems platform and compensated with class credit. Our target sample size was 100 participants in each of three experimental conditions. Because we anticipated data loss due to incomplete or low-effort responses, we recruited 359 participants and manually reviewed their responses before beginning analysis. After excluding 28 participants for poor data quality, we retained 100 the remaining participants from each condition at random. Characteristics of the 300 participants retained for analysis can be found in Table 1.

Overall, participants in the short condition ($M = 50.0$, $SD = 110$) took longer to complete the 60 word associations than the child ($M = 26.4$, $SD = 21.6$) or peer ($M = 24.0$, $SD = 16.9$) conditions, $t = 3.05$, $p < 0.01$. There was no significant difference in completion times between the child and peer conditions. This difference was tested by fitting a linear model contrasting the survey-duration time across conditions using Helmert contrasts to code conditions. This human subjects research was approved by the IRB at Louisiana State University.

Stimuli

Sixty cue words were selected from the MacArthur Bates Communicative Development Inventory (Fenson et al., 2007), which is a child vocabulary checklist that is comprised of 680 early acquired words. These are a subset of the cues utilized by Cox and Haebig (2023), chosen with the intention to capture the effects of context observed between their child and peer word association tasks. Because these cues previously elicited responses with significantly lower ages of acquisition (AoA), fewer letters and syllables, and higher frequency and contextual diversity in spoken language from participants in the child condition, they are useful for testing our hypotheses about the kind of semantic control behind the behavior observed by Cox and Haebig (2023).

Design and Procedure

Participants were directed to our Qualtrics (Qualtrics, Provo, UT) survey through SONA Systems and indicated if they are a native English speaker. After consenting to participate, they were randomly assigned to either the child context manipulation, the peer context manipulation, or the short context manipulation. Across all conditions, participants were prompted to list three words in response to each cue (De Deyne & Storms, 2008). Collecting multiple responses per cue builds richer response profiles for each cue, which is helpful when estimating the similarity structure among cues, and allows us to examine the generation process—do context effects wane as participants report increasingly weak associations?

Participants were randomly instructed to generate responses during the word association task 1) as if interacting with someone with the same knowledge and life experience as themselves (peer condition), 2) as if interacting with a toddler (child condition), or 3) by reporting the shortest words that come to mind (short condition). In the child condition, participants are given the cover story that they are playing a word association game with a toddler. Each of the participants were presented with the same 60 cues, regardless of condition, in random order. Finally, all participants were to avoid generating associations to their own prior responses (chaining) and to provide one-word responses.

Data Cleaning

We assessed each profile of responses provided by our participants to ensure that they were on-task and not providing low effort responses (e.g., responses that were single letters or non-words). We excluded one participant for low effort and 27 for incompleteness. Additionally, we manually corrected spelling and regularized to lemmas, while cross-referencing with databases containing psycholinguistic variables of interest: word frequency and contextual diversity from SUBTLEX (Brysbaert & New, 2009) and AoA data published by Kuperman et al. (2012). We managed these relationships in a SQLite database (Hipp, 2018) to facilitate cross-dataset links in a way that was robust to different lemmatization/regularization/spelling decisions by different authors.

Psycholinguistic Analysis

We used five dependent variables to examine individual responses: word length, log transformed word frequency and contextual diversity, age of acquisition (AoA), and number of syllables. Word frequency and contextual diversity were estimated by SUBTLEX and are operationally defined as the cumulative frequency of words in the corpus of subtitles and the number of unique contexts in which the words appear, respectively (Brysbaert & New, 2009). AoA data were provided by pre-collected normative data (Kuperman et al., 2012). Word length was calculated using a character count, and number of syllables was computed using the a *nsyllable* counting function in the R package *quanteda* (Benoit et al., 2018).

For each of our dependent variables, we fit a separate 3 (condition: child, peer, short) x 3 (response order: first, sec-

ond, or third) within-cue ANOVAs for each of our dependent variables using the *ez* package in R (Lawrence, 2016).

