# Measuring Verbal Fluency Task Performance of Indonesian Bilinguals

**FEISAL AZIEZ[1], HAZIM ALKHRISHEH[2], FURQANUL AZIEZ[3], AND MAULANA MUALIM[4]**

**Abstract**

Verbal fluency (VF) task is a tool that has been utilized in bilingual studies to measure executive performance of bilingual individuals. While many of the previous studies compared bilingual speakers' performance in VF to their monolingual counterparts, this descriptive study compares exclusively the cross-linguistic VF performance of Indonesian bilinguals in their first (L1), second (L2), and third language (L3). The aim of this study was to see whether or not there were any differences in the result of their performance in each of their languages. The participants of the current study were 25 non-native bilingual university students in Indonesia who speak English in at least intermediate level. Javanese is the L1 of the participants. Indonesian is their L2, whereas English is their L3. In the VF task, the participants were requested to generate, in a rapid fashion, semantic category and phonemic category in their L1, L2, and L3. The results showed that the participants' VF performance in English and standard Indonesian were significantly higher (p < .05) than Javanese. However, no significant difference was indicated when comparing the semantic category and phonemic category in all three languages. The findings of this study will be used as a basis for a forthcoming study on VF performance of Indonesian bilinguals.

**Keywords**: Indonesian bilinguals, phonemic category, semantic category, verbal fluency

[1.] A PhD. student at Multilingualism Doctoral School, University of Pannonia, Hungary. He is also a lecturer at English Education Department, Universitas Muhammadiyah Purwokerto, Indonesia; feisalaziez@ump.ac.id

[2.] A PhD. student at Multilingualism Doctoral School, University of Pannonia, Hungary; hkhrisheh@yahoo.com

[3.] Associate professor at Faculty of Teacher Training and Education, Universitas Muhammadiyah Purwokerto, Indonesia; f.aziez2010@gmail.com

[4.] Lecturer at English Education Department, IAIN Purwokerto, Indonesia; maulanamualim@iainpurwokerto.ac.id

### Introduction

There are several different ways in which bilingualism is described (Romaine, 1995). Haugen (1953) says that bilingualism is apparent when a language speaker has the ability produce meaningful statements in a different language. Bloomfield (1933) notes that bilingualism is possible only if the "native" mastery of both the languages has been acquired by the speaker. Researchers also characterize bilinguals by the speaker's level of proficiency in their languages. For example, the term "balanced" bilingual means someone who has relatively equal proficiency proficient in two or more languages. In comparison, bilinguals who are "non-balanced" have different proficiency in their languages meaning that they are usually more proficient in a language compared to their other language(s) (Romaine, 1995). Across the globe, it has been reported that the majority of the world's population routinely speak two languages or more (Moreno & Kutas, 2005) without apparent difficulty to bear the burden of bilingualism. In the United States, bilingualism is much less common, but the bilingual population (around one fifth of the whole population) is relatively large and increasing in a quick fashion (US Census, 2000). The nature of bilingualism offers a chance to explore their language production by questioning how bilingualism affects bilinguals' executive control which can be reflected in their ability to produce words in rapid fashion (e.g. verbal fluency).

Bilinguals may appear to be able to use two or more languages easily in everyday use at a high degree of proficiency. Nonetheless, there are some processing costs that have been related to bilingualism. According to some studies, bilinguals tend to score less compared to monolinguals in standardized tests including the Boston Naming Test (e.g. Gollan et al., 2007) and picture naming tasks (e.g., Silverberg, Gollan and Silverberg, 2001). They are also reported to display more tip-of-the-tongue (TOT) (e.g., Gollan and Silverberg, 2001) and have poor verbal fluency performance (e.g. Gollan et al., 2002). Crucially, even when they are tested on their dominant languages (e.g., Gollan et al., 2005), and native languages, bilingual people have been reported to be relatively less fluent than monolinguals (e.g., Ivanova & Costa, 2008).

As a multilingual country, there has been little study conducted in Indonesia in regards to the verbal fluency performance of the bilingual speakers in the country. In the current study, we assessed the verbal fluency performance of non-native bilingual university students in Indonesia who speak English in at least intermediate level, using a verbal fluency task—a commonly used neuropsychological task of lexical knowledge or lexical retrieval ability (Friesen et al., 2013). Their performance in their first, second, and third language were measured to see whether there was any indication of processing costs. On those bases, the questions that this study attempted to answer were (1) is there any difference in the speakers' VF performance in Javanese, Indonesian and English? (2) is there any difference between semantic categories and phonemic categories regardless of the languages? (3) is there any difference among all scores in the semantic and phonemic categories in all of the languages?

**Literature Review**

*Benefits and costs of bilingualism*

Many studies in bilingualism have pointed out the cognitive benefits of bilingualism both in children and adults (e.g., Lauchlan et al., 2012; and Antoniou et al., 2016). Lauchlan et al. (2012) investigated the benefits of bilingualism on the cognitive ability of children who are the speakers of Sardinian and Scottish Gaelic, two minority languages as well as Italian and English, which is their country's 'national' languages. Their study indicated that these bilingual children exceeded the monolingual children in the cognitive tasks to a significant degree. Antoniou et al. (2016), for example, compared the performance in executive control of bidialectal children to that of bilingual children (English–Greek speaking) in Cyprus and monolingual children (Standard Modern Greek-speaking) in Greece. They found that speaking more than one dialect bears similar benefits as speaking more than one language. Both bidialectal and bilingual children outperformed monolingual children in executive performance.

