Investigating Object Orientation Effects Across 14 Languages

Sau-Chin Chen[*1], Anna Szabelska[*2], Christopher R. Chartier[*3], Zoltan Kekecs[*4], Dermot Lynott[*5], Pablo Bernabeu[*5], Benedict Jones[*6], Lisa DeBruine[6], Carmel Levitan[*7], Kaitlyn M. Werner[*8], Kelly Wang[*8], Marina Milyavskaya[*8], Erica Musser[*9], Marietta Papadatou-Pastou[*10], Nicholas Coles[*11], Steve M. J. Janssen[*12], Asil Özdoğru[*13], Daniel Storage[*14], Harry Manley[*15], Benjamin Brown[*16], Krystian Barzykowski[*17], Thomas Rhys Evans[*18], Sara Álvarez Solas[*19], Manyu Li[*20], Leigh Ann Vaughn[*21], Balazs Aczel[*22], Attila Szuts[22], Carlota Batres[*23], William Chopik[*24], Kim Peters[*25], Jerome Olsen[*26], Martin Voracek[*26], Christian K. Tamnes[*27], Miroslav Sirota[*28], Dawn Liu[*28], Glenn P. Williams[*29], Arti Parganiha[*30], Babita Pande[*30], Sraddha Pradhan[*30], Noorshama Parveen[*30], Pratibha Kujur[*30], Priyanka Chandel[*30], Margaret Messiah Singh[*30], Chrystalle B. Y. Tan[31], John Protzko[32], Jack Arnal[34], Stefan Stieger[36], Marco Tullio Liuzza[37], Pavol Kačmár[38], Jozef Bavoľár[38], Monika Hricová[38], Gabriel Baník[39], Matúš Adamkovič[39], Ivan Ropovik[39], Peter Babinčák[39], Martin Seehuus[40], Vanja Ković[41], Kathleen Schmidt[42], & Elisabeth Oberzaucher[43]

[1] Department of Human Development and Psychology, Tzu-Chi University, Taiwan

[2] Institute of Cognition and Culture, Queen's University Belfast, UK

[3] Department of Psychology, Ashland University, USA

[4] Department of Psychology, Lund University, Sweden

[5] Department of Psychology, Lancaster University, United Kingdom

[6] Institute of Neuroscience & Psychology, University of Glasgow, Scotland, UK

[7] Department of Cognitive Science, Occidental College, USA

[8] Department of Psychology, Carleton University, Canada

[9] Department of Psychology, Florida International University, Miami, FL, USA

[10] National and Kapodistrian, University of Athens

[11] Department of Psychology, University of Tennessee, Knoxvile, TN, USA

[12] School of Psychology, University of Nottingham - Malaysia Campus, Malaysia

[13] Department of Psychology, Üsküdar University, Turkey

[14] Department of Psychology, University of Denver, USA

[15] Faculty of Psychology, Chulalongkorn University, Thailand

[16] School of Liberal Arts, Georgia Gwinnett College, USA

[17] Applied Memory Research Laboratory, Institute of Psychology, Jagiellonian University, Krakow, Poland

[18] School of Social, Psychological and Behavioural Sciences, Coventry University, United Kingdom

[19] Ecosystem Engineer, Universidad Regional Amazónica Ikiam, Tena, Ecuador

[20] Department of Psychology, University of Louisiana at Lafayette, Lafayette, LA, USA

[21] Department of Psychology, Ithaca College, Ithaca, NY, USA

[22] Institute of Psychology, ELTE, Eotvos Lorand University, Hungary

[23] Department of Psychology, Franklin and Marshall College, Lancaster, PA, USA

[24] Department of Psychology, Michigan State University, East Lansing, MI, USA

[25] School of Psychology, University of Queensland, Australia

[26] Faculty of Psychology, University of Vienna, Austria

[27] Department of Psychology, University of Oslo, Norway

[28] Department of Psychology, University of Essex, United Kingdom

[29] Division of Psychology, School of Social and Health Sciences, Abertay University, Dundee, United Kingdom

[30] School of Studies in Life Science, Pt. Ravishankar Shukla University, Raipur, India
[31] Department of Community and Family Medicine, Universiti Malaysia Sabah, Malaysia
[32] Department of Psychological and Brain Sciences, University of California, Santa Barbara, CA, USA
[34] Department of Psychology, McDaniel College, MD, USA
[36] Department of Psychology and Psychodynamics, Karl Landsteiner University of Health Sciences, Austria
[37] Catanzaro School of Medicine and Surgery, Magna Graecia University of Catanzaro, Italy
[38] Department of Psychology, Faculty of Arts, Pavol Jozef Šafarik University in Košice, Slovakia
[39] Faculty of Education, University of Presov, Slovakia
[40] Department of Psychology, Middlebury College, USA
[41] Department of Psychology, Faculty of Philosophy, University of Belgrade, Serbia
[42] Department of Psychology, Southern Illinois University Carbondale, USA
[43] Department of Evolutionary Anthropology, University of Vienna, Austria

Author Note

Author contributions: Sau-Chin Chen developed study concept and proposed the initial design. Sau-Chin Chen and Anna Szabelska wrote initial proposal draft. Dermot Lynott and Benedict C. Jones gave many suggestions on the improvement of study design and analysis plan. Zoltan Kekecs shared his opinions on the power analysis and sampling plan. Authors who coordinated proposal development are Christopher R. Chartier, Carmel A. Levitan, Kaitlyn M. Werner, Erica D. Musser, Marietta Papadatou-Pastou, Nicholas A. Coles, Pablo Bernabeu, Niklas K. Steffens, and Kelly Wang. Christopher R. Chartier organized and conducted pilot study. Nicholas A. Coles developed the codes for drawing the map of participated laboratories. Authors with asterisk contributed their efforts for the substantial proposal revisions. Sau-Chin Chen, Anna Szabelska and Christopher R. Chartier managed the project's OSF (view-only link: https://tinyurl.com/yb6sg2df).

Funding statement: Glenn P. Williams is supported by the Leverhulme Trust Research Project Grant (RPG-2016-093). Krystian Barzykowski is supported by National Science Centre, Poland (2015/19/D/HS6/00641). Erica D. Musser is supported by the National Institute on Mental Health (1R03MH110812-01). Nicholas Coles is supported by the National Science Foundation Graduate Research Fellowship (R010138018).

Ethical approval statement: Our leading researcher, Sau-Chin Chen, has obtained the approval from the Research Ethics Committee, National Taiwan University (NTU-REC No.: 201805ES025). The latest status of ethical approval for all the participating laboratories is available at our public sheet (https://tinyurl.com/ya2ssl4u).

