SPAML: Semantic Priming Across Many Languages

\* Corresponding author: Erin M. Buchanan ([ebuchanan@harrisburgu.edu](mailto:ebuchanan@harrisburgu.edu))

Erin M. Buchanan, Harrisburg University, Harrisburg, PA, USA, ebuchanan@harrisburgu.edu

Kelly M. Cuccolo, Alma College, Alma, MI, USA, cuccolokm@alma.edu

Nicholas Coles, Stanford University, CA, USA, ncoles797@gmail.com

Aishwarya Iyer, Montfort College, KA, India, iyeraishwarya.work@gmail.com

Neil Lewis Jr., Cornell University, NY, USA, nlewisjr@cornell.edu

Kim Peters, University of Exeter, UK, K.O.Peters2@exeter.ac.uk

Niels van Berkel, Aalborg University, Denmark, nielsvanberkel@cs.aau.dk

Anna E. van ’t Veer, Leiden University, Netherlands, a.e.van.t.veer@fsw.leidenuniv.nl

Jack E. Taylor, School of Psychology and Neuroscience, University of Glasgow, UK, j.taylor.3@research.gla.ac.uk

Maria Montefinese, IRCCS San Camillo Hospital, Venice, Italy, maria.montefinese@gmail.com

K. D. Valentine, Massachusetts General Hospital, Boston, MA, USA; Harvard Medical School, Boston, MA, USA, kvalentine2@mgh.harvard.edu

Nicholas P. Maxwell, University of Southern Mississippi, Hattiesburg, MS, USA, nicholas.maxwell@usm.edu

Martin R. Vasilev, Bournemouth University, Department of Psychology, United Kingdom, mvasilev@bournemouth.ac.uk

Andrei Dumbravă, G.Georgescu Institute of Cardiology Iași Romania; Alexandu Ioan Cuza University, Iași Romania, andrei.dumbrava@uaic.ro

Nienke Böhm, Vrije Universiteit Amsterdam, The Netherlands, Amsterdam, The Netherlands, nienkebohm@hotmail.com

Natalia Irrazabal, Universidad de Palermo - National Scientific and TechnicalÂ Research Council - Argentina, nirrazabal@psi.uba.ar

María Fernández-López, University of València, Spain, maria.fernandez@uv.es

Susana Ruiz-Fernandez, FOM University of Applied Sciences, Essen, Germany ; Leibniz-Institut für Wissensmedien, Tübingen, Germany; LEAD Research Network, Eberhard Karls University, Tübingen, Germany, susana.ruiz-fernandez@psychology-research.de

Mehmet Peker, Ege University, Turkey, mehmet.peker@ege.edu.tr

Jackson G. Lu, Massachusetts Institute of Technology, USA, lu18@mit.edu

Ulrich S. Tran, Department of Cognition, Emotion, and Methods in Psychology, Faculty of Psychology, University of Vienna, Vienna, Austria, ulrich.tran@univie.ac.at

Dina El-Dakhs, Prince Sultan University, Saudi Arabia , ddakhs@psu.edu.sa

David Moreau, University of Auckland, NZ, d.moreau@auckland.ac.nz

Suzanne L. K. Stewart, School of Psychology, University of Chester, United Kingdom, s.stewart@chester.ac.uk

Luisa Batalha, Australian Catholic University, Australia , luisa.batalha@acu.edu.au

Raluca D. Szekely-Copîndean, Romanian Academy, Cluj-Napoca, Romania, ralucaszekely@academia-cj.ro

Roxana E. Morariu, Babeș-Bolyai University, Cluj, m.roxana124@gmail.com

Hendrik Godbersen, FOM University of Applied Sciences, Essen, Germany, hendrik.godbersen@fom.de

Anna O. Kuzminska, University of Warsaw, Faculty of Management, Poland, a.kuzminska@uw.edu.pl

Patricia Garrido-Vásquez, University of Concepción, Chile, patricia.garrido@udec.cl

Miroslav Sirota, University of Essex, United Kingdom, msirota@essex.ac.uk

Chin Wen Cong, Department of Psychology and Counselling, Faculty of Arts and Social Science, Universiti Tunku Abdul Rahman (UTAR), Kampar, Perak, Malaysia, chinwencong@outlook.com

Fredrik Sigfrids, Åbo Akademi University, Finland, fredrik.sigfrids@abo.fi

Glenn P. Williams, School of Psychology, Faculty of Health Sciences and Wellbeing, University of Sunderland, United Kingdom, glenn.williams@sunderland.ac.uk

Max Korbmacher, Department of Biological and Medical Psychology, University of BergenÂ , Norway, Max.korbmacher@gmail.com

Izuchukwu L. G. Ndukaihe, Alex Ekwueme Federal University Ndufu-Alike, Nigeria, izumario@yahoo.co.uk

Krystian Barzykowski, Institute of Psychology, Jagiellonian University, Krakow, Poland, krystian.barzykowski@uj.edu.pl

Claudia Mazzuca, BalLab (Body, Action, Language Lab), Sapienza University of Rome, Italy, mazzuca.claudia@gmail.com

William E. Davis, Wittenberg University, USA, davisw4@wittenberg.edu

Gerit Pfuhl, Department of Psychology, UiT The Arctic University of Norway, gerit.pfuhl@uit.no

Zoran Pavlović, University of Belgrade, Faculty of Philosophy, Department of Psychology, Serbia, Europe, zoran.pavlovic@f.bg.ac.rs

Philip A. Grim II, Harrisburg University, Harrisburg, PA USA, pgrim@harrisburgu.edu

Leticia Micheli, Institute of Psychology, Würzburg University, Würzburg, Germany, leticia.micheli@uni-wuerzburg.de

Debora I. Burin, Universidad de Buenos Aires, Facultad de Psicologi­a / CONICET Argentina, dburin@psi.uba.ar

David C. Vaidis, Université de Paris, France, david.vaidis@u-paris.fr

Gerald H. Vineyard, University of East London, UK, gerald.vineyardofyork@gmail.com

Juan C. Oliveros-Chacana, Centro de Investigación en Ciencias Cognitivas, Facultad de Psicología, Universidad de Talca, Chile, juan.oliveros@utalca.cl

Paulo Roberto dos Santos Ferreira, UFGD, Brazil, paulorobertosantosferreira@gmail.com

Carlota Batres, Franklin and Marshall College, USA, cbatres@fandm.edu

Ekaterina Pronizius, Department of Cognition, Emotion, and Methods in Psychology, Faculty of Psychology, University of Vienna, Austria, ekaterina.pronizius@univie.ac.at

Steven Verheyen, Department of Psychology, Education and Child Studies, Erasmus University Rotterdam, Post Box 1738 3000 DR Rotterdam, The Netherlands, verheyen@essb.eur.nl

Vanessa Era, Department of Psychology, Sapienza University, Rome, Italy; IRCCS Fondazione Santa Lucia, Rome, Italy, vanessa.era90@gmail.com

Chris Isloi, Independent Scientist, barbu.chris@gmail.com

Aslan Karaaslan, Ege University, Turkey, aslankaraaslan@gmail.com

Mahmoud M. Elsherif, University of Birmingham, Department of Psychology, Birmingham, United Kingdom; University of Leicester, Department of Neuroscience, Psychology and Behaviour, Leicester, United Kingdom, mahmoud.medhat.elsherif@gmail.com

Sezin Öner, Department of Psychology, Kadir Has University, Istanbul, Turkey, sezin.oner@khas.edu.tr

John Protzko, Central Connecticut State University, USA, protzko@gmail.com

Magdalena Senderecka, Institute of Philosophy, Jagiellonian University, Krakow, Poland, magdalena.senderecka@uj.edu.pl

Alexandra I. Kosachenko, Ural Federal University, Ekaterinburg, Russia, alleshch7@gmail.com

Anna Maria Borghi, Department of Dynamic and Clinical Psychology and Health Studies, Sapienza University of Rome, Italy, anna.borghi@uniroma1.it

Belgüzar Nilay Türkan, Pamukkale University, Department of Psychology, Denizli, Turkey, bturkan@pau.edu.tr

Thomas Rhys Evans, School of Psychology and Counselling, University of Greenwich, England, UK, thomas.evans@greenwich.ac.uk

Sandra J. Geiger, Environmental Psychology, Department of Cognition, Emotion, and Methods, Faculty of Psychology, University of Vienna, 1010 Vienna, Austria, sandra.geiger@univie.ac.at

Bastian Jaeger, Vrije Universiteit Amsterdam, The Netherlands, bxjaeger@gmail.com

Jordan W. Suchow, School of Business, Stevens Institute of Technology, U.S.A., jws@stevens.edu

Dermot Lynott, Department of Psychology, Maynooth University, Ireland, dermot.lynott@mu.ie

Joseph P. McFall, State University of New York at Fredonia, United States, joseph.mcfall@fredonia.edu

Flavio Azevedo, Institute of Communication Science, Friedrich Schiller University, Jena, Germany, flavio.azevedo@uni-jena.de

Tijana Vesić Pavlović, University of Belgrade, Faculty of Mechanical Engineering, Serbia, Europe, tvesic@mas.bg.ac.rs

Emily Higgins, Dublin, Ireland, emilyhiggins17@gmail.com

Ewa Szumowska, Institute of Psychology, Jagiellonian University, Krakow, Poland, ewa.szumowska@uj.edu.pl

Luis Carlos Pereira Monteiro, Neuroscience and Cell Biology Graduate Program, Institute of Biological Sciences, Federal University of Pará, Belém, Pará, Brazil, luis.monteiro@icb.ufpa.br

Nwadiogo Chisom ARINZE, Alex Ekwueme Federal University Ndufu-Alike, Nigeria, arinzenwadiogo@gmail.com

Asil Ali Özdoğru, Department of Psychology, Üsküdar University, İstanbul, Turkey, asil.ozdogru@uskudar.edu.tr

Bruno Schivinski, School of Media and Communication, RMIT University, Australia, bruno.schivinski@rmit.edu.au

Cameron Brick, University of Amsterdam, Department of Psychology, PO Box 15900, 1001 NK Amsterdam, Netherlands, c.brick@uva.nl