However, in some other researches (e.g. Kirk et al., 2014), the results showed no significant difference in executive control performance between bilinguals and monolinguals. Kirk et al. (2014) suggested that there is a possibility that the inconsistent findings of bilingual advantages in terms of executive control are caused by the impact of interactional contexts and bilingual literacy, to which bilinguals are generally exposed to. Moreover, some research even reported that bilinguals scored less compared to monolinguals in standardized tests such as the Boston Naming Test (e.g., Gollan et al., 2007) and picture naming tasks (e.g. Gollan & Silverberg, 2001). They are also reported to display more tip-of-the-tongue (TOT) (e.g., Gollan & Silverberg, 2001) and have poor verbal fluency performance (e.g. Gollan et al., 2002). Crucially, even when they are assessed on their dominant languages (e.g., Gollan et al., 2005), and native languages, bilingual people have been reported to be relatively less fluent than monolinguals (e.g., Ivanova & Costa, 2008).

Despite the confirmation from recent research on the existence of bilingual disadvantage in the VF task, the explanation to this problem remains elusive (e.g., Bialystok et al., 2008; Portocarrero et al., 2007). In VF task (see Benton et al., 1983), speakers usually have to name, in one minute, as many words as possible in the tested language which belong to a semantic category (e.g. clothing items, furniture, animals, etc.) or letter category (e.g. words beginning with letter, S, F, or A). The most possible explanation on the disadvantages of bilingualism in VF task is that bilinguals may need to retrieve the expected items while, at the same time, attempting to suppress interruption from the non-target language(s) in their mind. Unintentional activation of non-target language words may interrupt the retrieval of target language words resulting in fewer correct responses than monolinguals (Sandoval et al., 2010). Another potential reason is that in retrieving target language words without any intervention from the non-target language, bilinguals are merely slower than monolinguals. (e.g., Gollan et al., 2008). Another alternative is the differences between monolinguals and

bilinguals in terms of language specific vocabulary knowledge. Although, overall, bilinguals may know more words than monolinguals, bilinguals may know fewer words than monolinguals in each language that they have (e.g., Bialystok et al., 2008). Undoubtedly, these different mechanisms may occur at the same time, affecting bilinguals' verbal fluency performance.

### *Verbal fluency task*

In measuring lexical knowledge or lexical retrieval ability, psychologists and psycholinguists commonly used a type of neuropsychological task called verbal fluency (VF) task (Friesen et al., 2013). Semantic fluency and phonemic fluency tasks are two types of VF tasks that are most commonly used (Shao et al., 2014). In a semantic fluency test, participants are asked to produce, in the target language, as many unique words as possible in a semantic category (e.g. animals, girl's names, etc.), whereas in a phonemic fluency test, they are asked to produce as many unique words as possible which starts with a particular letter (e.g. words starting with A, S, D, etc.) in a limited time which is typically 60 seconds (Lezak et al., 2004). The correct words produced within each category are then counted and used as the score. However, other studies use other measurements such as the number of repetitions, the length of words in each subcategory, the reaction time between the words, and so on (e.g., Shao et al., 2014; Troyer et al., 1997; Troyer et al., 1998). The COWAT (Controlled oral word association test) which is the initial letter fluency is the most common variant employed to investigate VF (Loonstra et al., 2001).

VF task is often used in clinical practice e.g. to support diagnoses of Alzheimer's disease (e.g., Zhao et al., 2013), Parkinson's disease (e.g., Henry & Crawford, 2004), Schizophrenia (e.g., Frith et al., 1995) and neuropsychological assessment (e.g. Vaucheret Paz et al., 2020). While in non-clinical research, the task is used, for example, to compare lexical retrieval ability in monolinguals and bilinguals (e.g., Friesen et al., 2013; Patra et al., 2020), between genders (e.g., Scheuringer, 2017), age (Gaillard et al., 2000) or a combination of different factors such as age and education (e.g., Tombaugh et al., 1999). In addition, VF tasks are also used to assess bilingual processing in psychological studies (Friesen et al., 2016). The validity of the task comes from evidence related to VF (in addition to executive control abilities) to brain damage (Schwartz & Baldo, 2001). Wysokinski et al. (2010: 438) defined VF as the individual's ability to utter words and expressions with required criteria associated with communication and functioning. Generating words in VF tasks usually requires individuals to exercise cognitive activation such as short-term memory retrievals, inhibitions, and organizing information.

VF tasks are believed to require search strategies from the test takers that utilizes their executive control during the lexical retrieving process and Friesen et al. (2013) argue that the demand for executive control is greater in phonemic category tasks. This statement is supported by some researches which indicate that individuals tend to produce fewer lexical items during phonemic tasks than semantic tasks (e.g., Gollan et al., 2002; and Kormi-Nouri et al., 2012; Alkhrisheh & de Bot, 2019). One potential explanation for this is that the role

specifications for semantic categories are consistent with the semantic memory structure; in one's mind, concepts are clustered based on semantic properties which help during the lexical retrieval process e.g. for speech production (Luo et al., 2010). On the other hand, generating words from a phonemic cue is an unusual lexical retrieval approach since lexical entries are not listed in alphabetical order (Strauss et al., 2006). Shao et al. (2014) attribute the higher cognitive performance demanded by the phonemic category to the lack of existing links between the concepts. The semantic category seems to provoke responses based on already existing links between the concepts. Thus, in the phonemic fluency tasks, participants are required to suppress any exploit semantic associations, and instead, turn to novel retrieval strategies. Furthermore, the VF task is widely spread because of its validity and association with executive control functioning in which mental access to the lexicon is required to fulfill the task (Fisk & Sharp, 2004). The aspects of executive control process in VF tasks require participants to focus on the task in order to generate words according to criteria while avoiding repetition. This process involves aspects of executive control abilities including working memory, inhibition, and self-initiation effort. The developmental dynamics of the two variants vary greatly. Studies show that the development of the semantic variant is more stable and rapid. It also shows a linear developmental sequence at the age of 12-13 whereas the phonemic variant seems to be less rapid and lasts until adolescence (Kavé, 2006; Sauzeon et al., 2004).