Correspondence concerning this article should be addressed to Sau-Chin Chen, No. 67, Jei-Ren St., Hualien City, Taiwan. E-mail: csc2009@mail.tcu.edu.tw

Abstract

Mental simulation theories of language comprehension propose that people automatically create mental representations of real objects. Evidence from sentence-picture verification tasks has shown that people mentally represent various visual properties such as shape, color, and size. However, the evidence for mental simulations of object orientation is limited. We report a study that investigates the match advantage of object orientation across speakers of different languages. This multi-laboratory project aims to achieve two objectives. First, we examine the replicability of the match advantage of object orientation across multiple languages and laboratories. Second, we will use a mental rotation task to measure participants' mental imagery after the sentence-picture verification task. The relationship between the participants' performance of the two tasks will provide a cross-linguistic examination of perceptual simulation processes. With the evaluation of individual mental imagery ability and potential linguistic moderators, we expect a robust estimation of match advantage of object orientation.

Keywords: mental simulation, object orientation, mental rotation, language comprehension

Word count: 3931

Investigating Object Orientation Effects Across 14 Languages

Introduction

Researchers have explored that how human mind represents objects through numerous methods and investigated the underlying cognitive processes of mental simulation during language comprehension. The sentence-picture verification task is a well-known method for assessing the mental simulation of object's physical features during sentence reading (Connell, 2007; De Koning, Wassenburg, Bos, & Van der Schoot, 2017; Stanfield & Zwaan, 2001; **ZwaanEmbodiedsentencecomprehension2005?**; **ZwaanRevisitingMentalSimulation2012?**; **ZwaanLanguageComprehendersMentally2002?**). The task requires participants to read a probe sentence (e.g., *He saw the eagle in the sky.*) and then verify if the target object in the picture was mentioned in the probe sentence. The target object (e.g., *eagle*) is presented in one of the two ways: with a matching feature (e.g., *flying eagle*) or with a mismatching feature (e.g., *standing eagle*). The match advantage refers to faster reaction time for verifying the target object. Because the match advantages resulted from the implied perceptual features (e.g., shape, color, and orientation) of the object while reading the sentence. These findings have been believed to serve as evidence for how the mental representations of objects are formed and activated during language processing (e.g., Barsalou, 1999, 2009).

Among the studied features, objects with matching shape (i.e., *flying eagle* versus *standing eagle*) and color (i.e., *chocolate ice cream* versus *vanilla ice cream*) have been shown to produce match advantages across many languages, including English (**ZwaanEmbodiedsentencecomprehension2005?**; **ZwaanRevisitingMentalSimulation2012?**), Dutch (De Koning, Wassenburg, Bos, & Van der Schoot, 2017; **RommersObjectshapeorientation2013?**), German (Koster, Cadierno, & Chiarandini, 2018), Croatian (**SeticNumericalCongruencyEffect2017?**), and Chinese (**LiERPsStudyMental2017?**). In contrast, match advantages of object orientation have been found in English (Stanfield & Zwaan, 2001; **ZwaanRevisitingMentalSimulation2012?**) and German (Koster, Cadierno, & Chiarandini, 2018), but was absent in the other languages, such as Dutch (De Koning, Wassenburg, Bos, & Van der Schoot, 2017; **EngelenPerceptualsimulationdeveloping2011?**; **RommersObjectshapeorientation2013?**). Several explanations have been proposed to explain these inconsistent findings, including the different experimental procedures in the earlier mentioned studies. Unlike the original study (Stanfield & Zwaan, 2001), the former studies did not require participants to recognize the probe sentences they had read. Without this memory task during verification trials, participants could pay less attention to the meaning of the probe sentence (Zwaan, 2014), in which case they were less likely to form a mental representation of the objects (e.g., Zwaan & van Oostendorp, 1993).

Extended Studies on Embodied Cognition

Among studies which focused on the conceptual representations of perception and action, there is a shared assumption that linguistic cues activate sensory and motor information processing outside the non-linguistic context (for a review, see **ZwaanRevisitingMentalSimulation2012?**; for a critique, see Mahon & Caramazza, 2008). A highly cited research topic related to it is the action–sentence compatibility effect that is measured by the sensibility judgment task (**GlenbergGroundinglanguageaction2002?**). This sensibility judgment task requires participants to verify the sensibility of sentences by pressing the response buttons toward or away from their bodies. The sentences implied the direction of body movement (e.g., *You gave the earring to Susan*; *Joe sends the card to you*). Participants made faster judgements of sentences that implied the congruent direction. Since the first study of action–sentence compatibility effect was published, researchers valiantly tested the related hypotheses and investigated similar effects with other methods. For example, Beilock, Lyons, Mattarella-Micke, Nusbaum, and Small (2008) found that specialised motor experience (via sport) enhances the mental processing of sentences that are related to the specific sport practised.

As like the action-sentence compatibility effect, some researchers have suggested that the match advantage of object orientation would associate with the activated motor information (**ConnellFunctionalRoleModalitySpecific2012?**; **PingUnderstandinggesturelistener2014?**). (**PingUnderstandinggesturelistener2014?**), for example, exposed individuals to the implied gesture video or to the blank screen between hearing the sentence and watching the object picture. They found that the facilitation effect appeared only in the condition where participants heard a sentence and watched the progressive action in the video and not in the condition where they heard a sentence only. Although these findings provide initial support for the association between mental representations of body movements and non-verbal communications (Alibali & Nathan, 2012; Novack & Goldin-Meadow, 2015; **NovackGesturerepresentationalaction2017?**), the tasks across these studies substantially differed. Without evaluating the true effect size of the match advantage, the results of subsequent studies would not only add little to the overall understanding of how body movements are mentally represented but also will not explain the inconsistency in the orientation effects across the studies.

Therefore, it is necessary to precisely estimate the true effect size in the original paradigm, before commencing with the theoretical and empirical probing for the orientation effect. Because single laboratory studies (De Koning, Wassenburg, Bos, & Van der Schoot, 2017; **RommersObjectshapeorientation2013?**; **ZwaanRevisitingMentalSimulation2012?**) hardly address the possible causes of inconsistent findings, it is imperative to collect data across multiple languages and laboratories as many as possible.This project accumulates materials translated into many languages by the labs recruited from a distributed laboratory network (Moshontz et al., 2018), using the materials and procedure created by Stanfield and Zwaan (2001).

Progressing Investigations on Match Advantage

The match advantage of object orientation has been investigated to date using objects with particular characteristics. As in the example illustrated in Figure 1, a picture of a "horizontal nail" matched the sentence "Tom hammered the nail into the wall" and a picture of a "vertical nail" matched the sentence "Tom hammered the nail into the floor." With this manipulation, the response data from two object orientations can be labeled the "matching pairs" and "mismatching pairs," respectively. Specifically, all the materials of Stanfield and Zwaan (2001) were manipulable objects, such as nails, and pens, and scissors. This choice restricted the object orientations to horizontal and vertical.