Claus Lamm, Department of Cognition, Emotion, and Methods in Psychology, Faculty of Psychology, University of Vienna, Vienna, Austria, claus.lamm@univie.ac.at

Oguz A. Acar, City, University of London, United Kingdom, oguz.acar@city.ac.uk

Ettore Ambrosini, Department of Neuroscience, University of Padova, Italy ; Department of General Psychology, University of Padova; Padova Neuroscience Center, University of Padova, ettore.ambrosini@unipd.it

Leigh Ann Vaughn, Ithaca College, USA, Lvaughn@ithaca.edu

Patrí­cia Arriaga, Iscte-Instituto Universitá¡rio de Lisboa, CIS-IUL Lisbon, Portugal, patricia.arriaga@iscte-iul.pt

Ogeday Çoker, Pamukkale University, Department of Psychology, Denizli, Turkey, ocoker@pau.edu.tr

Zainab A. Alsuhaibani, Imam Mohammad Ibn Saud Islamic University (IMSIU), Saudi Arabia, alsuhaibani.z@gmail.com

Carmel A Levitan, Department of Cognitive Science, Occidental College, USA, levitan@oxy.edu

Jan Philipp Röer, Department of Psychology and Psychotherapy, Witten/Herdecke University, Witten, Germany, jan.roeer@uni-wh.de

Barnaby J.W. Dixson, School of Health and Behavioural Sciences, University of the Sunshine Coast, QLD, 4502, Australia, bdixson@usc.edu.au

Maital Neta, Department of Psychology, University of Nebraska-Lincoln, mneta2@unl.edu

Chiara Fini Fini, Department of Dynamic and Clinical Psychology and Health Studies, Sapienza University of Rome, Italy, chiara.fini@uniroma1.it

Théo Besson, Université de Paris, France, theo.besson@parisdescartes.fr

Martín Martínez, University of Navarra, Spain, tininmartinez@gmail.com

Melanie Labusch, University of València, Spain, melanie.labusch@uv.es

Marietta Papadatou-Pastou, National and Kapodistrian University of Athens, Athens, Greece, marietta.papadatou-pastou@seh.oxon.org

Kathleen Schmidt, Southern Illinois University, United States, kathleen.schmidt@siu.edu

Ewa Ilczuk, Institute of Psychology, Jagiellonian University, Krakow, Poland, ewa.ilczuk@student.uj.edu.pl

Anjali Thapar, Bryn Mawr College, USA, athapar@brynmawr.edu

Ernest Baskin, Saint Joseph’s University, USA, ebaskin@sju.edu

Kate E. Mulgrew, School of Health and Behavioural Sciences, University of the Sunshine Coast, QLD, Australia, kmulgrew@usc.edu.au

Muhammad Mussaffa Butt, GC University, Lahore, Pakistan, mussaffa@gmail.com

Jacek Buczny, Department of Experimental and Applied Psychology, VU Amsterdam, The Netherlands, jbuczny@gmail.com

Yuri G. Pavlov, Ural Federal University, Ekaterinburg, Russia, pavlovug@gmail.com

Manuel Perea, University of València, Spain, manuel.perea@uv.es

Ali H. Al-Hoorie, Royal Commission for Jubail and Yanbu, Jubail Industrial City, Saudi Arabia, hoorie\_ali@hotmail.com

Dauren Kasanov, Department Of Psychology, Ural Federal University, Ekaterinburg, Russian Federation, dauren.kasa@gmail.com

Jacob J. Keech, School of Health and Behavioural Sciences, University of the Sunshine Coast, QLD, Australia, jkeech@usc.edu.au

Bradley J. Baker, Temple University, Philadelphia, PA, USA, bradley.baker@temple.edu

Peter R. Mallik, Ashland University, Department of Psychology, USA, pmallik@ashland.edu

Savannah C. Lewis, Ashland University, OH, USA, Slewis5920@gmail.com

**Abstract**

Semantic priming has been studied for nearly 50 years across various experimental manipulations and theoretical frameworks. These studies provide evidence of cognitive underpinnings of the structure and organization of semantic representation in both healthy and clinical populations. In this registered report, we propose to create a large database of semantic priming values, alleviating the sample size and limited language issues with previous studies in this area. Consequently, this database will include semantic priming data across multiple languages using an adaptive sampling procedure. This study will test the size of semantic priming effect and its variability across languages. Results will support semantic priming when reduced response latencies are found for related word-pair conditions in comparison to unrelated word-pair conditions. Differences in semantic priming across languages will be supported when priming effect confidence intervals do not overlap.

SPAML: Semantic Priming Across Many Languages

Semantic priming is a well-studied cognitive phenomenon whereby participants are shown a cue word followed by either a related or unrelated target word[1](https://www.zotero.org/google-docs/?3g451z). Semantic priming is defined as the decrease in response latency (i.e., reduced linguistic processing or facilitation) for target words that are related to their cue words in comparison to unrelated cue words[1](https://www.zotero.org/google-docs/?WN1tMB). Semantic priming research spans nearly 50 years of study as a tool to investigate cognitive processes, such as word recognition, and to elucidate the structure and organization of knowledge representation[2](https://www.zotero.org/google-docs/?90WY0P), often by using results from these studies to develop theoretical and computational models that capture empirical effects[3–6](https://www.zotero.org/google-docs/?l86DSG). Priming has also been used in studies of attention[7,8](https://www.zotero.org/google-docs/?As989I), studies of bi/multilingual persons[9,10](https://www.zotero.org/google-docs/?LfdLXq), on people with psychological difficulties[11–13](https://www.zotero.org/google-docs/?NFrcLP), and in a large body of neuroscience studies[14–16](https://www.zotero.org/google-docs/?vNDl6v). The purpose of this study is to leverage the power and network of the Psychological Science Accelerator (PSA)[17](https://www.zotero.org/google-docs/?QKMCSI) to create a cross-linguistic normed dataset of semantic priming, paired with other useful psycholinguistic variables (e.g., frequency, familiarity, concreteness). The PSA is a large network of research laboratories committed to large-scale data collection and open science principles.

Experimental psychologists have long understood that the stimuli in research studies are of great importance, and that controlled sets of normed information hold significant value for study control and allow for precision in measurement of effects. Often, stimuli are created in small pilot studies and then reused in many subsequent projects. However, both Lucas[18](https://www.zotero.org/google-docs/?OVYO17) and Hutchison[19](https://www.zotero.org/google-docs/?QyJdbw) provided evidence that these small pilot data should be carefully interpreted given larger, more reliable datasets. In recent years, researchers have begun to more frequently publish large datasets with experimental stimuli for reuse in future work[20](https://www.zotero.org/google-docs/?KLSQ4B). These datasets include lexical frequency[21,22](https://www.zotero.org/google-docs/?EDtfWY), large collections of text (e.g., corpora)[23](https://www.zotero.org/google-docs/?DE52fT), behavioral stimuli such as priming[24](https://www.zotero.org/google-docs/?G9NBKR) or response latencies,[25,26](https://www.zotero.org/google-docs/?xVr4ZI) and subjective ratings from participants on emotion[27–29](https://www.zotero.org/google-docs/?GIgMvs), concreteness[30](https://www.zotero.org/google-docs/?zaMAMH), or familiarity[31](https://www.zotero.org/google-docs/?mwG3cm). Advances in computational capability, the growth of large-scale online data collection, and the focus on replication and reproducibility have advanced this research area. The importance of normed stimuli for research cannot be overstated. Not only do they provide methodological standardization for studies using the stimuli, the stimuli themselves can be studied to gain insight into cognitive architecture and processes, such as attention, memory, perception, and language comprehension or production[32–35](https://www.zotero.org/google-docs/?UQyItt).

These datasets provide a wealth of information for studies on semantic priming. Facilitation in priming is based chiefly on semantic similarity or the related word-pair condition as contrasted to the unrelated word-pair condition. Traditionally, word-pairs were simply grouped into pairs that were face-value similar (DOG-CAT) and unrelated (DOG-SPOON), which was determined through pilot studies where word-pairs provided the expected statistical results. However, for reproducibility and methodological control, semantic similarity values should be defined before the results are known[36](https://www.zotero.org/google-docs/?f8fUJr). Semantic similarity has various conceptual and computational definitions that all generally describe the shared meaning between two words or texts[5](https://www.zotero.org/google-docs/?DoeKmM). The most common forms of similarity are feature-based similarity (i.e., number of shared features between words)[37–39](https://www.zotero.org/google-docs/?wbaOjo), association strength (i.e., the probability of a first word eliciting a second word when simply shown the first word)[33,40](https://www.zotero.org/google-docs/?jIk9MP), or text co-occurrence (i.e., words are similar because they appear in the same types of texts)[41–43](https://www.zotero.org/google-docs/?DJI9f3). Each of these computational definitions of similarity can be calculated from normed datasets or text corpora to provide a continuous measure of similarity distance from 0 (unrelated) to 1 (perfectly related).

The Semantic Priming Project comprised both large-scale database collection and a semantic priming study that used defined stimuli to create related word pairs[24](https://www.zotero.org/google-docs/?BQ1Avj). This project provided data for lexical decision and naming tasks for 1,661 English words and non-words, along with other psycholinguistic measures for future research. The results of the Semantic Priming Project showed 23 ms to 25 ms decreases in word response latencies (i.e., lexical decision or naming speed) for the related word-pair conditions compared to unrelated word-pair conditions. The proposed study seeks to expand this dataset and address a few key limitations with the Semantic Priming Project. First, Heyman et al.[44](https://www.zotero.org/google-docs/?Feucen) explored the reliability of item-level priming effects from the Semantic Priming Project, deeming them mostly unreliable. This result contrasts with Hutchison et al.’s[45](https://www.zotero.org/google-docs/?DGUy7d) study which demonstrated that priming effects can be predicted at the item-level, albeit with a smaller dataset. Heyman et al.[44](https://www.zotero.org/google-docs/?tr9BBI) noted that the required sample size necessary for reliable priming effects was much larger than the sample size used in the study, potentially explaining the differences between results as well as demonstrating the need for a larger dataset.