Even though VF tasks are commonly employed to assess cognitive deficiencies as mentioned earlier and demonstrated by the previous readings, yet other researchers used the task to assess individual differences considering variables such as gender, age, and the education level as factors that might contribute to the performance of the VF task. Heister (1982), for instance, conducted a study to compare males to females in word fluency and ideational fluency to find that females are superior to males in terms of performance in lexical access. Weiss et al. (2003), on the other hand, in the VF task results, did not notice any substantial discrepancies between males and females. The reported mixed results are attributed to many reasons such as the interference in lexical decision tasks and tip-of-the-tongue occurrences (Bialystok, 2009). Connectionists, for instance, argued that bilinguals exhibit weaker associations between words and associated notions in each language since information is stored in terms of weights and activation networks that often compete because they have to control multiple lexical reservoir (Michael & Gollan, 2005). Studies in language-switching have also been reporting disadvantages of bilinguals in terms of language specific knowledge (e.g., Altarriba & Basnight-Brown, 2009). Accordingly, bilingualism seems to have costs. Most of these costs seem to be deficits in language performance and lexical retrievals.

Researches in psycholinguistics also made use of VF tasks to detect any forms of cognitive advantage for being bilingual. However, in standard measures of vocabulary tasks, it was observed that bilinguals scored less than monolinguals, experienced more tongue slips, and were slower in naming objects displayed in pictures (Bialystok & Luk, 2012). In contrast, it has been reported that bilinguals outperformed monolinguals in executive functioning tasks (Bialystok et al., 2009). The reported results imply that there are advantages and

disadvantages of bilingualism represented in cognitive performance on the expense of language decline. The assumption that there are cognitive advantages of bilingualism in certain aspects of cognition does not eliminate the possibility that there are cognitive costs of bilingualism in other aspects of cognition. It has been reported that bilinguals activated multiple sets of lexical representations. As a result, the process of retrieving a suitable lexical candidate for translation took longer than expected, suggesting that interference seems to play a role in lexical retrieval (Kroll & Stewart, 1994). In the present study, the VF task is used to measure the lexical retrieval ability of Indonesian bilinguals in their mother tongue (i.e., Javanese), second language (i.e., Indonesian), and third language (i.e., English).

### Multilingual condition in Indonesia

Indonesia is a multilingual country with over 700 languages (2010 census). The national language of the country is Indonesian, a standardized register of Malay, which belongs to Austronesian language groups. It has been used for centuries as a lingua franca mainly for trades across the archipelago. Most of the population in the country use Indonesian as their daily mean of communication which makes it one of the languages in the world with the most speakers. In formal education, administrative works, governance, nearly all national mass media, judiciary, and other purposes, Indonesian is widely used in the country. Based on the 2010 census by the government, Indonesian is the native language of 42.8 million people and as a second language by 154.9 million people who speak it alongside their regional languages and dialects. This means the total number of its speaker is nearly 200 million. Indonesian is a common first language in urban areas while in more rural parts of Indonesia; it is spoken as a second language. As the official language, it is regulated by the Constitution of Indonesia in Chapter XV, 1945 which says: "The national language shall be Indonesian". It is also regulated in Chapter III, Section 25 to 45, Government regulation No. 24/2009. The language has an important role as a representation of national identity. It also has an imperative role as a unifying language in the diverse and multilingual Indonesian ethnic groups spread across the archipelago.

In 1945, the country established *Bahasa Indonesia* as their official language. The use of Indonesian as the national language of the country was uncommon when compared to most other post-colonial countries or states. Indonesia did not adopt Dutch as the official language. They also did not adopt Javanese as the language with most native speakers in the country. At the time, Indonesian was only about 5% of the population's native language, while Javanese was 42-48% of the population's mother tongue. It was a mixture of practical, nationalistic, and political reasons that ultimately led to Indonesian being adopted as a national language. Although at that time, Javanese has the most native speakers and it was a predominant language used in political, economic, judicial, religious, and literary settings, it lacked the fundamentals to unify the diverse population of Indonesia. Not only was that Javanese considered too difficult to learn by non-native speakers, Indonesian had also already widely spread in the archipelago at that time. Therefore, it could be more readily

embraced than any other language in the country. However, Javanese and its dialects are still widely used today in daily communication by more than 84 million Indonesians.

As stated earlier, apart from speaking the national language, the majority of Indonesians speak more than 700 indigenous regional languages fluently. The government even boasts that the number could reach 2500 if we include all the dialects of these languages. According to the 2010 survey by the government, 79.5 percent of the population of age above 5 use regional languages at home, while only 19.9 percent of the population use Indonesian. According to Ethnologue, here are ten regional languages with most speakers in Indonesia (van den Berg, 2014).

