Can you tell them apart? Using machine learning to classify bilinguals' and multilinguals' cognitive and linguistic performance

M. Paula Roncaglia-Denissen[1], Diana Chernichenko[1], Eriko Fukuda[1] and Peter Hendrix[1]

1 Department of Cognitive Science and Artificial Intelligences, Tilburg School of Humanities and Digital Sciences, Tilburg University, the Netherlands

Abstract

The debate of whether bilingualism provides a cognitive and or linguistic advantage is a lasting one. Underlying this debate is the idea that an additional language shapes cognition and linguistic processing. The current research analyzes a behavioral dataset containing individuals' performance in different general cognitive and linguistic tests using a machine learning approach to classify individuals as bilinguals or multilinguals based on their performance. Using an extreme gradient boosting model, we were able to achieve a balanced accuracy of 77%. High scores on a prescriptive grammar test, a verbal fluency test, and a picture naming test were predictive for multilingualism. The implications of the reported results for the field and future research are discussed.

Keywords: bilingualism, multilingualism, domain-general cognitive skills, language skills, machine learning

Introduction

The debate of whether there is a bilingual cognitive advantage has been present in the field literature for quite a long time (Bialystok et al., 2009; Lehtonen et al., 2018; Noort et al., 2006; Peal & Lambert, 1962; Saer, 1923). The lack of consensus about its existence may result, among others, from different methodology used by the studies addressing the matter (Calvo et al., 2016). While some studies provide evidence for a general cognitive advantage among bilinguals (Bialystok et al., 2004; Blom et al., 2014; Friesen et al., 2015), others argue for a lack thereof (Jones et al., 2021; Von Bastian et al., 2016), or for differences in linguistic processing (Bialystok et al., 2003; Bosch & Sebastián-Gallés, 2003; Jared & Szucs, 2002). Underlying this debate is **the idea that an additional language could affect one's general cognitive ability and language processing**, such as working memory (Blom et al., 2014), attention (Friesen et al., 2015), cognitive control (e.g., Blumenfeld & Marian, 2014) and **linguistic skills**, such as phonological awareness (Bialystok et al., 2003) and verbal fluency (Patra et al., 2020).

Despite the abundant literature about the effects of a second language on cognition and language processing, less attention has been dedicated to the impact of additional languages (Marx & Hufeisen, 2003). Only recently multilingualism has gained more interest from the field literature (Cenoz, 2013). Similarly to the plethora of definitions that the term "bilingualism" has received from the field literature, "multilingualism" has been defined, among others, taking one's language proficiency (Cook et al., 2011) and frequency use (Li, 2008) into account. For the current research, a bilingual is anyone who can communicate in or comprehend two languages, and a multilingual in more than two.

In light of the above discussed matters, the current study makes use of machine learning techniques to further investigate the effect of additional learned languages to individuals' general cognitive capacity and linguistic skills. To do so, machine learning models were used to extract the most significant features (individuals' cognitive and linguistic tests performance) in order to classify these individuals according to their self-reported bilingual or multilingual language background information.

The use of machine learning to analyze social-behavioral data has increased in recent years (Lv et al., 2020), and its advantage relies on its ability to capture the non-linear nature of the complex relationship between the use of additional language(s) and cognition might have (Jones et al., 2021). Here, we use a machine learning technique to gain insight into the full predictive power of a set of cognitive and linguistic measures for the classification of individuals as users of two or more languages.

Methods

Participants

The current research analyzed the behavioral dataset for individual differences in domain-general cognitive skills and language skills published by Hintz and associates (2020). The dataset consists of 112 native speakers of Dutch. All participants reported speaking at least one language in addition to Dutch, with 25 participants speaking exactly one additional language (bilinguals) and 87 participants speaking two or more additional languages (multilinguals). For an overview of all the languages spoken by the participants, please see Table 1. Henceforth, we refer to the former group of participants as bilinguals and to the latter group of participants as multilinguals. On average, bilinguals were 22.56 years old (*SD* = 3.42), whereas multilinguals were 22.21 years old (*SD* = 2.62).

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society*.