Stanfield and Zwaan (2001) ensured that the participants sufficiently understood the probe sentences and showed significant match advantages of object orientation. However, it is still unknown whether the match advantage associated with the general cognitive ability. Some further studies of mental simulation have suggested the relationship of mental rotation and spatial cognition (Chu & Kita, 2008; Pouw, de Nooijer, van Gog, Zwaan, & Paas, 2014). (**chenDoesObjectSize2018?**) investigated the relationship between the match advantage and mental rotation across three languages: English, Dutch, and Chinese. They introduced the picture-picture verification task to examine that how individuals process the target pictures with the same mental image, regardless of their native language. This picture-picture verification task was a modified form of the mental rotation paradigm (**CohenMentalRotationMental1993?**). In each trial of this task, two pictures appeared on opposite sides of the screen. Participants verify whether the pictures represent identical or different objects. The verification times for pictures of identical objects presented in the same orientation (that is, two identical pictures presented in horizontal orientation or vertical orientation) were shorter than those presented in different orientations (one horizontal; one vertical).

To evaluate the relationship between match advantage and mental rotation, (**chenDoesObjectSize2018?**) employed larger (e.g., *anchor*, *gate arms*) and smaller (e.g., *pen*, *nail*) objects. According to the findings about the relationship of mental rotation and spatial cognition (Chu & Kita, 2008; Pouw, de Nooijer, van Gog, Zwaan, & Paas, 2014), the authors assumed that response times would be longer for larger objects in both the sentence-picture verification task and picture-picture verification task. The results of the two tasks showed that large objects presented in different orientations take longer time to verify than small objects (**chenDoesObjectSize2018?**). However, the meta-analysis of (**chenDoesObjectSize2018?**) indicated variation in match advantages across different languages but a consistent effect of mental rotation across different languages. Their study left two primary questions to be answered through this project: (1) "How much of the match advantage of object orientation can be obtained within different languages?" and (2) "What is the correlation between the mental rotation of objects and the match advantage across different languages?"

Theory thinking?

tags: `PSA002`

Methods

Hypotheses and Design

Both the sentence-picture verification task and the picture-picture verification task have two independent variables - a shared between-participant independent variable and a task-specific within-participant independent variable. The shared variable is the primary languages of the participating laboratories. The task-specific variable in sentence-picture verification is whether or not sentence and picture have matching object orientations; the variable in picture-picture verification is whether picture and picture have identical object orientations. The dependent variable for both tasks is the response time the participant verify the target object.

In the sentence-picture verification task, we expect response time to be shorter when object orientations are matching than when mismatching. We expect to see the match advantage within each language but we make no hypotheses about the specific effect sizes in each language. In the picture-picture verification task, we expect shorter response time when the two pictures have identical orientations than different orientations. In addition, we compute an imagery score by taking the elapsed verification time between the orientation settings of critical object pictures (identical, different). If the mental rotation were the general cognitive aspect, we expect imagery scores to be the same across laboratories and languages. If the mental simulation shared the cognitive processing of mental rotation, the imagery scores could be the predictor as critical as the languages.

General Procedure

Participating laboratories will conduct both the sentence-picture verification task and the picture-picture verification task, taking 10 to 15 minutes in total (see Figure 2 for an outline of the general procedure). The sentence-picture verification task will start before or after the survey of Phills and Kekecs (in preparation). In the beginning of the sentence-picture verification task, participants will be instructed with an example. Each trial starts with a left-justified and horizontally centered fixation displayed point for 1000ms, immediately followed by the probe sentence. The sentence is presented until the participant presses the space key, thus acknowledging that they understood the sentence. Then the object picture is presented in the center of the screen until the participant responds. If no response is made, the object picture disappears from the screen after 2 seconds. Participants are instructed to verify the object picture mentioned in the probe sentence as quickly and accurately as they can. Following the original study (Stanfield &

Zwaan, 2001), a test to recognize the presented probe sentence will be conducted after every three to eight trials of sentence-picture verification. This is to make sure that the participants have read each sentence carefully.

The picture-picture verification task will use the same object pictures. The procedure is like the sentence-picture verification task, with one exception. In each trial, two pictures of objects will appear beside the central fixation point until either the participant indicates that the pictures display the same object or two different objects, or the time dedicated for the trial (2 seconds) has elapsed. Two pictures showing the same critical object will appear in each "yes" trial; two pictures showing two different objects from the filler items will appear in each "no" trial. All the procedures are compiled in OpenSesame scripts (**MathotOpenSesameopensourcegraphical2011?**). The instructions and experimental scripts are available in the project's repository (Chen, Chartier, & Szabelska, 2018).

Before the pandemic outbreaks happened in 2020 spring, 2/3 of participating laboratories had completed the studies. The rest of teams had to stop their data collection plans because of the local lockdown. The core project team decided to migrate the study from in the laboratories to on the Internet. To decrease the differences between laboratories and Internet circumstances, we migrated the original python codes to javascript codes. After months of tests, the revised script was able to collect the data through JATOS server(Lange, Kühn, & Filevich, 2015). The rest of teams conducted the Internet study between February to June 2021. The changes of the procedure had been priorly approved by Journal editor and reviewers in October, 22, 2020.

Participant Characteristics

This study were conducted within the Psychological Science Accelerator (PSA, Moshontz et al., 2018), a network of globally distributed psychological science laboratories. Before the pandemic outbreak, 2340 participants (1104 females; averagely 21.46 years old) from 33 laboratories joined and finished the study. After the study migrated to Internet, 1403 participants (926 females; averagely 23.75 years old) completed the study. All the participated laboratories had the approve from their local/institutional ethics review board or committee before the data collection. Appendix (#1) summarised the participant characteristics by laboratory.

Material

The latest update scripts are available at the project OSF repository (https://osf.io/e428p/). For the study on Internet, a recorded verbal brief played at first. Participants soon had a test for the confirmation they art the native speakers of the targeted language. All the verbal briefs were packaged in each language-specific scripts. Appendix (#1) described the deployment of the scripts and the results of participants' fluency test.

tags: `PSA002`

Results

No matter which data sources, the studies were conducted in site or on Internet, the sequential analysis were managed as the preregistered plan. Among the studies conducted in sites, 1979 participants finished sentence-picture verification task and passed the preregistered exclusion criterion (accuracy percentile > 70%); 2007 participants finished picture-picture verification task. -28 participants' rawdata files were lost because the experimenters did not upload the files or submitted the wrong files. Among the studies conduced on Internet, 1337 participants finished sentence-picture verification task and passed the preregistered exclusion criterion; 1402 participants finished picture-picture verification task. We excluded 1 participant from one laboratory (USA_033) because this participant did not complete the picture-picture verification. All the analysis scripts are available in the source files of this article (OSF link).