The second key limitation of the Semantic Priming Project is the data overlap with other normed datasets[46](https://www.zotero.org/google-docs/?U9D4B5). For example, if a researcher were to use similarity values[39](https://www.zotero.org/google-docs/?0AKhiV), concreteness ratings[30](https://www.zotero.org/google-docs/?jkW5Uy), valence[27,28](https://www.zotero.org/google-docs/?rqq4qW), and subjective/objective age of acquisition[31,47,48](https://www.zotero.org/google-docs/?WB2Udv) to predict priming, they may be limited to only a few hundred word-pairs because of the bounded overlap between each separate dataset. Last, the Semantic Priming Project only contains English data, thus limiting our understanding of priming and knowledge representation across languages and culture. Even with the increase in publication of normed datasets in non-English languages[20](https://www.zotero.org/google-docs/?4OIk9H), conducting cross-linguistic studies on the same concepts is challenging, as large-scale data in this area are sparse.

However, using newer computational techniques[49,50](https://www.zotero.org/google-docs/?3NDiur) and recently published corpora[23,51](https://www.zotero.org/google-docs/?sZZrt0), a broader coverage dataset in up to 44 languages is possible. Therefore, this study aims to provide data that complements and extends the published data, which would encourage research on methodology, item characteristics, models, cross-linguistic consistency in priming, and other theoretical areas that semantic priming has been applied to previously. The data will address the three limitations by increasing sample size, ensuring a complete dataset with multiple psycholinguistic variables, and expanding beyond the English language. From this openly shared data, two research questions will be assessed as detailed in Table 1:

1. Is semantic priming a non-zero effect? To assess this research question, we will examine the confidence interval of the semantic priming effect to determine if the lower limit of the confidence interval is greater than zero. Therefore, Hypothesis 1 predicts semantic facilitation with reduced response latencies for related word-pair conditions in relation to unrelated word-pair conditions. We will provide estimates across all languages and each language provided individually with corrections for the number of languages provided.
2. How does the semantic priming effect vary across languages? We will compare the confidence intervals of the semantic priming effect for each language to determine if they overlap. The effect will be considered equal when corrected confidence intervals overlap between languages and different when they do not overlap. For Hypothesis 2, we do not specify predicted directions for the effects but do expect meaningful differences between languages. It is logical to expect differences in language due to culture, orthography, alphabet, etc., and empirical data suggest meaningful differences between languages[52,53](https://www.zotero.org/google-docs/?mPIUtv).

This research crucially supplements the literature outlined above by focusing on several key components of psycholinguistic research. First, we will ensure adequate sampling for individual items by using accuracy in parameter estimation[54,55](https://www.zotero.org/google-docs/?vp3Ezd) to address the known reliability issues in item-level responding[44,56](https://www.zotero.org/google-docs/?ncpFMA) to support Hypothesis 1. The items will be selected using new computational techniques for addressing semantic similarity[49,50](https://www.zotero.org/google-docs/?87YvQh) with recently available large corpora of movie subtitles[23](https://www.zotero.org/google-docs/?FHPr4j) to appropriately match comparable items across languages. As noted in Buchanan et al.[20](https://www.zotero.org/google-docs/?Xji6Jy), research in non-English languages is expanding; however, stimuli matching is still sparse across published databases. By using large corpora, we will ensure items are matched not only in their similarity levels, but also for their frequency of use. Thus, differences in priming can be attributed to differences in linguistic structure or culture, rather than translation or poor item matching, supporting Hypothesis 2.

**Method**

**Ethics Information**

We will not collect any identifiable private or personal data as part of the experiment. This project was approved by Harrisburg University of Science and Technology conforming to all relevant ethical guidelines and the Declaration of Helsinki, with special care to conform to the General Data Protection Regulation (GDPR; eugdpr.org). Each research lab will obtain local ethical review, rely on the ethical review provided by Harrisburg University, or provide evidence of no required ethical review. The IRB approvals are available on our Open Science Framework (OSF) page: <https://osf.io/wrpj4/>. Participants may be compensated for their participation by course credit or payment depending on individual lab resources. No exclusion criteria will be used.

**Power Analysis**

For our power analysis, we first detail a background on how we plan to estimate sample size, explain accuracy in parameter estimation, provide two simulations based on previous research, and the final proposed sample size. We end this section by specifying why this procedure is superior to previous methods and the requirements for publication.

**Background**

One concern is how to estimate the sample size required for cue-target pairs, as the previous literature indicates variability in their results[44](https://www.zotero.org/google-docs/?o9mc2J). Sample sizes of *N* = 30 per study have often been used in an attempt to at least meet some perceived minimum criteria for the central limit theorem. We will focus on the lexical decision task for our procedure, wherein participants are simply asked if a concept presented to them is a word (NURSE) or non-word (LURSE). The dependent variable in this study is response latency, and we will use lexical decision data from the English Lexicon Project[25](https://www.zotero.org/google-docs/?BWo2L7) and the Semantic Priming Project[24](https://www.zotero.org/google-docs/?Je45qJ) to estimate the minimum sample size necessary for each item, as previous research has suggested an overall sample size may lead to unreliability in the item-level responses[44](https://www.zotero.org/google-docs/?SEy9WO). The English Lexicon Project contains lexical decision task data for over 40,000 concepts, while the Semantic Priming Project includes 1,661 target words.

**Accuracy in parameter estimation (AIPE)**

***AIPE description.*** In this approach, one selects a minimum sample size, a stopping rule, and a maximum sample size. A minimum sample size will be defined for all items based on data simulation below. For the stopping rule, we focused on finding a confidence interval around a parameter that would be “sufficiently narrow”[54,55,57](https://www.zotero.org/google-docs/?lkrp2M). These parameters are often tied to the statistical test or effect size for the study, such as correlation or contrast between two groups. In this study, we will pair accuracy in parameter estimation with a sequential testing procedure to adequately sample each item, rather than estimate an overall effect size. Therefore, we will use the previous lexical decision data to determine our sufficiently narrow confidence by finding a generalized standard error one should expect for well measured items. After the minimum sample size, each item’s standard error will be assessed to determine if the item has met the goals for accuracy in parameter estimation as our stopping rule. If so, the item will be sampled at a lower probability in relation to other items until all items reach the accuracy goals or a maximum sample size determined by our simulations below.

***Estimates from the English Lexicon Project*.** First, the response latency data for the English Lexicon Project were *z*-scored by participant and session as each participant has a somewhat arbitrary average response latency[58](https://www.zotero.org/google-docs/?42cOnG). The data was then subset for only real word trials that were correctly answered. The average sample size before data reduction was 32.69 (*SD =* 0.63) participants with an average retention rate of 84% and 27.41 (*SD* = 6.43) participants after exclusions. The retention rates are skewed due to the large number of infrequent words in the English Lexicon Project, and we will use the median retention rate of 91% for later sample size estimations. The median standard error for response latencies in the English Lexicon Project was 0.14 and the mean was 0.16. Because the retention rates are variable across items, we also calculated the average standard error for items that retained at least 30 participants at 0.12. This standard error rate would represent our potentially stopping rule.

The data was then sampled with replacement to determine the sample size that would provide that standard error value. One hundred words within the data were randomly selected, and samples starting at *n* = 5 to *n* = 200 were selected (increasing in units of five). The standard error for each of these samples was then calculated for the simulation, and the percent of samples with standard errors at or less than the estimated population value was then tabulated. In order to achieve 80% of items at or below the proposed standard error, we will need approximately 50 participants per word. This value will be used as our minimum sample size for a lexical decision task, and the accuracy standard error level will potentially be set at 0.12.

***Estimates from the Semantic Priming Project.*** This same procedure was examined with the Semantic Priming Project’s lexical decision data. The priming response latencies are expected to be variable, as this priming strength should be predicted by other psycholinguistic variables, such as word relatedness. Therefore, we aim to achieve an accurate representation of lexical decision times, from which priming can then be calculated. We used this data paired with the English Lexicon Project to account for the differences in a lexical decision only versus priming focused task. The average standard error in the Semantic Priming Project was less at 0.06, likely for two reasons: the data in the Semantic Priming Project are generally frequent nouns and only 1,661 concepts, as compared to the 40,000 in the English Lexicon Project. The retention rate for the Semantic Priming Project is less skewed than the English Lexicon Project at a median of 97% and mean of 96%. Using the same sampling procedure, we estimated sample sizes of *n* = 5 to *n* = 400 participants increasing by units of 5. In this scenario, we find the maximum sample size of 320 participants for 80% of the items to reach the smaller standard error of 0.06. Therefore, we will use 320 as our maximum sample size, and the average of the two standard errors found as our stopping rule, i.e., 0.09.

***Final sample size.*** Given our minimum, maximum, and stopping rule, we then estimated the final sample size per language based on study design characteristics. Participants would complete approximately 600 lexical decision trials per session and half of these concepts would be real words that are the target of this sample size analysis. Therefore, the target number of items (*n* = 1000 concepts) was multiplied by the minimum/maximum sample size, the data retention rate (a conservative estimate of 90%), and conditions (related word pair versus unrelated word pair) and divided by the total number of usable lexical decision trials per participant. The final estimate for sample size per language is 741 to 4741. The complete code and description of this process are detailed at: <https://osf.io/rxgkf/>.

This sample size estimation represents a major improvement from previous database collection studies, as many have used the traditional *N* = 30 to guess at minimum sample size. Because the variability of the sample size is quite large, we will employ a stopping procedure to ensure participant time and effort is maximized, and data collection is optimized. The minimum sample size will be 50 participants per word and the maximum will be 320, which results in 741 to 4,741 participants per language based on expected usable trials. Therefore, the total sample size will range from 3,705 to 23,705 participants for five languages. After 50 participants, each concept will be examined for standard error, and data collection for that concept will be decreased in probability when the standard error reaches our average criterion of 0.09. Item probability for selection will also be decreased when they reach the maximum proposed sample size (*n* = 320). This process will be automated online and checked in a scheduled subroutine.

While 44 languages have been identified for possible data collection, we plan to first publish the data when five languages have reached the appropriate sample size as outlined above based on recruitment of PSA partner labs. To date, we have recruited more than 100 researchers in 19 languages. Other data will be included in the online open dataset with notation on the progress towards completion. The five selected languages will be based on PSA lab recruitment and feasibility to reach sample size. We will target to include English, at least one Asian language, and the others will likely be European languages, also spoken in other world regions (e.g., Spanish).