Materials

Participants completed a battery of cognitive and linguistic tests in Dutch. The test battery included tests of general cognitive skills (e.g., working memory, non-verbal intelligence), linguistic experience. (e.g., vocabulary size, prescriptive grammar knowledge), and linguistic processing skills (e.g., word and sentence production, word and sentence comprehension). For the current study, only the tests in which participants' performance were of relevance for their binary classification into bilinguals or multilinguals will be further explained (please see the Results section). For more information about all the tests conducted and the associated experimental procedures, please see Hintz and associates (2020).

Each participant completed the test battery twice, with an interval of four weeks between both sessions. Here, we used the preprocessed data released by Hintz et al. (2020), which contain performance metrics for 33 cognitive tests in both sessions for each participant. For each test, we calculated average by-participant scaled test scores to obtain more robust estimates of participants' performance across tests. For an overview of all tests participants completed, please see Table 2.

Table 1: Number of participants for every self-reported language.

Language	Number of participants
Dutch (native language)	112
English	111
Croatian	1
Romanian	1
Spanish	13
Papiamento	1
Swahili	1
French	51
German	66
Swedish	1
Arabic	1
Dutch sign language	1
Afrikaans	1
Italian	1
Latin	1
Chinese	1

Table 2: Conducted linguistic experience, general cognitive and linguistic processing tests.

Linguistic experience tests
Stairs4Words
Peabody picture vocabulary test
Spelling test
Author recognition test
Idiom recognition test
Prescriptive grammar test
Syntest
General cognitive tests
Auditory simple reaction time test
Auditory choice reaction time test
Letter comparison test
Visual simple reaction time test
Visual choice reaction time test
Digit span test
Corsi block dicking test
Eriksen Flanker test
Antisaccade test
Raven's advanced progressive matrices test
Linguistic processing skills tests
Picture naming test
Rapid automatized naming
Antonym production
Verbal fluency test
Maximal speech rate
One-minute-test
Klepel test
Monitoring in noise in lists
Rhyme judgment
Auditory lexical decision
Semantic categorization
Phrase and sentence generation
Spontaneous speech
Gender cue activation in sentence comprehension
Verb semantics act. in sentence comprehension
Monitoring in noise in sentences

Analysis

Inferential Statistics

We carried out a series of Welch independent samples *t*-tests to establish which cognitive and linguistic tests revealed significant differences in performance between bilinguals and multilinguals. To control the false discovery rate, the *p*-values for all independent samples *t*-tests were corrected for multiple comparisons using a Benjamini-Hochberg correction (Benjamini & Hochberg, 1995).

Prediction Using Machine Learning

In the current study, we investigated to what extent it is possible to distinguish bilingual and multilingual participants on the basis of their cognitive and linguistic test scores. We technically formulated this investigation as a binary classification task, with participant status (bilingual vs. multilingual) as the response variable and by-participant average test scores for the different tasks as predictors.

We modelled the binary classification task using an extreme gradient boosting model (henceforth xgboost; Chen. et al., 2021)). The xgboost model is an extension of random forests. Like random forests, xgboost fits a sequence of decision trees to the data. Whereas the decision trees in a random forest are independent, however, each tree in an xgboost model is fit to the residual errors of the previous tree. As such, each tree is an expert at the shortcomings of its predecessor.

We fit an xgboost model to the data using the *caret* package for R (Kuhn, 2020), using balanced accuracy as a custom objective function to optimize performance across both bilinguals and multilinguals. The data contain 95 missing values (1.21% of the data). Prior to analysis these missing values were imputed using median imputation. The model was fit under leave-one-out cross validation. Under leave one-out cross validation, predictions for each observation are based on a model trained on all other observations. Random up-sampling of the minority class (bilinguals) was applied to treat class imbalance.

The xgboost model consisted of 1000 trees, with each tree being a stump (*max_depth* = 1). We tuned further hyperparameters for optimal performance. Following the hyperparameter tuning process, we used a learning rate of 0.2 (*eta = 0.2*) and considered a random subset of 70% of the predictors when building each tree (*colsample_by_tree* = 0.7). All other hyperparameters were set to their default.