Intra-lab analysis during data collection

During the data collection, the sample size justification of each laboratory depended on the Bayesian sequential analysis (**schonbrodtSequentialHypothesisTesting2017?**). Once a laboratory replied their progress, the latest results were updated on the public website. Till the end of data collection, only two laboratories (HUN_001, TWN_001) stopped the data collection because the sequential analysis indicated the latest collected data reached the preregistered criterion ($BF_{10} = 10\ or - 10$). Many laboratories stopped the data collection when they reached their registered sample size. Some laboratories did not reach the minimal criterion because of the following reasons: (1)their works were interrupted by the pandemic outbreak; (2)the participants performed worse through the online study; (3) the laboratories allocated the seats for the foreign participants; (4) A language (Serbian) has a various of scripts. At the end of data collection, there are 18 languages registered in this project. The collected data were from 47 laboratories and 3316 participants. The meta information and pre-processing results are available in Appendix (#2).

Inter-lab analysis after data collection

In addition to the preregistered variables included languages and laboratories, we also evaluated the data sources, in site/on Internet, in each preregistered analysis protocol. In the preregistered plan we decided two approaches to estimate the orientation effects. The first approach summarized the median reaction times by condition as the original studies (Stanfield & Zwaan, 2001; Zwaan & Pecher, 2012). The planned meta-analysis evaluated the global effect size based on the medians. The second approach used the raw response time then computed the fixed effects in the mixed-effect models (Baayen, Davidson, & Bates, 2008).

Identify the outliers in each data set. By each laboratory data set, we summarized how many response time points beyond the third quartile in that data set. 0 laboratories had no outliers. Among the data sets showed outliers, the averaged proportion of outliers was 0.25. Appendix (#3) illustrated the distribution of outliers by laboratory. In sum of the data excluded the outliers, Table 1 and Table 2 respectively summarized the orientation effects collected in sites and on Internet by each language.

(Insert Table 1 about here)

(Insert Table 2 about here)

Meta analysis of orientation effects across laboratories. Because the first preregistered analysis plan did not consider the data collected on Internet, we conducted three parts of meta-analysis to evaluate the global effect size. In the first part the data sources were combined. In the second and third parts we conducted the meta analysis on the data in sites and data on Internet respectively. In each part we computed the effect size by data set then estimated the global effect size. Considering the small sample size may bias the estimated effect size, 9 data sets were excluded from the analysis because the available sample size was less than 25.

Meta-analysis across all the data sets showed a positive but weak effect size. Figure 3 showed the larger variances happened to some laboratories in comparison to the others.

(Insert Figure 3 about here)

When we focused on the data collected in sites, the variation among the laboratories was small although the overall effect size was positive but weak. Figure 4 showed that only one laboratory (HUN_001) found a true positive orientation effect.

(Insert Figure 4 about here)

When we focused on the data collected on Internet, the variations among laboratories were obvious larger than the data collected in sites. The overall effect size tended to be negative. Figure 5 showed that only one laboratory (NZL_005) found a true positive orientation effect.

(Insert Figure 5 about here)

Evaluate orientation effect in mixed-effect models. As with the original plan for the mixed-effect models, all our works started from the model having the matching condition as the only one fixed effect. Because the final data came from sites and Internet respectively, we had to evaluate if one mixed-effect model sufficiently fit all the data. Otherwise we had to build the models for each data set. The results indicated that the data collected from sites and Internet had to be analyzed separately. Based on the last recommended practices (Bates, Kliegl, Vasishth, & Baayen, 2015; Brauer & Curtin, 2018), we decided the fitted model in terms of the maximal random structure and the least convergence problems. The final models indicated the interaction of languages and orientation effects from each data source. This section summarized the two final models from the two data sources as below. All the prior models are reported in Appendix (#4).

With the same sample size justification principle as the meta analysis, we excluded the languages with more than 25 participants in each data source before we conducted the mixed-effect models. Portuguese in the sites data and Norwegian in the Internet data were excluded from the analysis of final models. In each data set, we compared the fitness among the models with and without the random slope of matching condition. Both results indicated the model without the random slope had the best fitness. The model from the sites data indicated the fixed effect of orientation effect was positive but insignificant: coefficient = 5.73 ms, p = 03. However, the model from Internet data indicated the fixed effect of orientation effect was negative but insignificant: -3.92 ms, p = 13.

In terms of the final models, we plotted the predicted response times of matching condition in the function of languages. Figure 6 illustrated the estimated responses times based on the sites data. Greek tended to have the detectable orientation effect although the responses times were the longest among languages. Figure 7 illustrated the estimated responses times based on the sites data. Serbian showed a weak orientation effect but had the longest responses times. Many languages had inconsistent results between the studies in sites and on Internet, such as English, German, and Traditional Chinese.

(Insert Figure 6 about here)

(Insert Figure 7 about here)

Analysis on imagery scores. Imagery scores are the dependent measurement of picture-picture verification task. Dependent measurements were summarized by the elapsed responses times from different to identical orientation settings. Prior to the data collection, we assumed the imagery scores of every language group would be nearly equal. Because the real data came from sites and Internet, we managed the mixed-effect models to verify our prior assumption and to evaluate the confounding of data sources and laboratories. Each final model had the better fitness in consideration of the maximal random structure and the least convergence problem. The models indicated that the data sources and laboratories affected little on the imagery scores. In addition to summarize the fitted model of orientation setting and languages interaction as below, all the models are reported in Appendix (#5).

The best fitted model had the random intercepts of participants, targets, and laboratories, but had no slopes of orientation setting. The fixed effect of orientation setting was significant: coefficient = 33.35 ms, $p < .001$. The estimated response times illustrated in Figure 8 indicated that the imagery scores measured from each language were consistently positive. English result also showed the large sample size decreased the variations.

The last planned analysis evaluated the imagery scores as the predictor of the orientation effect. The above analyses confirmed that data sources did not alter the imagery scores but alter the orientation effect. Therefore we evaluate the regression models respectively. Because the random slopes of items in the analyses of the orientation effect were zero, the integrated data for the regression models were the summarized data by participant.