**Materials**

The following details the important facets of the materials. We will first explain the types of word-pair conditions in a semantic priming study (i.e., related, unrelated, and non-word). Next, we will detail how the related word-pair conditions were created using the OpenSubtitles corpora, new computational modeling techniques, and the selection procedure.

**Word-pair conditions**

In a semantic priming study, there are three types of word-pair conditions. In the related word-pair condition, cue-target pairs are chosen for their similarity or relatedness. Cosine distance is similar to correlation in representing relatedness; however, cosine distance is always positive. Therefore, a cosine distance of 1 represents the same numeric vectors (perfect similarity), while a cosine distance of 0 represents no similarity between vectors. To create the unrelated condition, cue-target pairs are shuffled so that the cue word is combined with a target word with which it has a negligible cosine distance similarity.

Finally, non-words pair conditions are created by changing one to two letters in a cue or target word to create a nonsense word (NURSE →LURSE), with the stipulation that they must be pronounceable and not pseudo-homophones (i.e., wherein the pronunciation sounds like a real word, KEEP → KEAP)[59](https://www.zotero.org/google-docs/?CmCDqd). All the character bigrams (i.e., two letter combinations, KITE → KI IT TE) for the cue-target pairs will be tabulated to determine valid bigram combinations for a language. A random letter from each word will be selected and the possible bigrams preceding and following this letter will be examined. For instance, from the word KITE, the letter T may be selected, and all bigrams starting with I and ending with E will be included as candidates for selection. A random selection of the union of these two bigram options will be generated, and that letter will replace the selected letter. Therefore, KITE may become KILE as IL and LE are valid English bigrams. Native speakers will check the selected concept’s non-word generation in order to ensure that each non-word is pronounceable and that the random generation did not produce another real word like CITE. In this case, the randomization will be generated until the non-word meets both these criteria. Special care will be taken with logographic languages (Japanese, Chinese, Korean, Hebrew), as these are often one character, and the bigram procedure described above will not suffice. We will consult with native speakers to change one stroke or radical such that the character(s) are a pronounceable word with no meaning. Only one non-word is presented in the non-word word-pair conditions.

**Similarity calculation**

***Corpora.*** As described in the introduction, the choice of related words based on similarity is key for the study. There are multiple measures of semantic similarity including the cosine between overlapping features[39](https://www.zotero.org/google-docs/?jPentL), free association probabilities[33,40,60](https://www.zotero.org/google-docs/?il95tW), and local/global coherence values from network models[35,61](https://www.zotero.org/google-docs/?8ZC2bi). However, the underlying data for these calculations is inconsistent across languages. Therefore, one solution is to use the data present in the OpenSubtitles datasets[23](https://www.zotero.org/google-docs/?W1kIEE) (i.e., a large collection of movie subtitles) to calculate word frequency and cosine distance similarity values. These datasets have been used to calculate word frequencies for the SUBTLEX projects, which have validated their use as strong predictors of cognitive related phenomena[21,62–69](https://www.zotero.org/google-docs/?fVYhpE). Cosine distance was selected over other similarity measures because of the availability of possible languages and models for this project, as described below.

The OpenSubtitles data includes 62 languages or language combinations (i.e., Chinese-English mix). We will use the 10,000 most frequent nouns, adjectives, adverbs, and verbs from each potential language without lemmatization (i.e., converting words into their dictionary form RUNS → RUN). The *udpipe* package[70](https://www.zotero.org/google-docs/?IkaYKW) is a natural language processing package that contains more than 100 treebanks to assist in part of speech tagging (i.e., labeling words as noun, verb, etc.), parsing (i.e., separating blocks of text into words and their relationship to other words in a text), and lemmatization. This package was selected for its large coverage of languages with reliable part-of-speech tagging. Cross-referencing the available languages in *udpipe* with the OpenSubtitles data allows for the possibility of 44 different languages in this project. See Figure 1 for the model selection process.

***Modeling.*** The *subs2vec* project[51](https://www.zotero.org/google-docs/?bn5NRy) used the OpenSubtitles data to create fastText[71](https://www.zotero.org/google-docs/?LlNeaG) computational representation for 55 languages. fastText is a distributional vector space model, an extension of word2vec[49,50](https://www.zotero.org/google-docs/?HTyPwi)**,** wherein each word in a corpus is converted to a vector of numbers that represents the relationship of that word to a number of dimensions. These dimensions can be imagined as a thematic or topic representation of the text. The relationship between these vectors represents the similarity between concepts, as words that have similar or related meanings will appear in similar places and dimensions in a text, and will, therefore, have similar numeric vectors[4,5](https://www.zotero.org/google-docs/?yCRyi2). We will use the existing models from *subs2vec* to extract related word concepts for the most frequent concepts identified using the top cosine distance between word vectors.

***Cue selection procedure.*** The procedure for stimuli selection can be viewed at <https://osf.io/2cmp4/> and is displayed graphically in Figure 1. If the language is available via *subs2vec*, the provided subtitle frequency counts will be examined. If the language has more than 50,000 unique concepts represented in the subtitle data, we will use the subtitle model only. If the subtitles do not provide enough linguistic information (i.e., fewer than 50,000 concepts in the corpus), we will use the combined Wikipedia and subtitle model CITE SUBS. *subs2vec* contains models with only the OpenSubtitles data, only Wikipedia for a given language, and a combined model of both. The subtitle data has shown to best represent a language[21,62](https://www.zotero.org/google-docs/?Pux4r1); however, not all subtitle projects contain a large enough corpus for the subtitles to cover the breadth of the possible concepts within that language (e.g., Afrikaans subtitles only represent approximately 18,000 words).

The selected token list will then be tagged for part-of-speech using *udpipe,* selecting tokens that are tagged as nouns, adjectives, adverbs, and verbs. All stopwords (i.e., commonly used words in a language with little semantic meaning such as THE, AN, OF), words with fewer than three characters, and words with numeric characters will be eliminated (i.e., 1 would be eliminated but not ONE). The stopword lists can be found in the *stopwords* package using the Stopwords ISO dataset[72](https://www.zotero.org/google-docs/?lYlnP6). This procedure will cover all but two languages in our list of 44 possible languages. For the final two languages, we will use *udpipe* to tag the OpenSubtitles directly and calculate word frequency. Additionally, fastText model using the same parameters as *subs2vec* will be trained for similarity calculation. The 10,000 most frequent concepts will be selected at this point.

***Target selection procedure.*** Using the fastText models for each language, we will select the top five cosine distance similarity values for each concept in each language, resulting in 50,000 possible cue-target pairs. These will be cross-referenced across languages using Google Translate. Native speakers will be recruited to ensure the accurate translation of word pairs using the PSA’s translation network. The related word pairs (*n* = 1,000) will be selected from the list using each cue only once, favoring pairs with translations in most languages. If a selected pair does not exist in a language, translation from a native speaker will be used to create that pair.

**Procedure**

We will describe the important components to the procedure in this section. First, we detail the implementation of the study, focusing on the timing software and adaptive stimuli section, as not all participants see all items. We then discuss the study procedure in order, as shown in Figure 2. First, participants will complete a demographic questionnaire, followed by the lexical decision task. We explain how our data compliments the Semantic Priming Project and finally, discuss additional data that we plan to collect.

**Implementation**

***Timing software.*** While participants will be naïve to the word pairings, the principal investigator will know the pair combinations during data collection and analysis. A small demonstration of the experiment can be found at: <https://psa007.psysciacc.org/>. The study will be programmed using lab.js[73](https://www.zotero.org/google-docs/?I0C8gI), which is an online, open-source, study-building software. Precise timing measurement is required for this study, and the lab.js team has documented the accuracy of measurement within their framework[74](https://www.zotero.org/google-docs/?f01nZ6), and previous work has shown no differences between lab and web-based data collection for response latencies[75](https://www.zotero.org/google-docs/?zGeBtI). In addition, SPALEX, a large lexical decision database in Spanish, was collected completely online[26](https://www.zotero.org/google-docs/?Ys9Jnw). We will recommend that research labs suggest Chrome as their browser for participants completing the study due to recommendations from the lab.js team. However, meta-information about the browser and operating system are saved when participants take the experiment to examine for potential implementation differences.

Participants will be directed to an online web portal to complete the study, and all data will be retained in the online platform with nightly backups to the server. Participants will be required to complete the study on a computer with a keyboard, rather than on a device with only a touch screen. This requirement allows for tracking of the display of the device which will indicate important aspects about screen size, browser, and timing accuracy. In order to enforce this requirement, participants will be asked to hit the spacebar to continue the study.

***Adaptive stimuli selection.*** At the start of data collection, all presented items will be randomly selected from the larger item pool by equalizing the probability of inclusion equal for all words and non-words (*p* = 1/1000 concepts). After the minimum sample size is collected, each word’s standard error will be checked to determine if the sample size for that item has reached our accuracy criteria. If so, the probability of sampling that item will be decreased by half. Once a concept has reached the maximum required sample size, the probability of sampling will also be decreased by half. This procedure will allow for random sampling of the items that still need participants without eliminating words from the item pool. Therefore, we will ensure that there are always words to randomly select from (i.e., to keep the same procedure and number of trials for all participants) and that the randomization is a sampled mix of words that reach accuracy quickly and words that need more participants (i.e., participants do not only see the unusual words at the end of data collection). Once all words have reached the stopping criteria or maximum sample size, the probabilities will be equalized. We have set minimum, maximum, and a stopping rule for the initial data collection; however, we will allow data collection after these have been reached and will post updates to the data using a DOI service to allow researchers to cite the specific dataset they used for their research (modeled after the Small World of Words Project[33](https://www.zotero.org/google-docs/?ihEkPu), which is ongoing). All data will be included in our dataset, and the analysis section describes how we will indicate potential data for exclusion. Therefore, data collection will occur in a repeated-measures design in which participants do not see all of the possible stimuli, but do see all the possible conditions (related, unrelated, and non-word pairs).

**Study Procedure**

***Demographics.*** Participants will be directed to select their first language, which will then direct them to the appropriate translation of the experiment. Participants will be asked to indicate their gender (i.e., male, female, other, prefer not to say), year of birth, and education level (i.e., none, elementary school, high school, bachelors, masters, doctorate) for demographic variables. A flow chart of the procedure is provided in Figure 2.

