Gender and the Letter Fluency Task: Evidence from Second Language Learners*

Justin M. Aronoff[1]
University of Southern California[2]
aronoff@usc.edu

The letter fluency task is a common test of executive control, but the factors that effect it are poorly understood. One of the factors that has generated a considerable amount of research is the gender of the participant. A number of studies have found that females produce more words in this task than males, but this is conflated with differences in vocabulary size and exposure. To examine this issue, two studies were conducted using the letter fluency task with classroom second language learners, an ideal group because of their uniform exposure to the test language. The results indicate that, even when exposure and vocabulary size are controlled, females still produce more words than males in the letter fluency task, thus indicating that this phenomenon cannot be reduced to an exposure effect of exposure or vocabulary size.

1. Introduction

The letter fluency task is a common behavioral/neuropsychological test in which patients are asked to name as many words as possible that start with a specific letter. This test is one of many tests of executive control often used to evaluate individuals with suspected prefrontal lobe damage or impairment. Individuals with prefrontal damage generally make more errors and produce less (non-repeated) words than healthy normal controls.

This task is difficult because it requires patients to ignore the more obvious (and more commonly used) semantic relationships, which rarely have the same first letter[3] as well as inhibit previously used responses. There have been a number of studies that have investigated the effect of a number of patient characteristics, such as age, education level, and gender on the letter fluency task (*e.g.*, Capitani *et al.*., 1998; Phillips, 1999; Tomer and Levin, 1993). Important for the current study are findings, such as those by Capitani *et al.*. (1998), which have demonstrated gender differences in the letter fluency task, with females producing more words than males (though not all studies have found the effect [*e.g.*, Auriacombe *et al.*, 2001]). Because the gender effects may be conflated with or hidden by a number of other factors such as productive vocabulary size or differences in exposures to categories of words, an important question is whether a gender effect would still be found when these are controlled.

Possible suggestions for the basis of the gender effects come from the semantic fluency task. In this task, participants are given a category such as *fruits* and asked to generate as many words belonging to that category as possible. Auriacombe *et al.* (2001) conducted a large scale study of participants 65 years and older and found that, after factoring out age, education, and occupation,

* The author would like to thank Parkland College and Elk Grove High School for each graciously permitting the use of eight of its Spanish classes and for the continual effort made by the schools, and specifically the Spanish departments, to further their understanding of the subjects they teach and the way that they are taught. In addition, thanks are due to Gail Gottfried, Elaine Andersen, and two anonymous reviewers for their helpful comments.
[1] This research was supported by NIH training grant 5T32MH20003-05.
[2] Data for this paper was collected while the author was at the University of Illinois at Urbana Champaign, part of which was used in a Senior Thesis.
[3] This is not to say that there is no correspondence between the first letter(s) of the word and the semantics of the word (e.g. many words that start with *sn* refer to less desirable words such as *sneer, snide, snore, etc.*), but this tends to be the exception, rather than the rule in English and Spanish.

females produced more words than males. In contrast, Capitani *et al.* (1999) demonstrated that gender effects exist for semantic retrieval tasks, but there is a female advantage for one category (fruits) and a male advantage for another (tools), indicating that the effect is not caused by the abilities of one gender over another but rather by differential exposure in one semantic area or another. Additional evidence for a category-specific gender effect comes from a number of studies comparing performance of young healthy males and females naming living and nonliving objects. For instance, Laws (1999) found that females respond faster to living than non-living objects and males the reverse.

Because category-specific effects of this sort may explain the gender differences for semantic categories, they may do the same for the letter fluency task. Although there have not (to the author's knowledge) been letter fluency studies that have tried to determine if the words starting with a given letter come from a specific category disproportionately, it is possible that the different category effects could be driving the gender-based effect for the letter fluency task. This could be true either based on the letters chosen or on an overall living-bias within vocabularies.