**Table 1.** *Regional languages in Indonesia with most speakers*

| No. | Language | Number of Speakers |
|-----|----------|--------------------|
| 1 | Javanese | 84.3 million |
| 2 | Sundanese | 34.0 million |
| 3 | Madurese | 13.6 million |
| 4 | Minangkabau | 5.5 million |
| 5 | Musi | 3.9 million |
| 6 | Manado Malay | 3.8 million |
| 7 | Bugis | 3.5 million |
| 8 | Bandar | 3.5 million |
| 9 | Acehnese | 3.5 million |
| 10 | Balinese | 3.3 million |

Some of the regional languages also have their own writing systems or alphabets such as *Batak, Javanese, Makassar, and Bugis* whereas *Aceh, Malay, and Wolio* adopted Arabic alphabets. However, the use of the regional languages' writing systems has been significantly decreasing in the last few decades. This is due to the widespread of the Indonesian writing system in education and mass media as the national language. Javanese, for example, has been taught in the Indonesian writing system instead of its writing system for decades. In the country, all regional languages are acknowledged as an important part of the culture which has to be respected and conserved. However, in practice, Indonesian as the national language is highly prioritized while the regional languages are given limited space in the curriculum or even completely neglected. Since 1995 there has been an effort to include regional languages as a local content in the curriculum. However, the teaching is limited to the elementary (grade 1-6) and junior high school (grade 7-9) level and only regional languages with a high number of speakers are included. Some minority languages such as *Wawonii, Kulisusu, Kamaru, and Busoa* are completely neglected. This happens due to government and school policies, lack of teaching materials, limited teaching staff, *etc.* (van den Berg, 2014). English is a compulsory subject taught at almost all levels of education in Indonesia (Lauder, 2008) for its role as a mean for global communication. The participants in this study were third-year university students of English language programs in a private university in Central Java in which Javanese is the mother tongue of the majority of the

population. Therefore, they have been learning English exclusively during the three years. They also still used Javanese in their daily communication while Indonesian is mainly used mostly in formal situations. The multilingual nature of the context is therefore interesting for exploring the VF performance across the languages.

**Methodology**

***Research design, participants, and locale of the study***

This study used a quantitative method in describing cross-linguistic and cross-categorical performances of Indonesian bilinguals. In this study, a homogeneous participant sample was selected through non-probability, convenience sampling (Bhattacherjee, 2012). Language History Questionnaire (LHQ) version 3.0 was used to ensure that the bilingual participants were approximately equal in proficiency in their languages. Since this study attempted to provide insight to the VF performance of one particular group of bilinguals in Indonesia and no comparison was made, no a priori claims should be made about the generalizability of this study. There were 25 participants involved in this research. They were third-year students of an English education department in a private university in Central Java, Indonesia. Their age range was from 19 to 21. The participants speak at least three languages; Javanese (L1), Indonesian (L2), and English (L3). They are mostly early bilinguals—having been exposed to Javanese and Indonesian at an early age. English is learned at school from the elementary school level. Their English proficiency levels at the time of data collection were B1 and B2 (CEFR).

***Data collection and analysis***

The VF task in the present study followed the existing steps that have been used in many previous studies (e.g., Friesen et al, 2013; Lezak, et al. 2004; Shao et al, 2014). First, participants were asked to produce, in rapid fashion, semantic category (i.e. job) and phonemic category (i.e., names of different words beginning with S) in their L1, L2, and L3. The participants were given 60 seconds to produce as many words as possible in each category and each language. Before performing the actual task, the participants were given one practice task with a different set of categories to allow them to get used to the task. The responses from the participants while performing the actual tasks were recorded. The scores for each participant were measured by counting the correct and unique words produced by her/him. To ensure the reliability of the correct words produced, two raters were asked to assess them. The scores then were analyzed using one-way ANOVA in SPSS 22. The task scores from all categories and all languages were compared to each other to see whether or not there is any significant difference. Once a significant difference was found, Tukey Post-Hoc Tests were performed to see which of the comparisons were significant.

### *Ethical considerations*

The present study, for ethical considerations, required full consent from the participants prior to the study. Participants were informed accordingly about the purpose of the present study and assured that participation was voluntary and they have rights to withdraw from the study at any stage. This study was self-funded and there was no conflict of interests. Information on all participants and organization are kept confidential except for the information used for the purpose of the data analysis with the informed consent from the respected parties. This study also ensured an adequate level of confidentiality of the research data.

### **Findings**

Abbreviations, for practical reasons, will be used to describe the variables in this study i.e., SemJav is semantic category in Javanese language; SemInd is semantic category in standard Indonesian; SemEng is semantic category in English; PhoJav is semantic category in Javanese language; PhoInd is phonemic category in standard Indonesian; and PhoEng is phonemic category in English.

### ***Comparing bilinguals' VF performance in Javanese, Indonesian, and English***

Table 2 shows that the participants produced more words in English (M=11.98, SD=3.74) on both semantic and phonemic categories, which is slightly higher than in Indonesian (M=11.80, SD=4.45), whereas they produced least words in their Javanese (M=7.34, SD=3.56).

**Table 2.** *Descriptive statistics words production by language*

| Language | Number of words produced | | |
|---|---|---|---|
| | *n* | *M* | *SD* |
| Javanese | 50 | 7.34 | 3.561 |
| Indonesian | 50 | 11.80 | 4.445 |
| English | 50 | 11.98 | 3.744 |

The result of the analysis of variance indicated a significant difference (F (22.309) = 2, p= .000) between the languages regardless of the semantic and phonemic categories as shown in the following table.