***Lexical Decision Task.*** Instructions on how to complete a lexical decision task will be shown on the next screen, followed by 10 practice trials. Each trial starts with a fixation cross (+) in the middle of the screen for 500 ms. The concept will then be displayed in the middle of the screen in uppercase san-serif 18-point font (i.e., Arial font, NURSE). On the bottom of the screen the possible responses will be shown as the traditional keys next to the *Shift* key depending on the most common keyboard layout for that language (i.e., Z and / on a QWERTY keyboard or > and - on a QWERTZ keyboard). These answer responses will also be presented in the reverse option, which will be randomly assigned by lab.js at the start of the study to counterbalance response options. Participants will make their choice for each concept, and then the next stimulus will appear with an intertrial interval of 500 ms (i.e., the time between the offset of the first concept response and onset of the next concept, when the fixation cross is showing). Responses will time out after 5 seconds and move on to the next trial. After 10 trials, participants will see the instruction screen again with a reminder that they will now be doing the real task.

After 100 trials, the participants will be shown a short break screen with the option to continue by hitting the spacebar after 10 seconds. After six blocks of 100 trials (600 words), the experiment will end with a thank you screen. On this screen, participants will indicate what type of credit they are receiving for the study (e.g., course credit, payment, no compensation), and they will be given instructions on how to indicate that they have completed the study to the appropriate lab. Participants will be allowed to take the study multiple times as items are randomly selected for inclusion. An estimate for the time required for the study is approximately 30 minutes inclusive of practice trials, instructions, and breaks. This estimate is based on previous studies of lexical decision times[25](https://www.zotero.org/google-docs/?FNiE4x), and pilot testing will be used to determine if the number of trials should be reduced to accommodate the 30-minute expected time.

***Comparison to the Semantic Priming Project.*** This procedure is a single stream lexical decision task wherein every concept (cue and target) is judged for lexicality (i.e., word/non-word). Many priming studies often present cue words for a short period of time prior to the presentation of target words for lexicality judgement. Evidence from the Semantic Priming Project suggests that the stimulus onset asynchrony (i.e., time between non-judged cue word and target word) does not affect overall priming rates (25 versus 23 ms for 200 ms and 1200 ms). Further, adding the lexicality judgment to each presented concept creates a less obvious link between cue and target to avoid potential conscious expectancy generation effects[76](https://www.zotero.org/google-docs/?X3eD3h). Even though they appear sequentially in the task, they are not explicitly paired by being a non-judged cue word followed by a judged target word. Therefore, this procedure varies from the data collected in the Semantic Priming Project; thus, extending their work to different conditions.

***Additional data.*** A primary goal of this project is to provide a complete dataset of priming and other important related linguistic variables. Lexical measures, such as length, frequency, part of speech, and the number of phonemes (i.e., sounds in a word) are easily created from the concept or the SUBTLEX projects. Subjective measures are concept characteristics that are rated by participants, such as age of acquisition[77–80](https://www.zotero.org/google-docs/?ocU0s2) (approximate age you learned a concept), imageability[81,82](https://www.zotero.org/google-docs/?yRz6iY) (how easy the concept comes to mind), concreteness[83](https://www.zotero.org/google-docs/?twN092) (how concrete is the concept), valence (how positive versus negative is the concept), arousal (how excited or calm a concept makes a person), dominance (the word denotes something that is weak/subordinate or strong/dominant)[27,29](https://www.zotero.org/google-docs/?rpRfkx), and familiarity (how well a person knows a concept)[84](https://www.zotero.org/google-docs/?QJfJTO). We will use published databases to create a master list of available values for each of these subjective ratings. For words that are missing these values in a target language, the same participants will be asked to provide ratings on a single metric (e.g., they would only see instructions for familiarity or arousal).

Each participant may be asked to provide 25-50 ratings of words, given the missing data present for a particular language, while also controlling for the length of the task to prevent fatigue in the experiment. These ratings will only be presented at the end of the experiment to prevent interactions with priming effects. We will use the available large databases of these variables to estimate a sample size necessary for these ratings using the same simulation procedure detailed above. The exact instructions for each task can be found at: https://osf.io/tcxk2/. Once subjective ratings for each item have reached an appropriate accuracy rating, that item will be discontinued from data collection, and the subjective rating task will be dropped once all items have been collected. The item and task selection for each participant is randomized. This section will be described as a bonus task, like a rating game, for participants, and they will be provided feedback on how their scores compare to others as an incentive to complete the ratings. This section will be discontinued when the data collection for these values is complete.

**Analysis Plan**

An example of the data and processing for English can be found at <https://osf.io/6jmzk/>.

**Trial level data**

Files containing the entire data from the experiment will be available for download from the experiment website. Each language will be saved in a separate file with an item specific trial identification number to allow for matching concepts across languages (i.e., CAT (English) → KATZE (German) → GATTA (Italian)). All data will be archived on our server, and we will use Zenodo (<https://zenodo.org/>) to release versions of the data with citable DOIs given the planned continuation of the project after the initial PSA support. Participants are expected to incorrectly answer trials, and these trials will be marked for exclusion. Further, computer errors or trials due to missing data (i.e., participant inattentiveness and timeout trials, internet disconnection, computer crashes) will be marked as such in the final data with missing values. No missing values will be imputed. The response latencies from each participant’s session will be *z*-scored in line with recommendations from Faust et al.[58](https://www.zotero.org/google-docs/?fGu7zA). We will not collect enough data to note if a person takes the experiment multiple times for privacy reasons, but as these would be considered different sessions, the recommended *z*-score procedure should control for participant variability at this level. Therefore, repeated participation would not be detrimental to data collection. Finally, participants' overall proportion of correct answers will be calculated, and participants who do not correctly answer at least 80% of 100 minimum trials will be excluded for item data, priming data, and analysis. The average error in the Semantic Priming Project ranged from 4% to 5%, and this criterion was chosen to include participants who were focused on the task.

**Item level data**

An item-level data file will also be prepared for publication and data release. The item file will contain lexical information about all stimuli calculated from the OpenSubtitles[23](https://www.zotero.org/google-docs/?2jJJUX) and *subs2vec*[51](https://www.zotero.org/google-docs/?HSfGZs) projects (length, frequency, orthographic neighborhood, bigram frequency, orthographic and phonographic Levenshtein distance). The descriptive statistics calculated from the trial level data will then be included: mean response latency, average standardized response latency, sample size, standard errors of response latencies, and accuracy rate. For mean and standard error calculation, the incorrect and missing trials will be excluded. No data will be excluded for being a potential outlier; however, we will recommend a cut-off criterion for absolute value *z*-score outliers at 2.5 and 3.0, and we will calculate these same statistics with those subsets of trials excluded. For all real words, the age of acquisition, imageability, concreteness, valence, dominance, arousal, and familiarity values will be included. These values do not exist for non-words.

**Priming data**

In a separate file, we will also prepare information about priming results. For each item, priming is defined as the average z-scored response latency when presented in the unrelated minus the related condition. Therefore, the timing for DOCTOR-NURSE would be subtracted from TREE-NURSE to indicate priming for the word NURSE. The similarity scores calculated during stimuli selection will be provided in this file, as well as other popular measures of similarity if they are available in that language. For example, semantic feature overlap norms are also available in Italian[85](https://www.zotero.org/google-docs/?pKAUFA), German[86](https://www.zotero.org/google-docs/?vHk4Pg), Spanish[26](https://www.zotero.org/google-docs/?kbfcpp), and Dutch[87](https://www.zotero.org/google-docs/?TUi2Ne)**.**

**Hypothesis 1**

Hypothesis information is presented in Table 1. Hypothesis 1 predicts semantic facilitation with reduced response latencies for related than unrelated words. Hypothesis 1 will be analyzed by calculating the average *z*-scored priming response latency for all languages and its 95% confidence interval. The priming response latency is calculated by taking the average of the unrelated pair *z*-scored response latency minus the related pair response latency within each item. Therefore, values that are positive and greater than zero (e.g., > 0.0001) indicate priming because the related pair had a faster response latency than the unrelated pair. We will determine support for Hypothesis 1 if the lower limit of the confidence interval is greater than zero (i.e., a directional comparison). This process will be repeated for average priming scores calculated without trials that were marked as 2.50 *z*-score outliers and 3.00 *z*-score outliers separately. Next, each language will be examined individually, using the same procedure by calculating the average scores and their confidence intervals. We will correct for the comparisons between languages by calculating the confidence interval at 1 - .05 / languages provided. Therefore, if five total languages are provided, we will calculate a 99% CI for these comparisons. The decision criteria will remain the same, and we will identify any differences in decisions based on outlier statistics (e.g., priming only occurs when X trials are removed in X language).

**Hypothesis 2**

Hypothesis 2 explores the extent to which these semantic priming effects vary across languages. Thus, the statistics from Hypothesis 1 will be used for this hypothesis, such that each language's confidence interval will be compared to that of every other language. The confidence intervals will be calculated based on the same correction procedure described above. Confidence intervals that overlap will be considered equal in their priming effect, while confidence intervals that do not overlap will be considered different levels of priming effects. This analysis will be repeated with the 2.50 *z*-score outliers and 3.0 *z*-score outliers excluded. We will depict the results from this analysis visually to denote the variation and any differences found between languages. Example code can be found at <https://osf.io/fbhr8/>.

**Data Availability**

All raw and processed data will be available for download from the website devoted to this project with backups provided on OSF and Zenodo.

**Code Availability**

All code used for study creation and delivery, data processing, and analyses will be available on OSF (https://osf.io/wrpj4/) and GitHub (https://github.com/SemanticPriming/SPAML).