To gain more insight into the predictive power of the xgboost model we extracted variable importances from the fitted model. We used the standard metric for variable importance in the *xgboost* library for R (Chen et al., 2021), which is the summed information gain achieved by splits on a predictor across all trees. For ease of interpretation we report both raw variable importances and re-scaled variable importances (0 - 100).

Results

Inferential Statistics

After correction of the *p*-values with Benjamini-Hochberg correction, the series of Welch independent *t*-tests fitted to the data revealed a significant difference between mean test scores for bilinguals and multilinguals in three tests: the prescriptive grammar test ($t(43.013) = -3.850$, $p = 0.011$), the verbal fluency test for categories ($t(36.853) = -3.748$, $p = 0.011$), and the picture naming test ($t(35.356) = 3.409$, $p = 0.019$). Group means for bilinguals and multilinguals with 95% confidence intervals for the test scores in the prescriptive grammar test, the verbal fluency test, and the picture naming test are presented in Figure 1.

The prescriptive grammar test required participants to indicate whether auditorily presented sentences were (grammatically) correct or not. As can be seen in the left panel of Figure 1, grammaticality judgments were more accurate for multilingual participants ($M = 0.718$, $SE = 0.131$) as compared to bilingual participants ($M = 0.620$, $SE = 0.022$). In the verbal fluency test for categories participants were asked to name as many animals (part 1) or food and drinks (part 2) as they could within one minute. Test scores are the average number of words named within a minute in both parts of the test. Again, test scores were higher for multilinguals ($M = 26.301$, $SE = 0.494$) as compared to bilinguals ($M = 22.500$, $SE = 0.886$) (see middle panel of Figure 1). The picture naming test consisted of 40 trials in which participants had to name a photograph of an object as fast as possible. As can be seen in the right panel of Figure 1, average response times were significantly shorter for multilingual participants ($M = 2.926$, $SE = 0.0005$) as compared to bilingual participants ($M = 2.967$, $SE = 0.011$). Across the three tests that revealed a significant difference

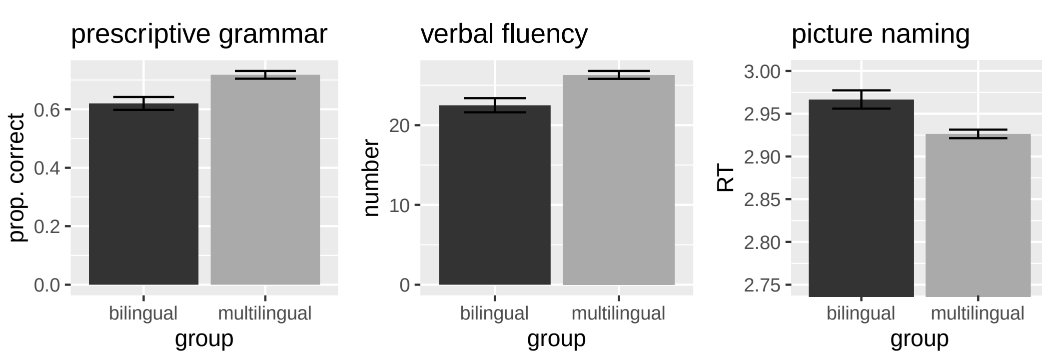

Figure 1. Average performance for bilinguals and multilinguals with 95% confidence intervals and standard errors in the prescriptive grammar (left panel), verbal fluency (middle panel), and picture naming (right panel) tests.

between bilinguals and multilinguals, multilinguals outperformed bilinguals.

Prediction Using Machine Learning

The machine learning model fit to the data achieved a macro average F1 score of 0.778 (majority baseline = 0.440) under leave-one-out cross validation. The confusion matrix for the model is shown in Table 3. Despite the fact that class imbalance was accounted for through up-sampling and a custom objective function, the F1 score model remains higher for multilinguals (0.903) than for bilinguals (0.653). Bilinguals therefore are harder to classify correctly as compared to multilinguals. The ROC AUC score - a measure of how well the model is able to separate both classes - was 0.777. This indicates that if one were to randomly select two participants, one bilingual and one multilingual, the probability of the model assigning a higher probability of multilingualism to the multilingual participant than to the bilingual participant is 77.7%.