**Table 3.** *ANOVA of words production by language*

|  | *df* | *SS* | *MS* | *F* | *p* |
|---|---|---|---|---|---|
| Between Groups | 2 | 690.893 | 345.447 | 22.309 | .000 |
| Within Groups | 147 | 2276.200 | 15.484 |  |  |

To see between which language(s) the difference is significant, post hoc analysis was carried out using a Tukey, a .05 level of significance was then achieved. We can see from the table below that when compared to Javanese (M=7.34, SD=3.56) the scores of English (M=11.98, SD=3.74) and Indonesian (M=11.80, SD=4.45) are significantly higher. However, no significant difference is found between English and Indonesian. The following table shows the results of all the multiple comparisons among the languages.

**Table 4.** *ANOVA comparisons of word production from three languages*

|  |  |  |  | Tukey's HSD Comparison | |
|---|---|---|---|---|---|
| Language | n | M | SD | Javanese | Indonesian |
| Javanese | 50 | 7.34 | 3.561 |  |  |
| Indonesian | 50 | 11.80 | 4.445 | < .001 |  |
| English | 50 | 11.98 | 3.744 | < .001 | >1.000 |

### Comparing bilinguals' VF performance in semantic and phonemic categories

On category level, we found that the participants produced slightly more words on phonemic category (M=10.99, SD=4.26) than semantic category (M=9.76, SD=4.61) as seen in table 3 below.

**Table 5.** *Descriptive statistics words production by category*

| Category | Number of words produced | | |
|---|---|---|---|
|  | n | M | SD |
| Semantic | 75 | 9.76 | 4.606 |
| Phonemic | 75 | 10.99 | 4.257 |

As seen in the following table, the result of the analysis of variance on semantic category and phonemic category regardless the languages showed no significant difference (F (2.87) = 1, p= .092).

**Table 6.** *ANOVA of words production by semantic and phonemic categories*

|                | df  | SS       | MS     | F     | p    |
|----------------|-----|----------|--------|-------|------|
| Between Groups | 1   | 56.427   | 56.427 | 2.869 | .092 |
| Within Groups  | 148 | 2910.667 | 19.667 |       |      |

### Comparing VF performance across languages and categories

Overall, the participants produced more words in phonemic category than semantic category of all languages. Phonemic category in Indonesian (PhoInd) (M=12.68, SD=3.62) is higher than all other categories with the lowest being semantic category in Javanese (SemJav) (M=6.84, SD=3.27). In semantic category, learners produced more words in semantic category in English (SemEng) (M=11.52, SD=3.95) than semantic category in Indonesian (SemInd) (M=10.92, SD=5.06) and semantic category in Javanese (SemJav) (M=6.84, SD=6.84). In phonemic category, as mentioned earlier, phonemic category in Indonesian (PhoInd) (M=12.68, SD=3.62) is higher than phonemic category in English (PhoEng) (M=12.44, SD=3.55) and phonemic category in Javanese (PhoJav) (M=7.84, SD=3.83). The following table shows the result from the descriptive statistics.

**Table 7.** *Descriptive statistics words production based on category and language*

| Category and Language | Number of words produced | | |
|-----------------------|------|-------|-------|
|                       | n    | M     | SD    |
| SemJav                | 25   | 6.84  | 3.275 |
| SemInd                | 25   | 10.92 | 5.057 |
| SemEng                | 25   | 11.52 | 3.949 |
| PhoJav                | 25   | 7.84  | 3.826 |
| PhoInd                | 25   | 12.68 | 3.625 |
| PhoEng                | 25   | 12.44 | 3.548 |

To see the statistical significance of the differences among the scores of all categories, an analysis of variance was performed. The result indicated a significant difference (F (9.79)= 5, p= .000) in the categories as shown in the following table.

**Table 8.** *ANOVA of words production by all categories and languages*

|  | df | SS | MS | F | p |
|---|---|---|---|---|---|
| Between Groups | 5 | 752.693 | 150.539 | 9.789 | .000 |
| Within Groups | 144 | 2214.400 | 15.378 |  |  |

To find out which of the categories are significantly different compared to others, Post hoc analysis was conducted using a Tukey and a .05 level of significance was achieved.

**Table 9.** *ANOVA comparisons of word production by all categories and languages*

| Category and Language | n | M | SD | Tukey's HSD Comparison | | | | |
|---|---|---|---|---|---|---|---|---|
|  |  |  |  | SemJav | SemInd | SemEng | PhoJav | PhoInd |
| SemJav | 25 | 6.84 | 3.275 |  |  |  |  |  |
| SemInd | 25 | 10.92 | 5.057 | .005 |  |  |  |  |
| SemEng | 25 | 11.52 | 3.949 | .001 | >1.000 |  |  |  |
| PhoJav | 25 | 7.84 | 3.826 | >1.000 | 0.93 | .001 |  |  |
| PhoInd | 25 | 12.68 | 3.625 | <.001 | >1.000 | >1.000 | <.001 |  |
| PhoEng | 25 | 12.44 | 3.548 | <.001 | >1.000 | >1.000 | >1.000 | >1.000 |

As seen in table 9, SemJav (M=6.84, SD=3.27) is significantly lower compared to other categories except for PhoJav (M=7.84, SD=3.83). PhoJav (M=7.84, SD=3.83) is also significantly lower in comparison to other categories except for SemJav (M=6.84, SD=3.27) and SemInd (M=10.92, SD=5.057). Although PhoInd (M=12.68, SD=3.63) is the highest among all, it is only significantly higher when compared to PhoJav (M=7.84, SD=3.83) and SemJav (M=6.84, SD=3.27). SemInd (M=10.92, SD=5.057) is only significantly higher than SemJav (M=6.84, SD=3.27). Similar to PhoInd, SemEng (M=11.52, SD=3.949) and PhoEng (M=12.44, SD=3.548) are significantly higher only when compared to PhoJav (M=7.84, SD=3.83) and SemJav (M=6.84, SD=3.27). The following table shows all the multiple comparisons among the categories.