The regression models should have the orientation effect as the dependent variable, languages and imagery scores as the independent variables. If the imagery could predict the orientation effect, the model with two independent variables has to fit better than the model with language only. For the sites data, the model with language only has a better fitness than the model of two independent variables. $b = 0.03$, 95% CI $[-0.13, 0.20]$, $t(1491) = 0.41$, $p = .680$. For the Internet data, the model of two independent variables has a better fitness than the model with language only. $b = -0.28$, 95% CI $[-1.68, 1.12]$, $t(1132) = -0.40$, $p = .691$.

tags: `PSA002`

Discussion

This project will achieve two goals after completion of the analysis of collected data. At first we will have the latest positive or negative evidence about the orientation effects being a general cognitive aspect among languages. A match advantage would support the embodied cognition studies that used the sentence-picture verification task as the paradigm in any particular language. On the other hand, a match disadvantage would serve as a warning for researchers who are planning to modify the sentence-picture verification for their studies. Our other contribution is to provide a platform to accumulate new data from the laboratories that will strictly follow this protocol to collect and analyze the data. With the continuous accumulation from more languages and laboratories, the mega empirical research framework, such as LeBel, McCarthy, Earp, Elson, and Vanpaemel (2018), will provide the robust information for who plan to measure the orientation effect in the circumstance unlike the original study.

tags: `PSA002`

Acknowledgement

Contributions: Chris Chartier and Jeremy Miller managed and monitor the progress. Erin Buchanan provide guides to improve the progress website management. Marton Kovacs and Balazs Aczel helped update the scripts of sequential analysis.

References

Alibali, M. W., & Nathan, M. J. (2012). Embodiment in Mathematics Teaching and Learning: Evidence From Learners' and Teachers' Gestures. *Journal of the Learning Sciences*, *21*, 247–286. https://doi.org/10.1080/10508406.2011.611446

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

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

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philos Trans R Soc Lond B Biol Sci*, *364*, 1281–1289. https://doi.org/10.1098/rstb.2008.0319

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious Mixed Models. Retrieved from http://arxiv.org/abs/1506.04967

Beilock, S. L., Lyons, I. M., Mattarella-Micke, A., Nusbaum, H. C., & Small, S. L. (2008). Sports experience changes the neural processing of action language. *Proceedings of the National Academy of Sciences*, *105*(36), 13269–13273. https://doi.org/10.1073/pnas.0803424105

Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, *23*(3), 389–411. https://doi.org/10.1037/met0000159

Chen, S.-C., Chartier, C. R., & Szabelska, A. (2018). Object Orientation cross languages: pilot. *OSF*. Retrieved from https://osf.io/e428p/

Chu, M., & Kita, S. (2008). Spontaneous gestures during mental rotation tasks: Insights into the microdevelopment of the motor strategy. *Journal of Experimental Psychology: General*, *137*(4), 706–723. https://doi.org/10.1037/a0013157

Connell, L. (2007). Representing object colour in language comprehension. *Cognition*, *102*, 476–485. https://doi.org/10.1016/j.cognition.2006.02.009

De Koning, B. B., Wassenburg, S. I., Bos, L. T., & Van der Schoot, M. (2017). Mental simulation of four visual object properties: Similarities and differences as assessed by the sentence-picture verification task. *Journal of Cognitive Psychology*, *29*(4), 420–432. https://doi.org/10.1080/20445911.2017.1281283

Koster, D., Cadierno, T., & Chiarandini, M. (2018). Mental simulation of object orientation and size: A conceptual replication with second language learners. *Journal of the European Second Language Association*, *2*(1). https://doi.org/10.22599/jesla.39

Lange, K., Kühn, S., & Filevich, E. (2015). "Just Another Tool for Online Studies" (JATOS): An Easy Solution for Setup and Management of Web Servers Supporting Online Studies. *PLOS ONE*, *10*(6), e0130834. https://doi.org/10.1371/journal.pone.0130834

LeBel, E. P., McCarthy, R. J., Earp, B. D., Elson, M., & Vanpaemel, W. (2018). A Unified Framework to Quantify the Credibility of Scientific Findings. *Advances in Methods and Practices in Psychological Science*, *1*(3), 389–402. https://doi.org/10.1177/2515245918787489

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

Moshontz, H., Campbell, L., Ebersole, C. R., IJzerman, H., Urry, H. L., Forscher, P. S., … Chartier, C. R. (2018). The Psychological Science Accelerator: Advancing Psychology Through a Distributed Collaborative Network: *Advances in Methods and Practices in Psychological Science*, *1*(4), 501–515. https://doi.org/10.1177/2515245918797607

Novack, M. A., & Goldin-Meadow, S. (2015). Learning from Gesture: How Our Hands Change Our Minds. *Educational Psychology Review*, *27*(3), 405–412. https://doi.org/10.1007/s10648-015-9325-3

Phills, C., & Kekecs, Z. (in preparation). *Gendered Social Category Representations.*

Pouw, W. T. J. L., de Nooijer, J. A., van Gog, T., Zwaan, R. A., & Paas, F. (2014). Toward a more embedded/extended perspective on the cognitive function of gestures. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00359

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, *12*(2), 153–156. https://doi.org/10.1111/1467-9280.00326

Zwaan, R. A. (2014). Replications Should Be Performed With Power and Precision: A Response to Rommers, Meyer, and Huettig (2013). *Psychological Science*, *25*(1), 305–307. https://doi.org/10.1177/0956797613509634

Zwaan, R. A., & Pecher, D. (2012). Revisiting Mental Simulation in Language Comprehension: Six Replication Attempts. *PLoS ONE*, *7*, e51382. https://doi.org/10.1371/journal.pone.0051382

Zwaan, R. A., & van Oostendorp, H. (1993). Do readers construct spatial representations in naturalistic story comprehension? *Discourse Processes*, *16*(1-2), 125–143. https://doi.org/10.1080/01638539309544832

tags: `PSA002`

Table 1

Summarized orientation effects by langauge. Data collected in sites.

Language	N	Mismatching	Matching	Orientation Effect
English	473	567(93.73)	561(94.63)	6.00
German	74	553(95.5)	555(95.61)	-2.00
Greek	73	701(90.53)	692(91.1)	9.00
Hebrew	109	556(97.09)	568(95.57)	-12.00
Hindi	59	594(87.57)	627(91.24)	-32.50
Magyar	97	607(95.45)	614(95.62)	-7.50
Norwegian	92	577(96.29)	581(96.47)	-3.75
Polish	37	561(94.82)	585(95.72)	-24.00
Portuguese	5	588(96.67)	580(98.33)	8.00
Simple Chinese	60	632(91.67)	619(92.78)	13.75
Slovak	103	589(96.6)	606(94.82)	-17.00
Spanish	95	646(92.63)	644(92.54)	2.00
Thai	37	637(90.77)	596(87.61)	41.50
Traditional Chinese	78	611(94.23)	616(93.27)	-5.25
Turkish	137	631(95.26)	618(94.89)	12.50

Table 2

Summarized orientation effects by langauge. Data collected on Internet.