Table 3. Confusion matrix for the xgboost model.

		Observed	
		bilingual	Multilingual
Model	bilingual	16	8
	multilingual	9	79

Figure 2 presents variable importances for the xgboost model, scaled from 0 to 100. Consistent with the results of the Welch independent samples *t*-tests, the picture naming (scaled importance: 100.000, raw importance: 0.153), prescriptive grammar (scaled importance: 86.748, raw importance: 0.133), and verbal fluency for categories (scaled importance: 46.740, raw importance: 0.071) tests provided substantial predictive power for participant status.

Interestingly, test scores in the phrase generation task contributed considerably to the predictive power of the xgboost model as well (scaled importance: 74.777, raw importance: 0.114), despite the fact that the performance for the bilinguals and multilinguals was not significantly different in this task ($t(31.904) = -0.431$, $p = 0.837$). In the phrase generation task, participants were asked to generate descriptions of objects varying in structure and complexity. A closer inspection of the data indicated that the standard deviation of the scores in this task was considerably higher for bilinguals ($SD = 0.082$) than for multilinguals ($SD = 0.061$). As such, the xgboost model may have been sensitive to the fact that the probability of a participant being multilingual is lower for extreme scores on the phrase generation test.

Discussion

In the current work we used machine learning techniques to classify individuals' general cognitive and linguistic performances based on their self-reported bilingual or multilingual language background. The use of machine

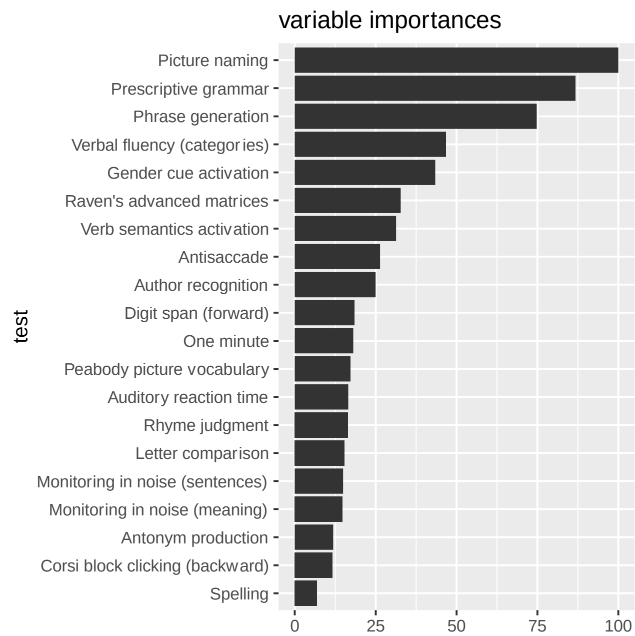

Figure 2. Variable importances for the 20 tests with the highest variable importance in the xgboost model.

learning techniques has recently become more popular in analyzing social behavioral data (Lv et al., 2020) and it can be helpful in extracting non-linear relations from complex data such as the interaction between language and cognition. Moreover, the current study contributes with additional evidence for the debate of whether the presence of an additional language affects general cognitive and linguistic performance. As far as these authors are concerned, this is the first study making use of machine learning techniques to classify individuals based on their bilingual and multilingual status using their cognitive and linguistic performance. The computed model could successfully classify bilinguals and multilinguals with an accuracy of 77%, which is considered a moderate effect

General cognitive performance were not important features in the classification of bilinguals and multilinguals, while linguistic performances were. Thus, it seems that additional languages do not affect one's general cognitive performance. Our findings are in line with some of the previous literature, providing evidence that additional language(s) may provide an advantage in linguistic processing (e.g., Patra et al., 2020) but not in general cognitive capacity (Jones et al., 2021; Von Bastian et al., 2016).

However, some caution is necessary to interpret these results since the general cognitive performance tests used in this study are limited to non-verbal processing speed (i.e., auditory and visual reaction time), working memory (i.e., auditory and visual-spatial domains), inhibition, and abstract reasoning skills. In order to achieve a full understanding of the cognitive abilities of bilinguals and multilinguals non-verbal and verbal cognitive tests should be considered.