### Discussion

There are three important points that can be drawn from the data found. Firstly, when comparing the languages regardless the semantic or phonemic category, it is found that the participants' VF performance in English and standard Indonesian was significantly higher than Javanese. Therefore, hypothesis (1) is accepted. This finding is interesting since Javanese is the participants' L1, which is mostly acquired before or alongside Indonesian (L2). In fact, the participants generated slightly more lexical items in English (M=11.98, SD=3.74) than Indonesian (M=11.80, SD=4.45) although no significance was found. There are some possible explanations for this finding. First, this could be a supportive argument which argues that bilinguals exhibit weaker links between words and associated concepts in each language since information is stored in terms of weights and activation networks that often compete as a result of having to manage more than one lexical reservoir (Michael & Gollan, 2005). Studies in language-switching also reported language specific deficits (e.g., Altarriba & Basnight-Brown, 2009). Since the participants were learning English at the time and Indonesian is also used in academic situation, their L1 is suppressed to allow executive control to focus more on L2 and L3.

Secondly, no significant difference between the semantic category and phonemic category regardless the language is found. Therefore, hypothesis (2) is rejected. This finding contradicts the studies where the participants tend to produce less lexical items in phonemic tasks than semantic tasks (e.g., Gollan et al, 2002; Kormi-Nouri et al, 2012). Despite no significance could be drawn, in fact, this study shows that the participants produced slightly more words on phonemic category (M=10.99, SD=4.26) than semantic category (M=9.76, SD=4.61). This might challenge Friesen's et al (2013) argument which stated that the demand for executive control is greater in phonemic category tasks. Of course, improving the number of participants in this study would be necessary for a stronger argument.

Finally, as it can be expected from the first analysis, there are several significant differences among the semantic and phonemic categories in all languages. Therefore, hypothesis (3) is accepted. To get a more detailed comparison among the categories and languages in the VF task, multiple comparisons were done. From the analysis, semantic category of Javanese (M=6.84, SD=3.27) is significantly lower compared to other categories except for phonemic category of Javanese (M=7.84, SD=3.83). On the other hand, phonemic category of standard Indonesian (M=12.68, SD=3.63) is the highest among all. However, it is only significant when compared to both semantic and phonemic category in Javanese.

Overall, the findings seem to support the recent findings on the costs of bilingualism. Even in their dominant language (i.e. Javanese), the participants produced less words than their less dominant language (i.e. English). This supports Gollan's et al. (2005) study which reported that bilinguals were less fluent than monolinguals even in the dominant and native languages (e.g., Ivanova and Costa, 2008). However, as it was previously mentioned, one cannot simply jump into conclusion on what causes this disadvantage (e.g.,

Portocarrero, Burright, and Donovick, 2007, Bialystok, Craik, and Luk, 2008). Although several possible explanations have been provided (i.e. retrieval slowing with interference between languages (Sandoval, Gollan, Ferreira, & Salmon, 2010); retrieval slowing without interference (Gollan, Montoya, Cera, & Sandoval, 2008); and the reduced vocabulary hypothesis (e.g., Bialystok et al., 2008)), assuming the actual cause can still be quite problematic.

## Conclusion and Recommendation

The findings in this study have shown some interesting insights into Indonesian bilinguals' VF performance across languages. They seem to support the recent findings on the costs of bilingualism. One important point is that the VF performance in L1 is lower than L2 and L3. This means that even in their dominant language (i.e. Javanese), the participants produced less words than their less dominant language (i.e. English). Secondly, there was no difference found in the production of lexical items in semantic and phonemic categories which are in contrast to common findings of several relevant studies. Although it has been suggested that inferring the possible cause(s) of the cost of bilingualism is still challenging, there is an indication from this study that mind conditioning may play a part in the VF performance. Since this study was conducted in an L2 English environment, the participants' mind might be affected by the linguistic environment which results in higher number of words produced in English. Therefore, future studies using some types of conditioning treatments and control group may reveal the extent of this factor. Moreover, since this study involves only one group of 25 English learners in a private university in Indonesia, the next study involving a larger number of participants would be crucial in providing more insights into this topic. Nonetheless, this study provides further insights on the cost of bilingualism particularly in the context of Indonesian bilinguals which may serve as a basis for future research in this topic.

## Disclosure statement

No conflict of interest was reported by the authors.

## Acknowledgments

We are deeply indebted to all reviewers and editors for accepting our manuscript.

## References

Alkhrisheh, H., & de Bot, K. (2019). The effect of maturity and gender on verbal fluency in Jordanian university students. *The Hungarian Journal of Applied Linguistics*, *19*(1), 1-19.