Language	N	Mismatching	Matching	Orientation Effect
Arabic	79	522(77.95)	544(76.69)	-21.50
Brazilian Portuguese	50	778(94)	779(94)	-0.75
English	630	644(93.16)	645(93.51)	-1.00
German	127	657(95.6)	668(95.21)	-11.50
Norwegian	16	844(96.88)	636(93.75)	208.25
Portuguese	40	672(96.46)	664(94.79)	9.00
Serbian	129	712(93.48)	722(94.38)	-9.50
Traditional Chinese	34	616(91.18)	664(95.34)	-47.75
Turkish	59	779(92.8)	780(90.68)	-1.00

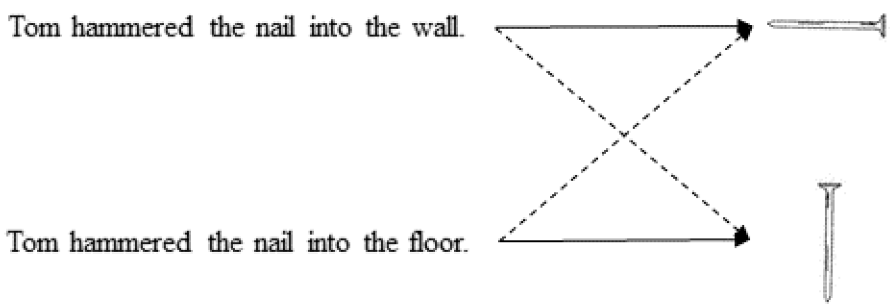

Figure 1. Example of matching (solid lines) and mismatching (dashed lines) sentence-picture pairs.

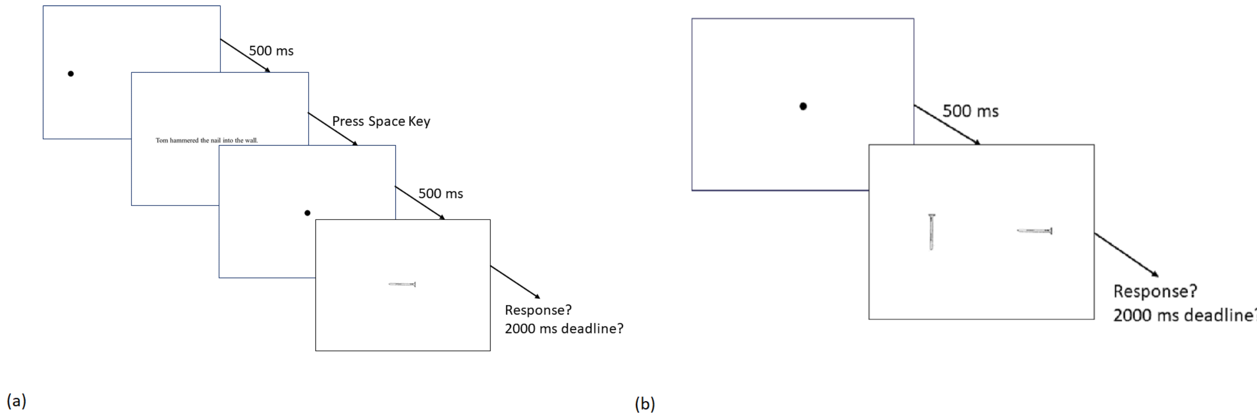

Figure 2. Procedures of stimuli presentation and response. (a) Sentence-picture verification task; (b) Picture-picture verification task.