Response Profile RSA Analysis

We were also interested to see if the three conditions would elicit different similarity structures between them using representational similarity analysis (RSA, Nili et al., 2014). To accomplish this, we first created a table for each condition with columns for each of the 60 cues and rows for each unique response provided. Cells in this matrix were filled with 1s if a given response was given for a cue and 0s otherwise. We then computed Pearson’s *r* for each pair of columns resulting in 60 cue x 60 cue correlation matrix for each condition. This was followed by computing the Spearman rank correlation between the lower triangles of each matrix for every combination of condition, resulting in the “representational similarity” between each condition. This correlation value will be low if the relationships among cues differ between conditions, and high if they are similar.

To analyze these Spearman rank correlations in reference to a null distribution, we combined the response data from two conditions by cue and randomly split the data for each cue 1,000 times. Again, this was done for each combination of conditions. The procedure above was repeated for each of these splits, resulting in a distribution of rank correlations representing a null distribution. We then compared rho for each of the groups to our null distribution to determine if either group attributed different meanings to the cues based on what condition they were in.

All analyses were conducted in R (R Core Team, 2024).

Results

Psycholinguistic Results

Condition by dependent variable descriptive statistics and omnibus model outputs can be found in Table 2 and Table 3, respectively. We also provide a graphical representation of t-values for each condition contrast for each dependent variable in Figure 1. Dashed lines represent two-tailed significance at $p < .05$. Within-cue ANOVAs revealed significant condition, response order, and condition by response order interactions for AoA. Overall, participants in the child condition produced lower AoA words than in the peer and short conditions across all three responses; $F(1, 59) = 116.16, p < .001, F(1, 59) = 111.59, p < .001$, respectively. AoA tended to increase across responses in all conditions, with the condition

Table 2: Description of associations by condition

Condition	Psycholinguistic Measure $M(SD)$				
	AoA	Letters	Log WF	Log CD	Syllables
Child	4.62(1.60)	5.06(1.77)	3.49(0.85)	3.13(0.64)	1.48(0.70)
Peer	4.85(1.82)	5.23(1.97)	3.44(0.89)	3.08(0.68)	1.54(0.77)
Short	4.86(1.82)	5.19(1.89)	3.43(0.89)	3.08(0.68)	1.53(0.75)

Note. AoA = Age of Acquisition, Log WF = \log_{10} Contextual Diversity, Log WF = \log_{10} Word Frequency

Table 3: Omnibus ANOVA Output Table

	df_N	df_D	F	p	η^2
AoA					
Condition	2	118	86.34	< .001	0.06
Response	2	118	67.71	< .001	0.14
C x R	4	236	2.83	< .01	0.003
Letters					
Condition	2	118	46.39	< .001	0.03
Response	2	118	27.81	< .001	0.08
C x R	4	236	0.49	0.74	0.001
Log CD					
Condition	2	118	45.75	< .001	0.009
Response	2	118	27.20	< .001	0.02
C x R	4	236	1.98	0.10	0.00
Syllables					
Condition	2	118	30.00	< .001	0.01
Response	2	118	14.97	< .001	0.04
C x R	4	236	0.10	0.98	0.00
Log WF					
Condition	2	118	29.11	< .001	0.004
Response	2	118	32.49	< .001	0.02
C x R	4	236	2.47	< .05	0.00

Note. AoA = Age of Acquisition, Log CD = \log_{10} Contextual Diversity, Log WF = \log_{10} Word Frequency, df_N = Numerator Degrees of Freedom, df_D = Denominator Degrees of Freedom

effect attenuating in the child condition as first, second, and third responses were made. There were no significant AoA differences between the short and peer conditions.

Additionally, there were condition and response order effects in the length of responses with no interaction, allowing us to interpret these as main effects. Here, participants in the child condition produced shorter words than in the peer and short conditions; $F(1, 59) = 76.39, p < .001, F(1, 59) = 66.74, p < .001$, respectively. As with AoA, word length also increased as more responses were made. The significant difference between the child and short conditions indicates that the child condition influences responses beyond just a short response “filtering” strategy. There were also significant condition and response-order effects in the number of syllables. Responses in the child condition contained fewer syllables than the peer and short conditions; $F(1, 59) = 45.77, p < .001, F(1, 59) = 46.31, p < .001$, respectively. No other condition comparisons were significant. Second and third responses tended to have more syllables than the first response.