Moreover, no information has been provided about possible participants' disorders which may compromise cognitive performance, such as ADHD. Previous literature reports that bilinguals with ADHD have a decreased executive functioning in comparison to ADHD monolinguals, suggesting that an additional language could provide an extra cognitive burden for those individuals (Mor et al., 2015). Consequently, future studies should consider such diagnostic information in their data collection.

Regarding their linguistic abilities, multilingual individuals were significantly better than bilinguals in the p**icture naming, prescriptive grammar and verbal fluency tests.**

In the picture naming test, participants were instructed to name a photograph of an object as fast as possible. While in the verbal fluency test, participants must correctly name as many animals, food and drinks as they could within one minute. Regarding the picture naming and verbal fluency tests, multilinguals were significantly faster and produced more words than bilinguals respectively. This is in line with findings from the previous literature reporting a better performance of bilinguals in comparison to monolinguals in word production when the lexical representation of the concept at hand is known in both languages (Gollan et al., 2005; Potter et al., 1984). Similar reasoning could be applied to our results. Multilinguals show a facilitation effect from the other non-activated languages on the activation of the lexical representation in the target language. As such, the more languages one individual has encoded the more lexical representations could help to activate the lexical representation in the target language.

Concerning the prescriptive grammar test, in which participants must judge whether auditory presented Dutch sentences were grammatically correct or not, multilinguals also outperformed bilinguals. A greater metalinguistic awareness has been previously reported for bilinguals in comparison to monolinguals (Bialystok, 1988; Bialystok et al., 2003) and the more languages an individual speaks the more metalinguistic awareness and strategy is used to acquire additional languages (Jessner, 2014; Kemp, 2007). This greater linguistic awareness in multilinguals in comparison to bilinguals could explain their better performance in the grammatical judgment test.

One interesting result of our analysis is the higher accuracy of the xgboost model in classifying multilinguals in comparison to bilinguals. When considering the performances of bilinguals and multilinguals in the three most significant linguistic processing tests for the model, it becomes clear that the standard deviation of the bilingual group is greater than of the multilinguals. This could have led the model to be less accurate when classifying this population in comparison to multilinguals.

Moreover, the model accuracy could have been improved if the information about participants' language proficiency and their precise age of acquisition for each language would be available and added as features. As previous literature reports, language proficiency (Athanasopoulos, 2007) and age of acquisition (Bylund et al., 2019) may interact with cognitive processes and language skills.

Furthermore, in this study, multilinguals are defined as those who can communicate in or comprehend more than two languages. Consequently, individuals who can use three languages or more were all categorized in the same multilingual group. Differences in performance, especially, in the general cognitive domain, may start to appear as the number and the proficiency of mastered languages increase, which should be further investigated in future studies.

It may also be beneficial to improve the model performance to differentiate the multilinguals based on their language combination. That is, similar languages might have a positive transfer in one's linguistic skills and performance (e.g. Hipfner-Boucher et al., 2016), while a combination of very distinct languages might have a negative one (e.g., Robertson, 2000). This could be the case because language similarities could yield a transfer effect from one language to the other, facilitating the target language activation. Even though language information is available in the dataset (please see Table 2), its size together with the various language combinations would compromise the analysis power if language subgroups would be created. Larger datasets should be used instead.

In addition to that, while there might be a linguistic processing difference between the bilinguals and multilinguals of the tested population, i.e., young adults, whether similar effects could be found in aging populations remains unknown. Previous literature has linked the use of an additional language to higher cognitive reserves (Schweizer et al., 2012) and slower cognitive decay (Gold et al., 2013). In this sense, it would be interesting to investigate whether differences in general cognitive and linguistic abilities could be subject to additional languages in aging populations. Perhaps the various learning strategies used by multilinguals in the long term (Kemp, 2007), could better retain their cognitive abilities and linguistic skills at the old age in comparison to bilinguals.

Therefore, additional research should be carried out using machine learning techniques to analyze the cognitive and linguistic performance of bilinguals and multilinguals of different age, language background, proficiency and age of acquisition. As such, a better understanding of the complex interaction between language and cognition can be gained.