In terms of more general gender effects, a number of researchers have suggested that gender-based differences may exist across a number of diverse cognitive tasks; for instance, a number of studies have demonstrated a female advantage for a variety of linguistically related tasks. Such experiments have shown that females perform better than males on tasks requiring phonological retrieval (Hines, 1990), synonym generation (Halpern and Wright, 1996), and speech articulation (Kimura and Hampson, 1994). Differences also surface for skills that are not language related, such as mood, manual dexterity, spatial ability, mathematical prowess, and perceptual speed (Hedges and Nowell, 1995; Montello *et al.,* 1999).

The above gender-related cognitive differences not withstanding, studies that have found gender-based differences on specific semantic categories (*e.g.*, Capitani *et al.*, 1999) indicate that experience-based factors may be driving the gender effects in the letter fluency task. The logic behind this argument is that, for example, males may have more exposure to a category such as *tools* than females. If the words that start with a given letter tend to be predominantly from categories that members of one gender generally have more exposure to, then it is possible that there will be a gender effect, although it is not related to more general cognitive differences.

Classroom-based second language acquisition is an ideal environment in which to determine whether the gender effect on the letter fluency task is caused solely by difference in exposure or if it is driven by a more general difference between males and females. The reason for this is that exposure to vocabulary is largely controlled in the classroom setting with all students being exposed to the same small set of common words. This means that any differences that may be found would not be caused by a difference in vocabulary exposure.

2. Study 1

2.1 Participants

Participants were 68 students at a midwestern community college. All students were enrolled in Spanish classes. Two classes participated from each of the four beginning levels of Spanish classes. Based on information given by the participants, data were excluded for any subjects that did not fulfill the following criteria: right handed according to a modified version of the Eidenburgh Handedness Inventory[4] (Oldfield, 1971), native speaker of English, English spoken at home, and Spanish first learned after age 8. The age of acquisition criteria was used to create a homogenous group because, as data from Chee *et al.* (1999), Dehaene *et al.* (1997), and Kim *et al.* (1997) indicate, early bilinguals (*i.e.*, individuals who acquired a second language before around age eight) have a biologically different representation of their second language than those who start acquiring a second language later in life. In addition, second language abilities of late learners are quite different than that of early

[4] This survey is a comprehensive measure for determining handedness. Participants are asked to list the hand they use for a variety of tasks such as using a toothbrush and striking a match. Based on their responses, they receive a score ranging from –20 (completely left handed) to +20 (completely right handed). For the purposes of the current study, anyone receiving a score above +5 was considered right handed.

learners (*e.g.*, Webber-Fox and Nevile, 1996; Johnson and Newport, 1989; Birdsong and Molis, 2001). In addition, limiting the participants to native English speakers who were English dominant reduced the chances that certain participants would have an advantage in this task through the use of cognates from different languages. In addition, only right handed individuals were used to create a more homogenous group in terms of the likely lateralization of the neurological representation of language. As a result of these criteria, data from sixty-four participants was used (37 females and 27 males). Table 1 shows the age and class level of the participants.

Level	Females	Males	Combined
1	26.5 (12)	24.4 (9)	25.5 (10.4)
2	20 (2.1)	23.4 (5.4)	21.2 (3.9)
3	19.6 (1.9)	23.3 (5.8)	21.3 (4.4)
4	21.1 (2.3)	22.9 (5.9)	21.9 (4.2)
Total	23.5 (6.3)	21.6 (6.2)	22.4 (6.3)

Table 1: Mean age (SD) by level and gender.

2.2. Procedure

Participants were first given a written letter fluency task. After completing this task they were asked to fill out a personal background sheet that asked for information such as their gender, age, the number of years they had been studying Spanish, native language, and handedness. The stimulus for the letter fluency task was a selected letter of the alphabet. The task was to generate (in written form) Spanish words that started with the selected letter. This task was repeated twice, first with the letter *p* and then with the letter *c*. These letters were chosen because they are the two most commonly occurring word-initial letters in Spanish based on a selection of the most common words in the *Diccionario de frecuencias de las unidades lingüísticas del castellano* (Alameda, 1995), a Spanish frequency count dictionary. Participants were instructed not to use proper nouns or multiple conjugations of verbs or nouns (he walks, they walked, car, cars, etc.). They were given two minutes to complete this task for each letter. After the time had elapsed they were instructed to finish the word they were currently writing and stop. Participants were not permitted to return to this task after the time had elapsed.