There were also condition, response order, and condition

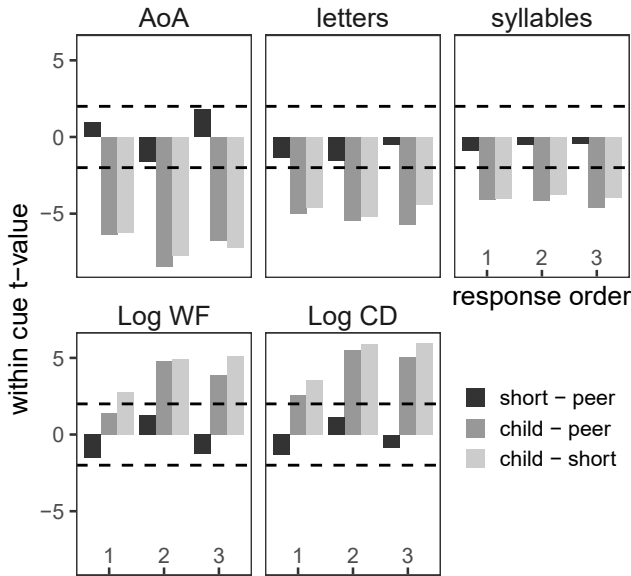


Figure 1: Difference of psycholinguistic characteristics between associations provided in each pair of conditions, reported as within-cue t-values. The x-axis represents the response order.

by response order interactions for word frequency. Here, the child condition elicited higher frequency words than the short and peer conditions; $F(1, 59) = 50.77, p < .001, F(1, 59) = 31.64, p < .001$, respectively. Again, this condition effect was only significant between the child condition and the other two conditions. In terms of response order, word frequency tended to decrease as more responses were made, with this decrease being more pronounced in the peer and short conditions. Finally, contextual diversity also differed between conditions and response orders, with no interaction. Participants in the child condition produced more contextually diverse responses than in the peer and short conditions; $F(1, 59) = 53.84, p < .001, F(1, 59) = 76.88, p < .001$, respectively. Contextual diversity also declined as participants provided their second and third associative responses.

Response Profile RSA Results

Associations elicited in the child and short conditions are starkly different. In particular, the finding that providing a semantically-coherent context in which short words are appropriate yields consistently shorter words than providing an explicit but non-semantic goal suggests that these two conditions tap different cognitive control mechanisms. However, it does not necessarily suggest that the mechanism involves *warping* the semantic space. Warping would imply not only the increased availability of context-appropriate words, but change in the representational similarity among concepts. Figure 2A plots true correlations between conditions (points) in comparison to their accompanying simulated null distributions (violin plots; see Methods). Figure 2B plots the z-scored difference of the true correlation between conditions

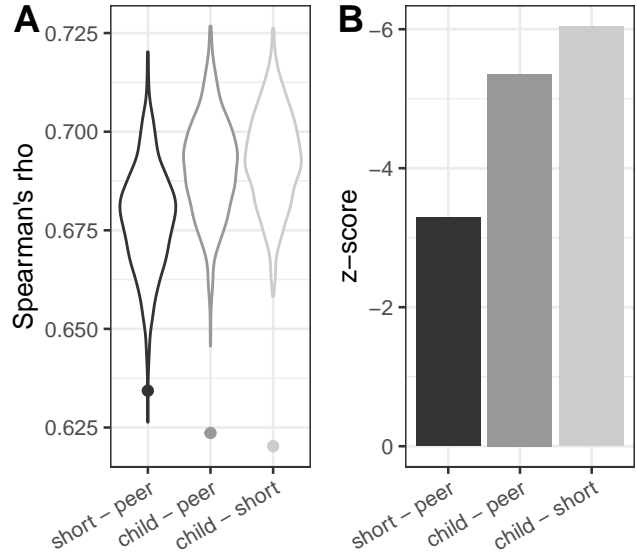


Figure 2: Representational Similarity Analysis. A) The representational similarity between each pair of conditions relative to the corresponding simulated null distribution. B) Standardized differences of the representational similarity of each pair of conditions from the mean of the simulated null, divided by the standard deviation of the null distribution.

and the mean of the simulated null distribution for each condition comparison (negative plotted up). Here, differences from zero indicate that the correlation of cues between conditions is lower than what you would expect from the simulated null distribution of random swaps between conditions.