References

Athanasopoulos, P. (2007). Interaction between grammatical categories and cognition in bilinguals: The role of proficiency, cultural immersion, and language of instruction. *Language and Cognitive Processes*, *22*(5), 689–699. https://doi.org/10.1080/01690960601049347

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. *Developmental Psychology*, *24*(4), 560–567. https://doi.org/10.1037/0012-1649.24.4.560

Bialystok, E., Craik, F. I. M., Green, D. W., & Gollan, T. H. (2009). Bilingual Minds. *Psychological Science in the Public Interest*, *10*(3), 89–129. https://doi.org/10.1177/1529100610387084

Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, Aging, and Cognitive Control: Evidence From the Simon Task. *Psychology and Aging*, *19*(2), 290–303. https://doi.org/10.1037/0882-7974.19.2.290

Bialystok, E., Majumder, S., & Martin, M. M. (2003). Developing phonological awareness: Is there a bilingual advantage? *Applied Psycholinguistics*, *24*(1), 27–44. https://doi.org/10.1017/S014271640300002X

Blom, E., Küntay, A. C., Messer, M., Verhagen, J., & Leseman, P. (2014). The benefits of being bilingual: Working memory in bilingual Turkish–Dutch children. *Journal of Experimental Child Psychology*, *128*, 105–119. https://doi.org/10.1016/j.jecp.2014.06.007

Blumenfeld, H. K., & Marian, V. (2014). Cognitive control in bilinguals: Advantages in Stimulus–Stimulus inhibition*. *Bilingualism: Language and Cognition*, *17*(3), 610–629. https://doi.org/10.1017/S1366728913000564

Bosch, L., & Sebastián-Gallés, N. (2003). Simultaneous Bilingualism and the Perception of a Language-Specific Vowel Contrast in the First Year of Life. *Language and Speech*, *46*(2–3), 217–243. https://doi.org/10.1177/00238309030460020801

Bylund, E., Abrahamsson, N., Hyltenstam, K., & Norrman, G. (2019). Revisiting the bilingual lexical deficit: The impact of age of acquisition. *Cognition*, *182*, 45–49. https://doi.org/10.1016/j.cognition.2018.08.020

Calvo, N., García, A. M., Manoiloff, L., & Ibáñez, A. (2016). Bilingualism and Cognitive Reserve: A Critical Overview and a Plea for Methodological Innovations. *Frontiers in Aging Neuroscience*, *7*. https://www.frontiersin.org/article/10.3389/fnagi.2015.00249

Cenoz, J. (2013). Defining Multilingualism. *Annual Review of Applied Linguistics*, *33*, 3–18. https://doi.org/10.1017/S026719051300007X

Chen., T., He, T., Benesty, M., Khotilovich, V., Tang, V., & Cho, H. (2021). *XGBoost: Extreme gradient boosting*. (Version 04-2 2015: 1-4.) [Computer software].

Cook, R. in A. L. V., Cook, V., & Bassetti, B. (2011). *Language and Bilingual Cognition*. Psychology Press.

Friesen, D. C., Latman, V., Calvo, A., & Bialystok, E. (2015). Attention during visual search: The benefit of bilingualism. *International Journal of Bilingualism*, *19*(6), 693–702. https://doi.org/10.1177/1367006914534331

Gold, B. T., Kim, C., Johnson, N. F., Kryscio, R. J., & Smith, C. D. (2013). Lifelong Bilingualism Maintains Neural Efficiency for Cognitive Control in Aging. *Journal of Neuroscience*, *33*(2), 387–396. https://doi.org/10.1523/JNEUROSCI.3837-12.2013

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition*, *33*(7), 1220–1234. https://doi.org/10.3758/BF03193224

Hintz, F., Dijkhuis, M., van 't Hoff, V., McQueen, J. M., & Meyer, A. S. (2020). A behavioural dataset for studying individual differences in language skills. *Scientific Data*, *7*(1), 429. https://doi.org/10.1038/s41597-020-00758-x