Altarriba, J. & Basnight-Brown, M. D. (2009). Empirical approaches to the study of code-switching in sentential contexts. In L. Isurin, D. Winford, & K. de Bot. (Eds.),

*Multidisciplinary approaches to code switching* (pp. 3-25). Amsterdam: John Benjamins Publishing Company.

Antoniou, K., Grohmann, K. K., Kambanaros, M., & Katsos, N. (2016). The effect of childhood bilectalism and multilingualism on executive control. *Cognition, 149*, 18-30.

Badan Pusat Statistik [Central Bureau of Statistics]. (2011). *Penduduk Indonesia hasil sensus penduduk 2010* [*Indonesian population results from the 2010 population census*]. Retrieved from http://sp2010.bps.go.id/index.php

Badan Pusat Statistik [Central Bureau of Statistics]. (2011). *Kewarganegaraan, suku bangsa, agama, dan bahasa sehari-hari penduduk Indonesia hasil sensus penduduk 2010* [The citizenship, ethnicity, religion, and everyday language of the Indonesian population based on the 2010 population census]. Retrieved from https://www.bps.go.id/publication/2012/05/23/55eca38b7fe0830834605b35/kewarganegaraan-suku-bangsa-agama-dan-bahasa-sehari-hari-penduduk-indonesia.htm

Benton, A. L., Hamsher, K. & Sivan, A. B. (1983*). Multilingual aphasia examination, 3rd ed.* Iowa City, IA: AJA Associates.

Bhattacherjee, A. (2012). *Social science research: Principles, methods, and practices*. Tampa: Creative Commons Attribution

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, *12*(1), 3–11.

Bialystok, E., Craik, F. I., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological science in the public interest*, *10*(3), 89-129.

Bialystok, E., Craik, F. I. M. & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *3*4(4), 859–873.

Bialystok, E., & Luk, G. (2012). Receptive vocabulary differences in monolingual and bilingual adults. *Bilingualism: Language and Cognition, 15*(2), 397-401.

Bloomfield, L. (1933). *Language.* New York: Holt.

Fisk, J. E., & Sharp, C. A. (2004). Age-related impairment in executive functioning: updating, inhibition, shifting, and access. *J Clin Exp Neuropsychol, 26*, 874–890.

Friesen, D.C., Luo, L., Luk, G., & Bialystok, E. (2013). Proficiency and control in verbal fluency performance across the lifespan for monolinguals and bilinguals. *Language, Cognition and Neuroscience*, *30*(3), 238–250.

Frith, C.D., Friston, K.J., Herold, S., Silbersweig, D., Fletcher, P.J., Cahill, C.M., Dolan, R.J., Frackowiak, R.S., & Liddle, P.F. (1995). Regional brain activity in chronic schizophrenic patients during the performance of a verbal fluency task. *The British journal of psychiatry: the journal of mental science*, *167*(3), 343-9.

Gaillard, W.D., Hertz-Pannier, L., Mott, S., Barnett, A., Lebihan, D., & Theodore, W. (2000). Functional anatomy of cognitive development: fMRI of verbal fluency in children and adults. *Neurology*, *54*(1), 180-5.

Gollan, T. H., Bonanni, M. P. & Montoya, R. I. (2005). Proper names get stuck on bilingual and monolingual speakers' tip-of-the-tongue equally often. *Neuropsychology, 19*, 278–287.

Gollan, T. H., Fennema-Notestine, C., Montoya, R. I. & Jernigan, T. L. (2007). The bilingual effect on Boston Naming Test performance. *Journal of the International Neuropsychological Society, 13*, 197–208.

Gollan, T. H., Montoya, R. I., & Werner, G. A. (2002). Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology, 16*, 562-576.

Gollan, T. H., Montoya, R. I., Cera, C. M. & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*, 787–814.

Gollan, T. H. & Silverberg, N. B. (2001) Tip-of-the-tongue states in Hebrew–English bilinguals. *Bilingualism: Language and Cognition, 4*, 63–83.

Gorter, D. & Cenoz J. (2012). Regional minorities, education and language revitalization. In A. Blackledge, A. Creese, M. Martin-Jones. (Eds.) *The Routledge handbook of multilingualism* (pp. 201-215). London: Routledge.

Haugen, E. (1953). The Norwegian language in America: A study in bilingual behavior. Philadelphia: University of Pennsylvania Press.

Heister, G. (1982). Sex differences in verbal fluency: A short note. *Current Psychology*, *2*(4), 257-260.

Henry, J.D., & Crawford, J.R. (2004). Verbal fluency deficits in Parkinson's disease: A meta-analysis. *Journal of the International Neuropsychological Society*, *10* (4), 608-22.

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica, 127*, 277–288.

Kavé, G. (2006). The development of naming and word fluency: evidence from Hebrew-speaking children between ages 8 and 17. *Developmental Neuropsychology*, *29*(3), 493-508.

Kirk, N. W., Fiala, L., Scott-Brown, K. C., & Kempe, V. (2014). No evidence for reduced Simon cost in elderly bilinguals and bidialectals. *Journal of Cognitive Psychology*, *26*(6), 640–648.

Kormi-Nouri, R., Moradi, A.R., Moradi, S., Akbari-Zardkhaneh, S., & Zahedian, H. (2012). The effect of bilingualism on letter and category fluency tasks in primary school children: Advantage or disadvantage? *Bilingualism: Language and Cognition*, *15*, 351-364

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of memory and language*, *33*(2), 149.