Each permutation of condition comparison (Child vs. Peer, Child vs. Short, and Peer vs. Short) resulted in z-scores that were statistically different than zero, $z = -4.96, p < .001; z = -5.56, p < 0.001; z = -3.12, p < .001$, respectively. We also observed that responses from the child condition produce especially different representational similarity structure than either of the other two conditions. This is made evident by the larger z-scores when the condition contrast includes the child condition. Additionally, we can see that the distance between the true correlations and the distribution of null correlations is larger when the child condition is being compared to either of the other two conditions.

Discussion

The current study examined association behavior under a variety of induced contexts to get a better understanding of what mechanisms are being implemented when we use semantic control. We found that participants in the child condition provided more child-directed responses. Specifically, responses in this condition were shorter, more frequent and contextually diverse, and had a lower AoA than any other condition. This effect held over each response order, although responses tended to become less child-directed as more responses were made. This is likely due to the fact that second and third

responses have been thought to reflect different kinds of information and are presumably less automatic (De Deyne & Storms, 2008). For instance, De Deyne and Storms (2008) found that second and third responses in their free-association task had a higher proportion of adjectives and verbs than the first response.

The main finding in this study is that participants in the child condition provided shorter responses overall than participants in the short condition. This indicates that the participants in the child condition are exercising some form of semantic control other than general inhibition of long words and selection of appropriate short alternatives. Instead, it seems like participants are using pMTG-mediated semantic control to utilize the contextual information to influence how representations are structured and pulling out responses that fit that context. This is a suitable interpretation given the fact that the pMTG has been identified as a component in semantic integration where its purpose is to act as a mediator between frontal control systems and temporal representation (Jackson, 2021; Murphy et al., 2023). That is, this area sustains the context activated from frontal control regions and allows for the efficient access to stored information (Noonan et al., 2013).

Additionally, we found that the representational similarity structures between all conditions differed. That is, the semantic profile of responses in each condition conveyed different relational structure amongst the cue words. Interestingly, this effect is boosted in contrasts where the child condition is being compared to either of the other two conditions. This suggests that the similarity structure in this condition is especially more unique than the other two conditions. Again, this finding agrees with the interpretation that participants are flexibly shifting their semantic space to conform to the context. The set of responses for each cue within the child “space” is different than what you would expect to find in either of the other conditions.

This interpretation is further bolstered by the fact that participants in the short condition took roughly twice as long to complete the task than participants in the other two conditions. A similar study by Playfoot et al., 2018 saw marked increases in word association response times when participants were told to enact a rule as opposed to free-associate. They concluded that when participants are given a rule to follow, time to response is slowed by comparison of plausible alternatives before selecting your response. Connecting this to our short word “rule”, it is possible that participants in this condition were doing a similar kind of multi-associate comparison before making their response. This was not necessary, however, in the child and peer conditions because the context-mediated shifting of semantic space made it more likely that the first associate to come to mind was context-appropriate. One caveat, however, is that the Playfoot et al., 2018 obtained response time data for every response, not the cumulative time it took to complete the task.

Not only is this evidence for a semantic system that is flexible to an imposed context, but it is also evidence that we

can manipulate that context and have measured effects on behavior. Previous studies looking at the effects of context on semantics have focused on how sentences phrased to emphasize a particular meaning of a word, or single word primes change how we interact with homographs (Gorfein & Berger, 2000; Planchuelo et al., 2022; Zeelenberg et al., 2003). These studies have largely found that participants are more likely to gravitate towards meanings of words that coincide with the context they are provided with. This research contributes to those findings in that it elevates the context from a more “local” (i.e., sentence, single word) level to a more “situational” (i.e., global) level which is more common in our everyday lives as we shift between contexts.