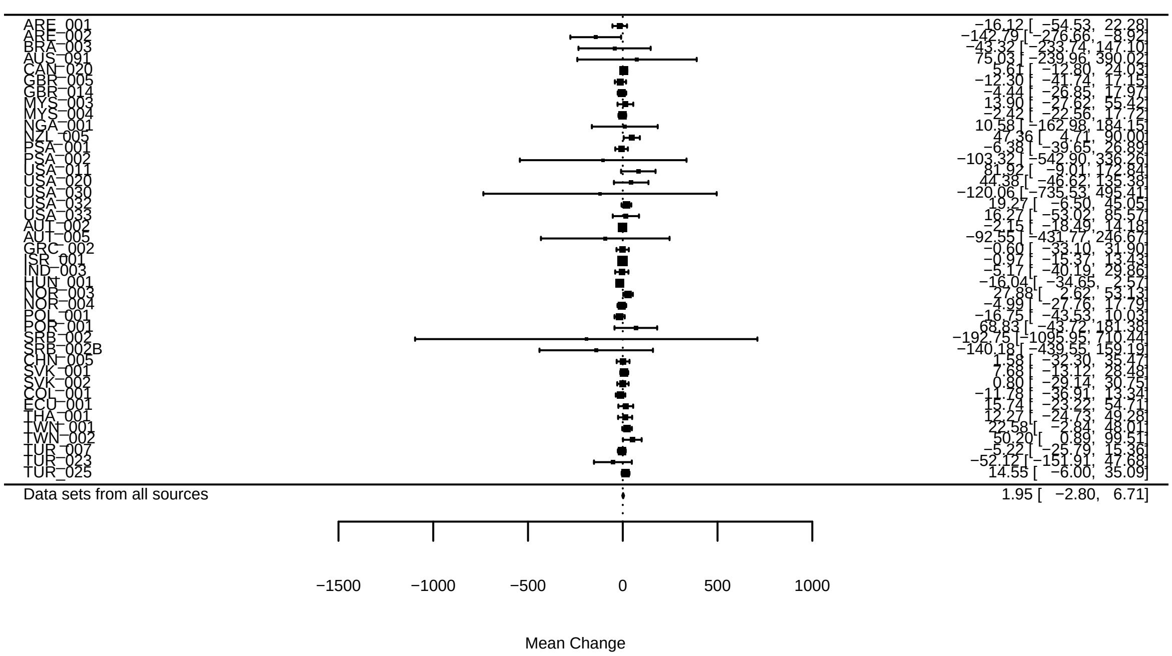

Figure 3. Meta analysis for all data

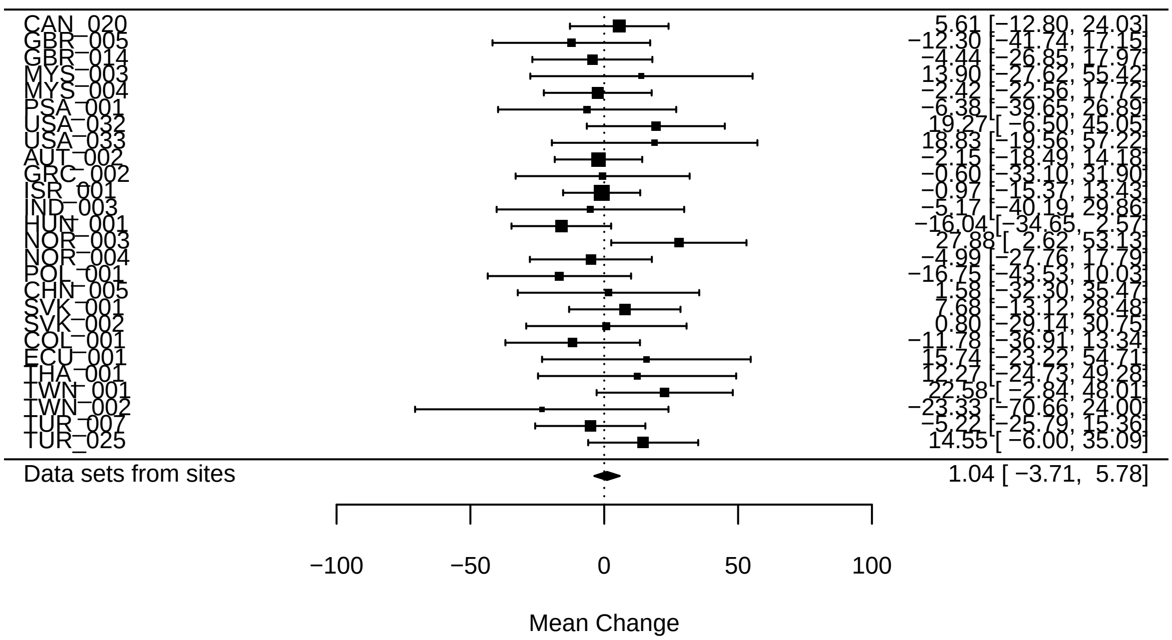

Figure 4. Meta analysis for site data

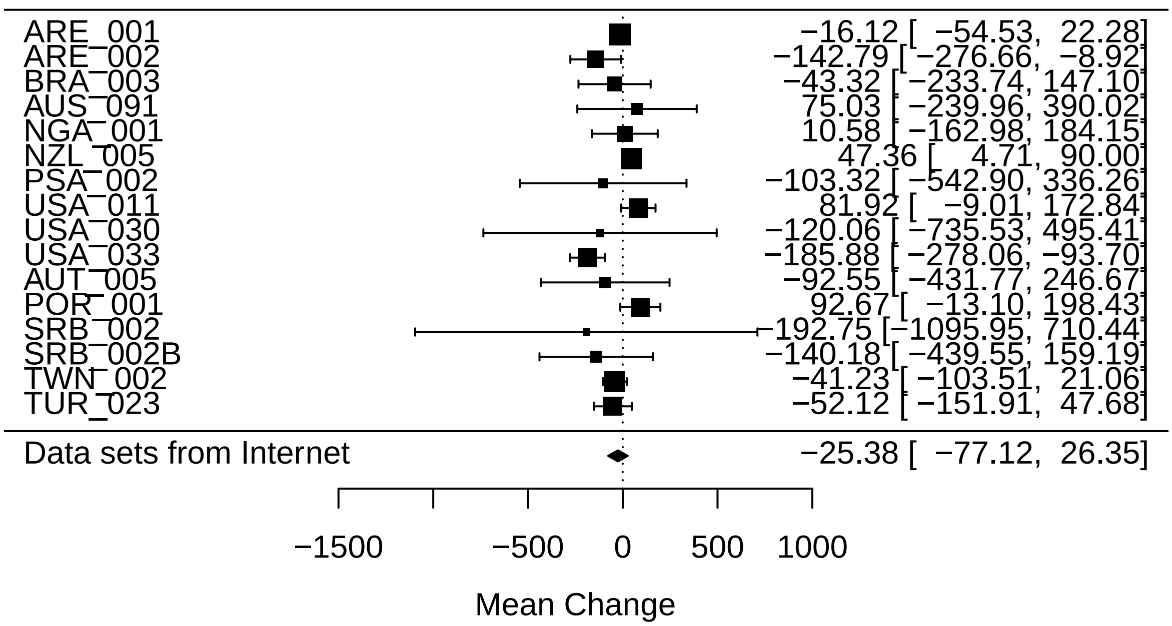

Figure 5. Meta analysis for Internet data

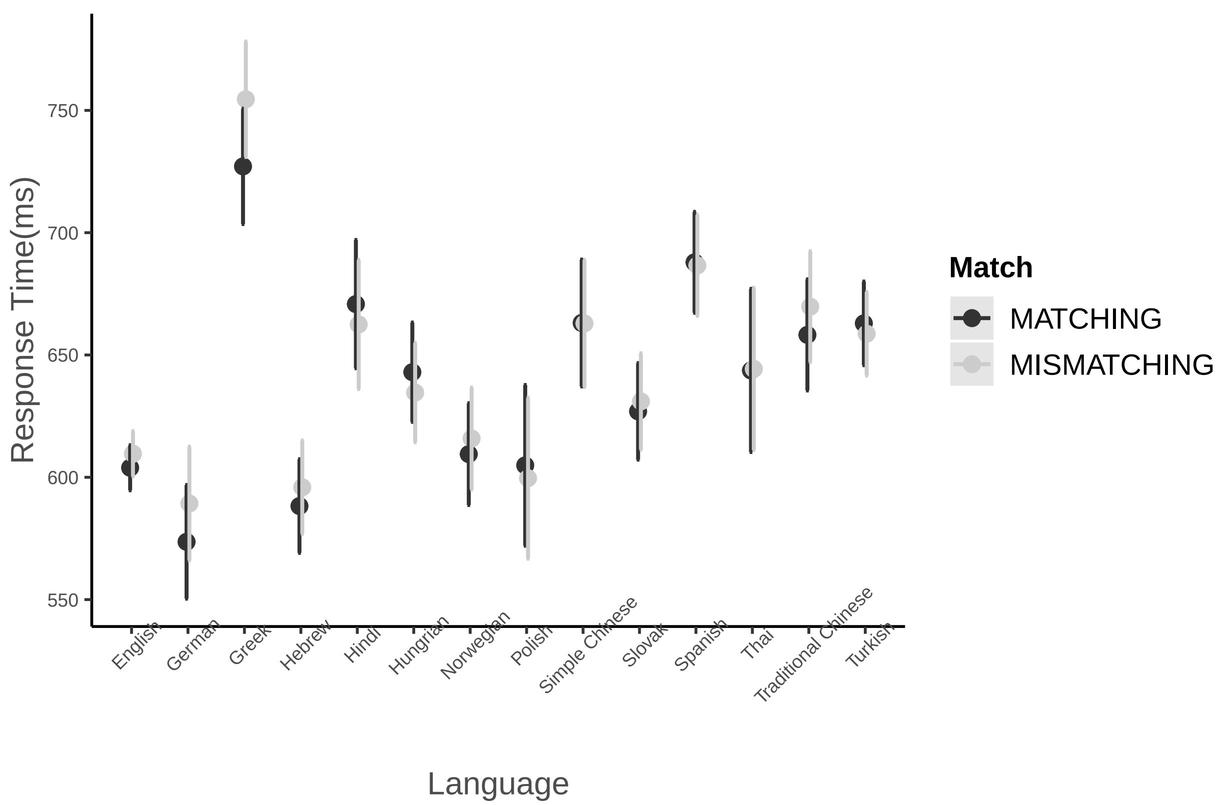

Figure 6. Estimated response times and 95% CI of sentence-picture verification in the function of langauges (by sites data)