Lauchlan, F., Parisi, M., & Fadda, R. (2013). Bilingualism in Sardinia and Scotland: Exploring the cognitive benefits of speaking a 'minority' language. *International Journal of Bilingualism*, *17*(1), 43-56.

Lauder, A. (2008). The status and function of English in Indonesia: A review of key factors. *Makara Sosial Humaniora*, *12*(1), 9-20.

Lezak, M. D., Howieson, D. B., Loring, D. W., & Fischer, J. S. (2004). *Neuropsychological assessment*. New York: Oxford University Press.

Loonstra, A. S., Tarlow, A. R., & Sellers, A. H. (2001). COWAT metanorms across age, education, and gender. *Applied neuropsychology*, *8*(3), 161-166.

Luo, L., Luk. G., & Bialystok, E. (2010). Effects of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition*, *114*, 29-41.

Melinger, A. (2018). Distinguishing languages from dialects: A litmus test using the picture-word interference task. *Cognition*, *172*, 73-83.

Michael, E. B., & Gollan, T. H. (2005). Being and becoming bilingual: Individual differences and consequences for language production. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 389–407). New York: Oxford University Press.

Moreno, E. M. & Kutas, M. (2005). Processing semantic anomalies in two languages: An electrophysiological exploration in both languages of Spanish–English bilinguals. *Cognitive Brain Research*, *22*, 205–220.

Portocarrero, J. S., Burright, R. G. & Donovick, P. J. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology, 22*, 415–422.

Romaine, S. (1995). *Bilingualism* (2nd ed.). Cambridge, MA: Blackwell.

Sandoval, T. C., Gollan, T. H., Ferreira, V. S., & Salmon, D. P. (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition*, *13*(2), 231–252.

Sauzéon, H., Lestage, P., Raboutet, C., N'Kaoua, B., & Claverie, B. (2004). Verbal fluency output in children aged 7–16 as a function of the production criterion: Qualitative analysis of clustering, switching processes, and semantic network exploitation. *Brain and Language*, *89*(1), 192-202.

Scheuringer, A., Wittig, R., & Pletzer, B. (2017). Sex differences in verbal fluency: the role of strategies and instructions. *Cognitive Processing*, *18*(4), 407-417.

Schwartz, S., & Baldo, J. (2001). Distinct patterns of word retrieval in right and left frontal lobe patients: a multidimensional perspective. *Neuropsychologia, 39*, 1209–1217.

Strauss, E., Sherman, E.M.S., & Spreen, O. (2006). *A compendium of neuropsychological tests: administration, norms, and commentary* (3rd. Ed.). New York: Oxford University Press.

Shao, Z., Janse, E., Viseer, K., & Meyer, A.S. (2014) What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in Psychology*, *5*(3), 772.

Tombaugh, T.N., Kozak, J., & Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Archives of clinical neuropsychology: the official journal of the National Academy of Neuropsychologists*, *14*(2), 167-177.

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: evidence from younger and older healthy adults. *Neuropsychology*, *11*(1), 138.

Troyer, A. K., Moscovitch, M., Winocur, G., Leach, L., & Freedman, M. (1998). Clustering and switching on verbal fluency tests in Alzheimer's and Parkinson's disease. *Journal of the International Neuropsychological Society*, *4*(2), 137-143.

US Census Bureau (2003, February 25). *Table 1. Language use, English ability and linguistic isolation for the population 5 years and over by state: 2000.* United States Census 2000. Retrieved July 16, 2020 from www.census.gov/population/cen2000/phc-t20/tab01.pdf

Van den Berg, René. (2014). Juara satu dan dua: Membandingkan situasi kebahasaan Indonesia dan Papua Nugini. *Linguistik Indonesia, 32*(2), 103-129.

Weiss, E. M., Siedentopf, C., Hofer, A., Deisenhammer, E. A., Hoptman, M. J., Kremser, C., & Delazer, M. (2003). Brain activation pattern during a verbal fluency test in healthy male and female volunteers: a functional magnetic resonance imaging study. *Neuroscience Letters, 352*(3), 191-194.

Wysokinski, A., Zboralski, K., Orzechowska, A., Galecki, P., Florkowski, A., & Talarowska, M. (2010). Normalization of the verbal fluency test on the basis of results for healthy subjects, patients with schizophrenia, patients with organic lesions of the chronic nervous system and patients with type 1 and 2 diabetes. *Archives of Medical Science*, *6*(3), 438-446.

Zhao, Q., Guo, Q., & Hong, Z. (2013). Clustering and switching during a semantic verbal fluency contribute to differential diagnosis of cognitive impairment. *Neuroscience Bulletin, 29*, 75–82.

**Biographical notes**

**FEISAL AZIEZ** is a PhD student at Multilingualism Doctoral School, University of Pannonia, Hungary. He is also a lecturer at English Education Department, Universitas Muhammadiyah Purwokerto, Indonesia; feisalaziez@ump.ac.id

**HAZIM ALKHRISHEH** is a PhD student at Multilingualism Doctoral School, University of Pannonia, Hungary; hkhrisheh@yahoo.com

**FURQANUL AZIEZ** is an associate professor at Faculty of Teacher Training and Education, Universitas Muhammadiyah Purwokerto, Indonesia; f.aziez2010@gmail.com

**MAULANA MUALIM** is a lecturer at English Education Department, IAIN Purwokerto, Indonesia; maulanamualim@iainpurwokerto.ac.id