References

[1. Meyer, D. E. & Schvaneveldt, R. W. Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology* **90**, 227–234 (1971).](https://www.zotero.org/google-docs/?qrzPMM)

[2. McNamara, T. P. *Semantic Priming*. (Psychology Press, 2005). doi:10.4324/9780203338001.](https://www.zotero.org/google-docs/?qrzPMM)

[3. Mandera, P., Keuleers, E. & Brysbaert, M. Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language* **92**, 57–78 (2017).](https://www.zotero.org/google-docs/?qrzPMM)

[4. Cree, G. S. & Armstrong, B. C. Computational Models of Semantic Memory. in *The Cambridge Handbook of Psycholinguistics* (eds. Spivey, M., McRae, K. & Joanisse, M.) 259–282 (Cambridge University Press, 2012). doi:10.1017/CBO9781139029377.014.](https://www.zotero.org/google-docs/?qrzPMM)

[5. McRae, K. & Jones, M. *Semantic Memory*. (Oxford University Press, 2013). doi:10.1093/oxfordhb/9780195376746.013.0014.](https://www.zotero.org/google-docs/?qrzPMM)

[6. Rogers, T. T. Computational Models of Semantic Memory. in *The Cambridge Handbook of Computational Psychology* (ed. Sun, R.) 226–266 (Cambridge University Press, 2001). doi:10.1017/CBO9780511816772.012.](https://www.zotero.org/google-docs/?qrzPMM)

[7. Frings, C., Schneider, K. K. & Fox, E. The negative priming paradigm: An update and implications for selective attention. *Psychon Bull Rev* **22**, 1577–1597 (2015).](https://www.zotero.org/google-docs/?qrzPMM)

[8. Spruyt, A., De Houwer, J., Everaert, T. & Hermans, D. Unconscious semantic activation depends on feature-specific attention allocation. *Cognition* **122**, 91–95 (2012).](https://www.zotero.org/google-docs/?qrzPMM)

[9. McDonough, K. & Trofimovich, P. *Using Priming Methods in Second Language Research*. (Routledge, 2011). doi:10.4324/9780203880944.](https://www.zotero.org/google-docs/?qrzPMM)

[10. Singh, L. One World, Two Languages: Cross-Language Semantic Priming in Bilingual Toddlers. *Child Dev* **85**, 755–766 (2014).](https://www.zotero.org/google-docs/?qrzPMM)

[11. Copland, D. The basal ganglia and semantic engagement: Potential insights from semantic priming in individuals with subcortical vascular lesions, Parkinson’s disease, and cortical lesions. *J Int Neuropsychol Soc* **9**, 1041–1052 (2003).](https://www.zotero.org/google-docs/?qrzPMM)

[12. Haverkort, M. Linguistic Representation and Language Use in Aphasia. in *Twenty-First Century Psycholinguistics: Four Cornerstones* 57–68 (Routledge/Taylor & Francis Group, 2005).](https://www.zotero.org/google-docs/?qrzPMM)

[13. Tan, E. J., Neill, E. & Rossell, S. L. Assessing the Relationship between Semantic Processing and Thought Disorder Symptoms in Schizophrenia. *J Int Neuropsychol Soc* **21**, 629–638 (2015).](https://www.zotero.org/google-docs/?qrzPMM)

[14. Kiefer, M. *et al.* Neuro-cognitive mechanisms of conscious and unconscious visual perception: From a plethora of phenomena to general principles. *Advances in Cognitive Psychology* **7**, 55–67 (2011).](https://www.zotero.org/google-docs/?qrzPMM)

[15. Steinhauer, K., Royle, P., Drury, J. E. & Fromont, L. A. The priming of priming: Evidence that the N400 reflects context-dependent post-retrieval word integration in working memory. *Neuroscience Letters* **651**, 192–197 (2017).](https://www.zotero.org/google-docs/?qrzPMM)

[16. Liu, B., Wu, G., Meng, X. & Dang, J. Correlation between prime duration and semantic priming effect: Evidence from N400 effect. *Neuroscience* **238**, 319–326 (2013).](https://www.zotero.org/google-docs/?qrzPMM)

[17. Moshontz, H. *et al.* The Psychological Science Accelerator: Advancing Psychology Through a Distributed Collaborative Network. *Advances in Methods and Practices in Psychological Science* **1**, 501–515 (2018).](https://www.zotero.org/google-docs/?qrzPMM)

[18. Lucas, M. Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review* **7**, 618–630 (2000).](https://www.zotero.org/google-docs/?qrzPMM)

[19. Hutchison, K. A. Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin & Review* **10**, 785–813 (2003).](https://www.zotero.org/google-docs/?qrzPMM)

[20. Buchanan, E. M., Valentine, K. D. & Maxwell, N. P. LAB: Linguistic Annotated Bibliography – a searchable portal for normed database information. *Behav Res* **51**, 1878–1888 (2019).](https://www.zotero.org/google-docs/?qrzPMM)

[21. New, B., Brysbaert, M., Veronis, J. & Pallier, C. The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics* **28**, 661–677 (2007).](https://www.zotero.org/google-docs/?qrzPMM)

[22. Gimenes, M. & New, B. Worldlex: Twitter and blog word frequencies for 66 languages. *Behav Res* **48**, 963–972 (2016).](https://www.zotero.org/google-docs/?qrzPMM)

[23. Lison, P. & Tiedemann, J. Opensubtitles2016: Extracting large parallel corpora from movie and tv subtitles. (2016).](https://www.zotero.org/google-docs/?qrzPMM)

[24. Hutchison, K. A. *et al.* The semantic priming project. *Behav Res* **45**, 1099–1114 (2013).](https://www.zotero.org/google-docs/?qrzPMM)

[25. Balota, D. A. *et al.* The English Lexicon Project. *Behavior Research Methods* **39**, 445–459 (2007).](https://www.zotero.org/google-docs/?qrzPMM)

[26. Aguasvivas, J. A. *et al.* SPALEX: A Spanish Lexical Decision Database From a Massive Online Data Collection. *Front. Psychol.* **9**, 2156 (2018).](https://www.zotero.org/google-docs/?qrzPMM)

[27. Bradley, M. M. & Lang, P. J. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry* **25**, 49–59 (1994).](https://www.zotero.org/google-docs/?qrzPMM)

[28. Warriner, A. B., Kuperman, V. & Brysbaert, M. Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behav Res* **45**, 1191–1207 (2013).](https://www.zotero.org/google-docs/?qrzPMM)

[29. Bradley, M. M. & Lang, P. J. *Affective norms for English words (ANEW): Instruction manual and affective ratings*. (1999).](https://www.zotero.org/google-docs/?qrzPMM)

[30. Brysbaert, M., Warriner, A. B. & Kuperman, V. Concreteness ratings for 40 thousand generally known English word lemmas. *Behav Res* **46**, 904–911 (2014).](https://www.zotero.org/google-docs/?qrzPMM)

[31. Stadthagen-Gonzalez, H. & Davis, C. J. The Bristol norms for age of acquisition, imageability, and familiarity. *Behavior Research Methods* **38**, 598–605 (2006).](https://www.zotero.org/google-docs/?qrzPMM)

[32. De Deyne, S., Navarro, D. J., Perfors, A. & Storms, G. Structure at every scale: A semantic network account of the similarities between unrelated concepts. *Journal of Experimental Psychology: General* **145**, 1228–1254 (2016).](https://www.zotero.org/google-docs/?qrzPMM)

[33. De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M. & Storms, G. The “Small World of Words” English word association norms for over 12,000 cue words. *Behav Res* **51**, 987–1006 (2019).](https://www.zotero.org/google-docs/?qrzPMM)

[34. Vankrunkelsven, H., Verheyen, S., Storms, G. & De Deyne, S. Predicting Lexical Norms: A Comparison between a Word Association Model and Text-Based Word Co-occurrence Models. *Journal of Cognition* **1**, 45 (2018).](https://www.zotero.org/google-docs/?qrzPMM)

[35. Vitevitch, M. S., Goldstein, R., Siew, C. S. Q. & Castro, N. Using complex networks to understand the mental lexicon. *Yearbook of the Poznan Linguistic Meeting* **1**, 119–138 (2014).](https://www.zotero.org/google-docs/?qrzPMM)

[36. Kerr, N. L. HARKing: Hypothesizing After the Results are Known. *Pers Soc Psychol Rev* **2**, 196–217 (1998).](https://www.zotero.org/google-docs/?qrzPMM)

[37. Cree, G. S., McRae, K. & McNorgan, C. An Attractor Model of Lexical Conceptual Processing: Simulating Semantic Priming. *Cognitive Science* **23**, 371–414 (1999).](https://www.zotero.org/google-docs/?qrzPMM)

[38. Zannino, G. D., Perri, R., Pasqualetti, P., Caltagirone, C. & Carlesimo, G. A. Analysis of the semantic representations of living and nonliving concepts: A normative study. *Cognitive Neuropsychology* **23**, 515–540 (2006).](https://www.zotero.org/google-docs/?qrzPMM)

[39. Buchanan, E. M., Valentine, K. D. & Maxwell, N. P. English semantic feature production norms: An extended database of 4436 concepts. *Behav Res* **51**, 1849–1863 (2019).](https://www.zotero.org/google-docs/?qrzPMM)

[40. Nelson, D. L., McEvoy, C. L. & Schreiber, T. A. The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers* **36**, 402–407 (2004).](https://www.zotero.org/google-docs/?qrzPMM)

[41. Landauer, T. K., Foltz, P. W. & Laham, D. An introduction to latent semantic analysis. *Discourse Processes* **25**, 259–284 (1998).](https://www.zotero.org/google-docs/?qrzPMM)

[42. Landauer, T. K. & Dumais, S. T. A solution to Plato’s problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review* **104**, 211–240 (1997).](https://www.zotero.org/google-docs/?qrzPMM)

[43. Lund, K. & Burgess, C. Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers* **28**, 203–208 (1996).](https://www.zotero.org/google-docs/?qrzPMM)

[44. Heyman, T., Bruninx, A., Hutchison, K. A. & Storms, G. The (un)reliability of item-level semantic priming effects. *Behav Res* **50**, 2173–2183 (2018).](https://www.zotero.org/google-docs/?qrzPMM)

[45. Hutchison, K. A., Balota, D. A., Cortese, M. J. & Watson, J. M. Predicting Semantic Priming at the Item Level. *Quarterly Journal of Experimental Psychology* **61**, 1036–1066 (2008).](https://www.zotero.org/google-docs/?qrzPMM)

[46. Buchanan, E. M., Maxwell, N. P. & Valentine, K. D. Using the Semantic Priming Project to understand variability in priming. (2018).](https://www.zotero.org/google-docs/?qrzPMM)