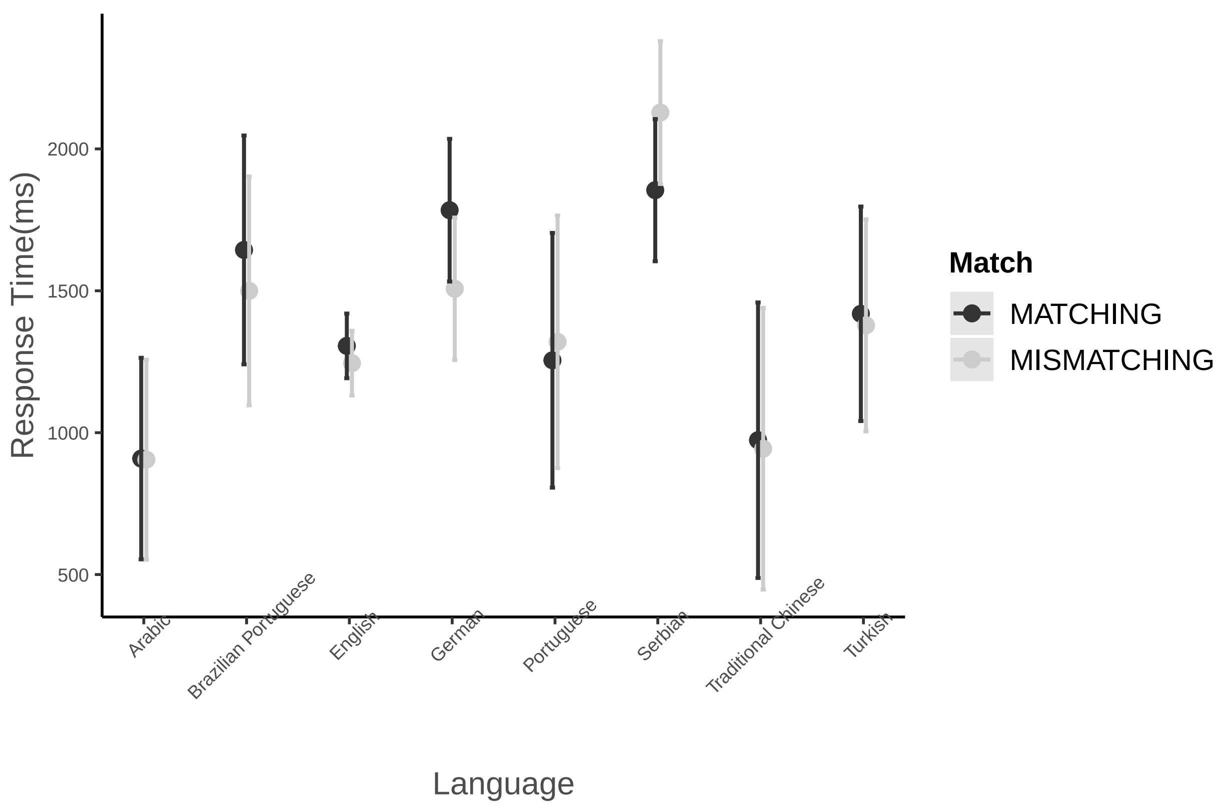

Figure 7. Estimated response times and 95% CI of sentence-picture verification in the function of langauges (by Internet data)

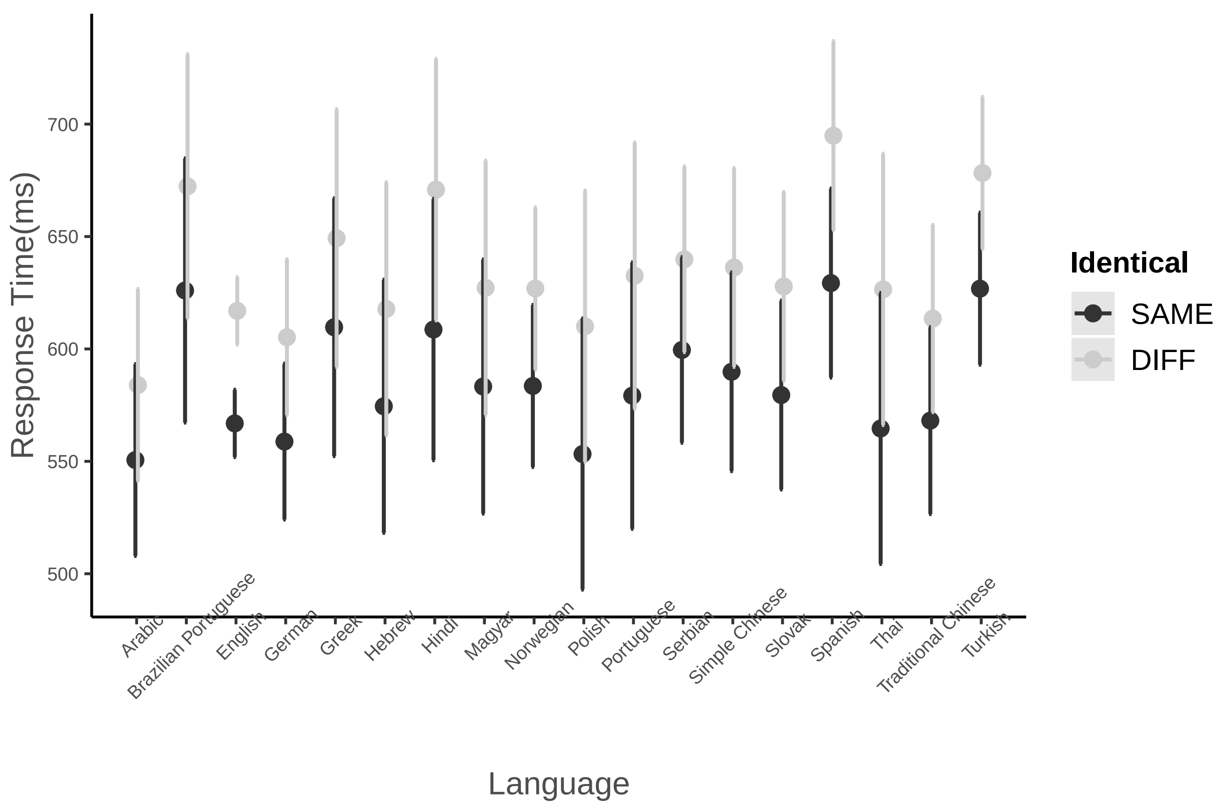

Figure 8. Estimated response times and 95% CI of picture-picture verification in the function of langauges (by all data)