Hipfner-Boucher, K., Pasquarella, A., Chen, X., & Deacon, S. H. (2016). Cognate Awareness in French Immersion Students: Contributions to Grade 2 Reading Comprehension. *Scientific Studies of Reading*, *20*(5), 389–400. https://doi.org/10.1080/10888438.2016.1213265

Jared, D., & Szucs, C. (2002). Phonological activation in bilinguals: Evidence from interlingual homograph naming. *Bilingualism: Language and Cognition*, *5*(03), 225–239. https://doi.org/10.1017/S1366728902003024

Jessner, U. (2014). On Multilingual Awareness or Why the Multilingual Learner is a Specific Language Learner. In M. Pawlak & L. Aronin (Eds.), *Essential Topics in Applied Linguistics and Multilingualism: Studies in Honor of David Singleton* (pp. 175–184). Springer International Publishing. https://doi.org/10.1007/978-3-319-01414-2_10

Jones, S. K., Davies-Thompson, J., & Tree, J. (2021). Can Machines Find the Bilingual Advantage? Machine Learning Algorithms Find No Evidence to Differentiate Between Lifelong Bilingual and Monolingual Cognitive Profiles. *Frontiers in Human Neuroscience*, *15*. https://www.frontiersin.org/article/10.3389/fnhum.2021.621772

Kemp, C. (2007). Strategic Processing in Grammar Learning: Do Multilinguals Use More Strategies? *International Journal of Multilingualism*, *4*(4), 241–261. https://doi.org/10.2167/ijm099.0

Kuhn, M. (2020). *Caret: Classification and Regression Training. R package* (6.0-86.) [Computer software].

Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin*, *144*(4), 394–425. https://doi.org/10.1037/bul0000142

Li, W. (2008). Research perspective in bilingualism and multilingualism. In M. G. Moyer & W. Li (Eds.), *The Blackwell Guide to Research Methods in Bilingualism and Multilingualism*. John Wiley & Sons.

Lv, Z., Qiao, L., & Singh, A. K. (2020). Advanced machine learning on cognitive computing for human behavior analysis. *IEEE Transactions on Computational Social Systems*.

Marx, N., & Hufeisen, B. (2003). » Multilingualism: Theory, Research Methods and Didactics «. *New Visions in Foreign and Second Language Education*, 178–203.

Mor, B., Yitzhaki-Amsalem, S., & Prior, A. (2015). The Joint Effect of Bilingualism and ADHD on Executive Functions. *Journal of Attention Disorders*, *19*(6), 527–541. https://doi.org/10.1177/1087054714527790

Noort, M. W. M. L. van den, Bosch, P., & Hugdahl, K. (2006). Foreign Language Proficiency and Working Memory Capacity. *European Psychologist*, *11*(4), 289–296. https://doi.org/10.1027/1016-9040.11.4.289

Patra, A., Bose, A., & Marinis, T. (2020). Performance difference in verbal fluency in bilingual and monolingual speakers. *Bilingualism: Language and Cognition*, *23*(1), 204–218. https://doi.org/10.1017/S1366728918001098

Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychological Monographs: General and Applied*, *76*(27), 1–23. https://doi.org/10.1037/h0093840

Potter, M. C., So, K.-F., Eckardt, B. V., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior*, *23*(1), 23–38. https://doi.org/10.1016/S0022-5371(84)90489-4

Robertson, D. (2000). Variability in the use of the English article system by Chinese learners of English. *Second Language Research*, *16*(2), 135–172. https://doi.org/10.1191/026765800672262975

Saer, D. J. (1923). The effect of bilingualism on intelligence. *British Journal of Psychology: General Section*, *14*, 25–38.

Schweizer, T. A., Ware, J., Fischer, C. E., Craik, F. I. M., & Bialystok, E. (2012). Bilingualism as a contributor to cognitive reserve: Evidence from brain atrophy in Alzheimer's disease. *Cortex*, *48*(8), 991–996. https://doi.org/10.1016/j.cortex.2011.04.009

Von Bastian, C. C., Souza, A. S., & Gade, M. (2016). No evidence for bilingual cognitive advantages: A test of four hypotheses. *Journal of Experimental Psychology: General*, *145*(2), 246–258. https://doi.org/10.1037/xge0000120