[47. Morrison, C. M., Chappell, T. D. & Ellis, A. W. Age of Acquisition Norms for a Large Set of Object Names and Their Relation to Adult Estimates and Other Variables. *The Quarterly Journal of Experimental Psychology Section A* **50**, 528–559 (1997).](https://www.zotero.org/google-docs/?qrzPMM)

[48. Kuperman, V., Stadthagen-Gonzalez, H. & Brysbaert, M. Age-of-acquisition ratings for 30,000 English words. *Behav Res* **44**, 978–990 (2012).](https://www.zotero.org/google-docs/?qrzPMM)

[49. Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S. & Dean, J. Distributed representations of words and phrases and their compositionality. in *Advances in neural information processing systems* 3111–3119 (2013).](https://www.zotero.org/google-docs/?qrzPMM)

[50. Mikolov, T., Chen, K., Corrado, G. & Dean, J. Efficient Estimation of Word Representations in Vector Space. *arXiv:1301.3781 [cs]* (2013).](https://www.zotero.org/google-docs/?qrzPMM)

[51. van Paridon, J. & Thompson, B. subs2vec: Word embeddings from subtitles in 55 languages. *Behav Res* **53**, 629–655 (2021).](https://www.zotero.org/google-docs/?qrzPMM)

[52. Tzelgov, J. & Eben-ezra, S. Components of the between-language semantic priming effect. *European Journal of Cognitive Psychology* **4**, 253–272 (1992).](https://www.zotero.org/google-docs/?qrzPMM)

[53. Kirsner, K., Smith, M. C., Lockhart, R. S., King, M. L. & Jain, M. The bilingual lexicon: Language-specific units in an integrated network. *Journal of Verbal Learning and Verbal Behavior* **23**, 519–539 (1984).](https://www.zotero.org/google-docs/?qrzPMM)

[54. Kelley, K. Sample size planning for the coefficient of variation from the accuracy in parameter estimation approach. *Behavior Research Methods* **39**, 755–766 (2007).](https://www.zotero.org/google-docs/?qrzPMM)

[55. Kelley, K., Darku, F. B. & Chattopadhyay, B. Accuracy in parameter estimation for a general class of effect sizes: A sequential approach. *Psychological Methods* **23**, 226–243 (2018).](https://www.zotero.org/google-docs/?qrzPMM)

[56. Heyman, T., Hutchison, K. A. & Storms, G. Is semantic priming (ir)rational? Insights from the speeded word fragment completion task. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **42**, 1657–1663 (2016).](https://www.zotero.org/google-docs/?qrzPMM)

[57. Maxwell, S. E., Kelley, K. & Rausch, J. R. Sample Size Planning for Statistical Power and Accuracy in Parameter Estimation. *Annu. Rev. Psychol.* **59**, 537–563 (2008).](https://www.zotero.org/google-docs/?qrzPMM)

[58. Faust, M. E., Balota, D. A., Spieler, D. H. & Ferraro, F. R. Individual differences in information-processing rate and amount: Implications for group differences in response latency. *Psychological Bulletin* **125**, 777–799 (1999).](https://www.zotero.org/google-docs/?qrzPMM)

[59. Keuleers, E. & Brysbaert, M. Wuggy: A multilingual pseudoword generator. *Behavior Research Methods* **42**, 627–633 (2010).](https://www.zotero.org/google-docs/?qrzPMM)

[60. De Deyne, S., Navarro, D. J. & Storms, G. Better explanations of lexical and semantic cognition using networks derived from continued rather than single-word associations. *Behav Res* **45**, 480–498 (2013).](https://www.zotero.org/google-docs/?qrzPMM)

[61. Siew, C. S. Q. & Vitevitch, M. S. Spoken word recognition and serial recall of words from components in the phonological network. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **42**, 394–410 (2016).](https://www.zotero.org/google-docs/?qrzPMM)

[62. Brysbaert, M. & New, B. 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 (2009).](https://www.zotero.org/google-docs/?qrzPMM)

[63. van Heuven, W. J. B., Mandera, P., Keuleers, E. & Brysbaert, M. Subtlex-UK: A New and Improved Word Frequency Database for British English. *Quarterly Journal of Experimental Psychology* **67**, 1176–1190 (2014).](https://www.zotero.org/google-docs/?qrzPMM)

[64. Keuleers, E., Brysbaert, M. & New, B. SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods* **42**, 643–650 (2010).](https://www.zotero.org/google-docs/?qrzPMM)

[65. Cai, Q. & Brysbaert, M. SUBTLEX-CH: Chinese Word and Character Frequencies Based on Film Subtitles. *PLoS ONE* **5**, e10729 (2010).](https://www.zotero.org/google-docs/?qrzPMM)

[66. Cuetos, F., Glez-Nosti, M., Barbón, A. & Brysbaert, M. SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicológica* **32**, 133–143 (2011).](https://www.zotero.org/google-docs/?qrzPMM)

[67. Brysbaert, M. *et al.* The Word Frequency Effect: A Review of Recent Developments and Implications for the Choice of Frequency Estimates in German. *Experimental Psychology* **58**, 412–424 (2011).](https://www.zotero.org/google-docs/?qrzPMM)

[68. Dimitropoulou, M., Duñabeitia, J. A., Avilés, A., Corral, J. & Carreiras, M. Subtitle-Based Word Frequencies as the Best Estimate of Reading Behavior: The Case of Greek. *Front. Psychology* **1**, (2010).](https://www.zotero.org/google-docs/?qrzPMM)

[69. Mandera, P., Keuleers, E., Wodniecka, Z. & Brysbaert, M. Subtlex-pl: subtitle-based word frequency estimates for Polish. *Behav Res* **47**, 471–483 (2015).](https://www.zotero.org/google-docs/?qrzPMM)

[70. Wijffels, J. *et al. udpipe: Tokenization, Parts of Speech Tagging, Lemmatization and Dependency Parsing with the ‘UDPipe’ ‘NLP’ Toolkit*. (2021).](https://www.zotero.org/google-docs/?qrzPMM)

[71. Bojanowski, P., Grave, E., Joulin, A. & Mikolov, T. Enriching Word Vectors with Subword Information. *arXiv preprint arXiv:1607.04606* (2016).](https://www.zotero.org/google-docs/?qrzPMM)

[72. Benoit, K., Muhr, D. & Watanabe, K. *stopwords: Multilingual Stopword Lists*. (2021).](https://www.zotero.org/google-docs/?qrzPMM)

[73. Henninger, F., Shevchenko, Y., Mertens, U. K., Kieslich, P. J. & Hilbig, B. E. *lab.js: A free, open, online study builder*. https://osf.io/fqr49 (2019) doi:10.31234/osf.io/fqr49.](https://www.zotero.org/google-docs/?qrzPMM)

[74. Henninger, F., Shevchenko, Y., Mertens, U. K., Kieslich, P. & Hilbig, B. E. Who said browser-based experiments can’t have proper timing? Implementing accurate presentation and response timing in browser. (2018).](https://www.zotero.org/google-docs/?qrzPMM)

[75. Hilbig, B. E. Reaction time effects in lab- versus Web-based research: Experimental evidence. *Behav Res* **48**, 1718–1724 (2016).](https://www.zotero.org/google-docs/?qrzPMM)

[76. Neely, J. H., Keefe, D. E. & Ross, K. L. Semantic priming in the lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **15**, 1003–1019 (1989).](https://www.zotero.org/google-docs/?qrzPMM)

[77. Johnston, R. A. & Barry, C. Age of acquisition and lexical processing. *Visual Cognition* **13**, 789–845 (2006).](https://www.zotero.org/google-docs/?qrzPMM)

[78. Ghyselinck, M., Lewis, M. B. & Brysbaert, M. Age of acquisition and the cumulative-frequency hypothesis: A review of the literature and a new multi-task investigation. *Acta Psychologica* **115**, 43–67 (2004).](https://www.zotero.org/google-docs/?qrzPMM)

[79. Juhasz, B. J. Age-of-Acquisition Effects in Word and Picture Identification. *Psychological Bulletin* **131**, 684–712 (2005).](https://www.zotero.org/google-docs/?qrzPMM)

[80. Brysbaert, M. & Ellis, A. W. Aphasia and age of acquisition: are early-learned words more resilient? *Aphasiology* **30**, 1240–1263 (2016).](https://www.zotero.org/google-docs/?qrzPMM)

[81. Richardson, J. T. E. Imageability and concreteness. *Bull. Psychon. Soc.* **7**, 429–431 (1976).](https://www.zotero.org/google-docs/?qrzPMM)

[82. Richardson, J. T. E. Concreteness and Imageability. *Quarterly Journal of Experimental Psychology* **27**, 235–249 (1975).](https://www.zotero.org/google-docs/?qrzPMM)

[83. Paivio, A., Walsh, M. & Bons, T. Concreteness effects on memory: When and why? *Journal of Experimental Psychology: Learning, Memory, and Cognition* **20**, 1196–1204 (1994).](https://www.zotero.org/google-docs/?qrzPMM)

[84. Wilson, M. MRC psycholinguistic database: Machine-usable dictionary, version 2.00. *Behavior Research Methods, Instruments, & Computers* **20**, 6–10 (1988).](https://www.zotero.org/google-docs/?qrzPMM)

[85. Montefinese, M., Ambrosini, E., Fairfield, B. & Mammarella, N. Semantic memory: A feature-based analysis and new norms for Italian. *Behav Res* **45**, 440–461 (2013).](https://www.zotero.org/google-docs/?qrzPMM)

[86. Kremer, G. & Baroni, M. A set of semantic norms for German and Italian. *Behav Res* **43**, 97–109 (2011).](https://www.zotero.org/google-docs/?qrzPMM)

[87. Ruts, W. *et al.* Dutch norm data for 13 semantic categories and 338 exemplars. *Behavior Research Methods, Instruments, & Computers* **36**, 506–515 (2004).](https://www.zotero.org/google-docs/?qrzPMM)

**Acknowledgements**

The authors received no specific funding for this work. Krystian Barzykowski and Ewa Ilczuk were supported by a grant from the National Science Centre, Poland (2019/35/B/HS6/00528).