Future studies will investigate what neural systems are at work when participants undergo a contextualized word association task. Specifically, it would be interesting to see whether the child word association task produces activation in the pMTG, whereas a more rule-based task (i.e. produce short words) would elicit activity in locations associated with more domain-general cognitive control (e.g., vIPFC).

In conclusion, the current study presents evidence that a global context manipulation can influence the semantic system beyond a simple rule-based schema. Instead, we posit that our pattern of results reflect a warping of semantic space that change what associations are most relevant given the context.

References

- Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., & Matsuo, A. (2018). Quanteda: An R package for the quantitative analysis of textual data. *Journal of Open Source Software*, 3(30), 774.
- Brysbaert, M., & New, B. (2009). Moving beyond kučera and francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, 41, 977–990.
- Cox, C. R., & Haebig, E. K. (2023). Child-oriented word associations improve models of early word learning. *Behavior Research Methods*, 1–31.
- Davey, J., Thompson, H. E., Hallam, G., Karapanagiotidis, T., Murphy, C., De Caso, I., Krieger-Redwood, K., Bernhardt, B. C., Smallwood, J., & Jefferies, E. (2016). Exploring the role of the posterior middle temporal gyrus in semantic cognition: Integration of anterior temporal lobe with executive processes. *NeuroImage*, 137, 165–177.
- De Deyne, S., & Storms, G. (2008). Word associations: Norms for 1,424 Dutch words in a continuous task. *Behavior Research Methods*, 40(1), 198–205.
- Fenson, L., et al. (2007). *MacArthur-bates communicative development inventories*. Paul H. Brookes Publishing Company Baltimore, MD.
- Gennari, S. P., MacDonald, M. C., Postle, B. R., & Seidenberg, M. S. (2007). Context-dependent interpretation of

- words: Evidence for interactive neural processes. *NeuroImage*, 35, 1278–1286.
- Gorfein, D. S., & Berger, S. (2000). The selection of homograph meaning: Word association when context changes.
- Hipp, R. D. (2018). SQLite.
- Jackson, R. L. (2021). The neural correlates of semantic control revisited. *NeuroImage*, 224, 117444.
- Jackson, R. L., Rogers, T. T., & Lambon Ralph, M. A. (2021). Reverse-engineering the cortical architecture for controlled semantic cognition. *Nature Human Behaviour*, 5, 774–786.
- Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 english words. *Behavior Research Methods*, 44, 978–990.
- Lawrence, M. A. (2016). *Ez: Easy analysis and visualization of factorial experiments* [R package version 4.4-0].
- Murphy, E., Forseth, K. J., Donos, C., Snyder, K. M., Rollo, P. S., & Tandon, N. (2023). The spatiotemporal dynamics of semantic integration in the human brain. *Nature Communications*, 14(1).
- Nili, H., Wingfield, C., Walther, A., Su, L., Marslen-Wilson, W., & Kriegeskorte, N. (2014). A Toolbox for Representational Similarity Analysis. *PLoS Computational Biology*, 10(4).
- Noonan, K. A., Jefferies, E., Visser, M., & Lambon Ralph, M. A. (2013). Going beyond inferior prefrontal involvement in semantic control: Evidence for the additional contribution of dorsal angular gyrus and posterior middle temporal cortex. *Journal of Cognitive Neuroscience*, 25(11), 1824–1850.
- Planchuelo, C., Buades-Sitjar, F., Hinojosa, J. A., & Duñabeitia, J. A. (2022). The nature of word associations in sentence contexts. *Experimental Psychology*, 69, 104–110.
- Playfoot, D., Balint, T., Pandya, V., Parkes, A., Peters, M., & Richards, S. (2018). Are word association responses really the first words that come to mind? *Applied Linguistics*, 39(5), 607–624.
- R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria.
- Snijders, T. M., Petersson, K. M., & Hagoort, P. (2010). Effective connectivity of cortical and subcortical regions during unification of sentence structure. *NeuroImage*, 52, 1633–1644.
- Zeelenberg, R., Pecher, D., Shiffrin, R. M., & Raaijmakers, J. G. W. (2003). Semantic context effects and priming in word association.