**Author contributions**

* Erin M. Buchanan: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Software, Supervision, Validation, Visualization, Writing - Original Draft, Writing - Review & Editing
* Kelly M. Cuccolo: Investigation, Project Administration, Supervision, Writing - Review & Editing
* Nicholas Coles: Project Administration, Writing - Review & Editing
* Aishwarya Iyer: Project Administration, Writing - Review & Editing
* Neil Lewis Jr.: Project Administration, Writing - Review & Editing
* Kim Peters: Project Administration, Writing - Review & Editing
* Niels van Berkel: Project Administration, Software, Writing - Review & Editing
* Anna E. van ’t Veer: Project Administration, Writing - Review & Editing
* Jack E. Taylor: Conceptualization, Methodology, Writing - Original Draft, Writing - Review & Editing
* Maria Montefinese: Conceptualization, Methodology, Resources, Writing - Original Draft, Writing - Review & Editing
* K. D. Valentine: Conceptualization, Writing - Original Draft, Writing - Review & Editing
* Nicholas P. Maxwell: Conceptualization, Writing - Review & Editing
* Martin R. Vasilev: Investigation, Resources, Writing - Review & Editing
* Andrei Dumbravă: Investigation, Resources, Writing - Review & Editing
* Nienke Böhm: Investigation, Resources, Writing - Review & Editing
* Natalia Irrazabal: Investigation, Resources, Writing - Review & Editing
* María Fernández-López: Investigation, Writing - Review & Editing
* Susana Ruiz-Fernandez: Investigation, Resources, Writing - Review & Editing
* Mehmet Peker: Investigation, Resources, Writing - Review & Editing
* Jackson G. Lu: Investigation, Writing - Review & Editing
* Ulrich S. Tran: Investigation, Writing - Review & Editing
* Dina El-Dakhs: Investigation, Writing - Review & Editing
* David Moreau: Investigation, Resources, Writing - Review & Editing
* Suzanne L. K. Stewart: Investigation, Writing - Review & Editing
* Luisa Batalha: Investigation, Writing - Review & Editing
* Raluca D. Szekely-Copîndean: Investigation, Resources, Writing - Review & Editing
* Roxana E. Morariu: Investigation, Writing - Review & Editing
* Hendrik Godbersen: Investigation, Resources, Writing - Review & Editing
* Anna O. Kuzminska: Investigation, Resources, Writing - Review & Editing
* Patricia Garrido-Vásquez: Investigation, Writing - Review & Editing
* Miroslav Sirota: Investigation, Writing - Review & Editing
* Chin Wen Cong: Investigation, Resources, Writing - Review & Editing
* Fredrik Sigfrids: Investigation, Resources, Writing - Review & Editing
* Glenn P. Williams: Investigation, Writing - Review & Editing
* Max Korbmacher: Investigation, Writing - Review & Editing
* Izuchukwu L. G. Ndukaihe: Investigation, Writing - Review & Editing
* Krystian Barzykowski: Investigation, Resources, Writing - Review & Editing
* Claudia Mazzuca: Investigation, Writing - Review & Editing
* William E. Davis: Investigation, Writing - Review & Editing
* Gerit Pfuhl: Investigation, Resources, Writing - Review & Editing
* Zoran Pavlović: Investigation, Writing - Review & Editing
* Philip A. Grim II: Investigation, Writing - Review & Editing
* Leticia Micheli: Investigation, Resources, Writing - Review & Editing
* Debora I. Burin: Investigation, Resources, Writing - Review & Editing
* David C. Vaidis: Investigation, Writing - Review & Editing
* Gerald H. Vineyard: Investigation, Writing - Review & Editing
* Juan C. Oliveros-Chacana: Investigation, Resources, Writing - Review & Editing
* Paulo Roberto dos Santos Ferreira: Investigation, Writing - Review & Editing
* Carlota Batres: Investigation, Writing - Review & Editing
* Ekaterina Pronizius: Investigation, Writing - Review & Editing
* Steven Verheyen: Investigation, Resources, Writing - Review & Editing
* Vanessa Era: Investigation, Writing - Review & Editing
* Chris Isloi: Investigation, Writing - Review & Editing
* Aslan Karaaslan: Investigation, Resources, Writing - Review & Editing
* Mahmoud M. Elsherif: Investigation, Resources, Writing - Review & Editing
* Sezin Öner: Investigation, Resources, Writing - Review & Editing
* John Protzko: Investigation, Writing - Review & Editing
* Magdalena Senderecka: Investigation, Resources, Writing - Review & Editing
* Alexandra I. Kosachenko: Investigation, Resources, Writing - Review & Editing
* Anna Maria Borghi: Investigation, Writing - Review & Editing
* Belgüzar Nilay Türkan: Investigation, Writing - Review & Editing
* Thomas Rhys Evans: NA, Investigation, Writing - Review & Editing
* Sandra J. Geiger: Investigation, Writing - Review & Editing
* Bastian Jaeger: Investigation, Writing - Review & Editing
* Jordan W. Suchow: Investigation, Writing - Review & Editing
* Dermot Lynott: Investigation, Writing - Review & Editing
* Joseph P. McFall: Investigation, Writing - Review & Editing
* Flavio Azevedo: Investigation, Writing - Review & Editing
* Tijana Vesić Pavlović: Investigation, Resources, Writing - Review & Editing
* Emily Higgins: Investigation, Writing - Review & Editing
* Ewa Szumowska: Investigation, Resources, Writing - Review & Editing
* Luis Carlos Pereira Monteiro: Investigation, Writing - Review & Editing
* Nwadiogo Chisom ARINZE: Investigation, Writing - Review & Editing
* Asil Ali Özdoğru: Investigation, Resources, Writing - Review & Editing
* Bruno Schivinski: Investigation, Writing - Review & Editing
* Cameron Brick: Investigation, Writing - Review & Editing
* Claus Lamm: Investigation, Writing - Review & Editing
* Oguz A. Acar: Investigation, Writing - Review & Editing
* Ettore Ambrosini: Investigation, Resources, Writing - Review & Editing
* Leigh Ann Vaughn: Investigation, Writing - Review & Editing
* Patrí­cia Arriaga: Investigation, Resources, Writing - Review & Editing
* Ogeday Çoker: Investigation, Writing - Review & Editing
* Zainab A. Alsuhaibani: Investigation, Resources, Writing - Review & Editing
* Carmel A Levitan: Investigation, Writing - Review & Editing
* Jan Philipp Röer: Investigation, Writing - Review & Editing
* Barnaby J.W. Dixson: Investigation, Writing - Review & Editing
* Maital Neta: Investigation, Resources, Writing - Review & Editing
* Chiara Fini Fini: Investigation, Writing - Review & Editing
* Théo Besson: Investigation, Resources, Writing - Review & Editing
* Martín Martínez: Investigation, Resources, Writing - Review & Editing
* Melanie Labusch: Investigation, Writing - Review & Editing
* Marietta Papadatou-Pastou: Investigation, Resources, Writing - Review & Editing
* Kathleen Schmidt: Investigation, Writing - Review & Editing
* Ewa Ilczuk: Investigation, Writing - Review & Editing
* Anjali Thapar: Investigation, Writing - Review & Editing
* Ernest Baskin: Investigation, Writing - Review & Editing
* Kate E. Mulgrew: Investigation, Writing - Review & Editing
* Muhammad Mussaffa Butt: Investigation, Resources, Writing - Review & Editing
* Jacek Buczny: Investigation, Resources, Writing - Review & Editing
* Yuri G. Pavlov: Investigation, Resources, Writing - Review & Editing
* Manuel Perea: Investigation, Writing - Review & Editing
* Ali H. Al-Hoorie: Investigation, Resources, Writing - Review & Editing
* Dauren Kasanov: Investigation, Resources, Writing - Review & Editing
* Jacob J. Keech: Investigation, Writing - Review & Editing
* Bradley J. Baker: Investigation, Writing - Review & Editing
* Peter R. Mallik: Investigation, Writing - Review & Editing
* Savannah C. Lewis: Project Administration, Resources, Writing - Review & Editing

**Competing interests**

The authors declare no competing interests.

**Figure 1.** Stimuli selection method flow chart. Circles represent the data or models used in the decision tree. Diamonds represent a decision criterion for the data selected. Squares represent coding processes or data reduction for the final stimuli set.

Diagram

Description automatically generated

**Figure 2.** Flow chart of the procedure for the study. Within the lexical decision task, participants are given short breaks after 100 trials (i.e., each answer given). The answer choices for that language will always be displayed on the bottom of the screen during the lexical decision task.

Diagram, schematic

Description automatically generated

**Table 1.**

**Design Table**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Question** | **Hypothesis** | **Sampling plan (e.g., power analysis)** | **Analysis Plan** | **Interpretation given to different outcomes** |
| Is semantic priming a non-zero effect? | HA: Response latencies will be faster for related word-pairs in comparison to unrelated word pairs.  H0: Response latencies for related word-pairs will be slower or equal to those for unrelated word-pairs. | We will sample participants on items until they reach a desired accuracy in parameter estimation confidence interval width (SE = .09). | We will calculate the mean and 95% confidence interval for the priming effect subtracting related word conditions from unrelated word conditions at the item level.  This analysis will be calculated on the entire dataset **and** each language separately. Then, these calculations will be repeated for the data with 2.5 *z*-score outlier trials excluded and 3.0 *z*-score outlier trials excluded. | The results will support HA when the lower limit of the confidence interval is **positive and non-zero > 0.0001**  The results will be inconclusive when the lower limit of the confidence interval is **negative or zero ≤ 0.0001.** |
| How does the semantic priming effect vary across languages? | HA: Priming response latencies will be different between languages.  H0: Priming response latencies will not be different between languages. | We will sample participants on items until they reach a desired accuracy in parameter estimation confidence interval width (SE = .09). | We will use the individual language confidence intervals calculated above on the full data, 2.5 *z*-score outliers excluded, and 3.0 *z*-score outliers excluded. | The results will support HA when the confidence intervals **do not overlap between languages calculated to four decimal places.**  The results will be inconclusive when the confidence intervals **do overlap between languages calculated to four decimal places.** |

**Supplementary Information**

All processing code and other files can be found at <https://osf.io/wrpj4/>.