2.3. Coding and Data Analysis

All words generated in the first task were reviewed and verified against a Spanish monolingual dictionary. A number of responses, though composed of actual words, were not considered valid. Among those were proper nouns and phrases. All morphological variants of a word were counted as one word, whereas each lexical variant was counted as a separate word (i.e. *plane* and *planes* counted as one word but *plane* and *plain* counted as two). The dependent variable was the total number of (acceptable) generated words.

Four different instructional levels were included to obtain a representative sample of language learners. However, because the effect of instructional level upon performance was not the main question of interest, the levels were combined for the purpose of subsequent statistical analyses. See Figures 1 and 2 for the mean score of males and females at each level for the letters *p* and *c*, respectively.

2.4. Results and Discussion

Because Student's t-test has been noted to perform poorly under conditions of non-normality, whereas a modified version of Student's t-test, Welch's test, has been demonstrated to function well under a wide range of normal and non-normal conditions, the first analysis used a Welch's test to compare the number of words generated by males and females for the letter *p*. This resulted in a significant female advantage, $t\ (61) = 2.84$, $p = .006$, such that females produced significantly more valid words than males. To verify the female advantage effect, the same analysis was carried out using the words generated with the letter *c*. Again, a significant female advantage was found, $t(59)=3.03$, $p = .004$.

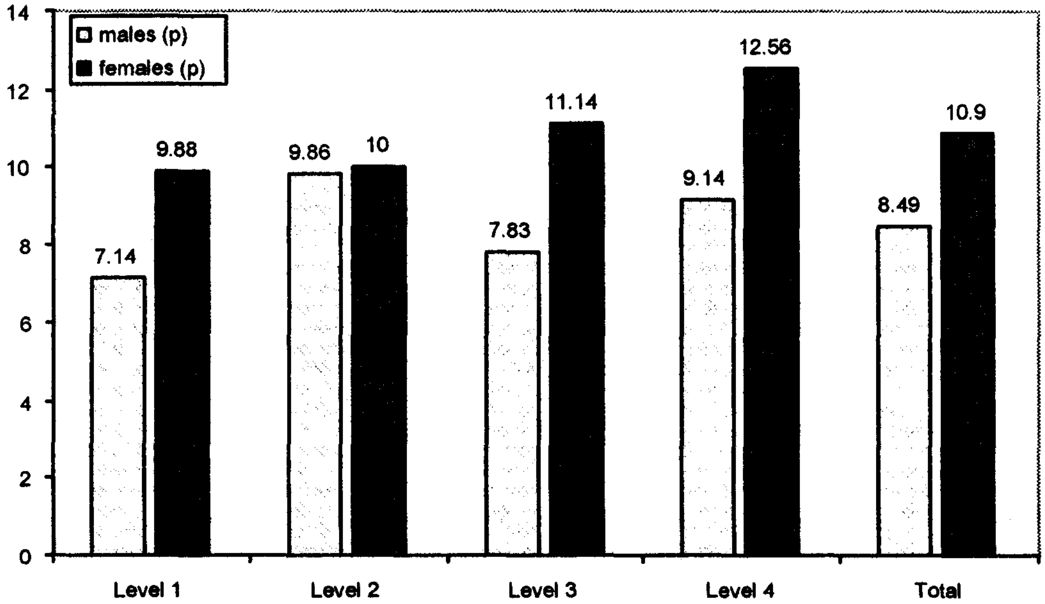

Figure 1. Average number of valid words generated beginning with the letter p.

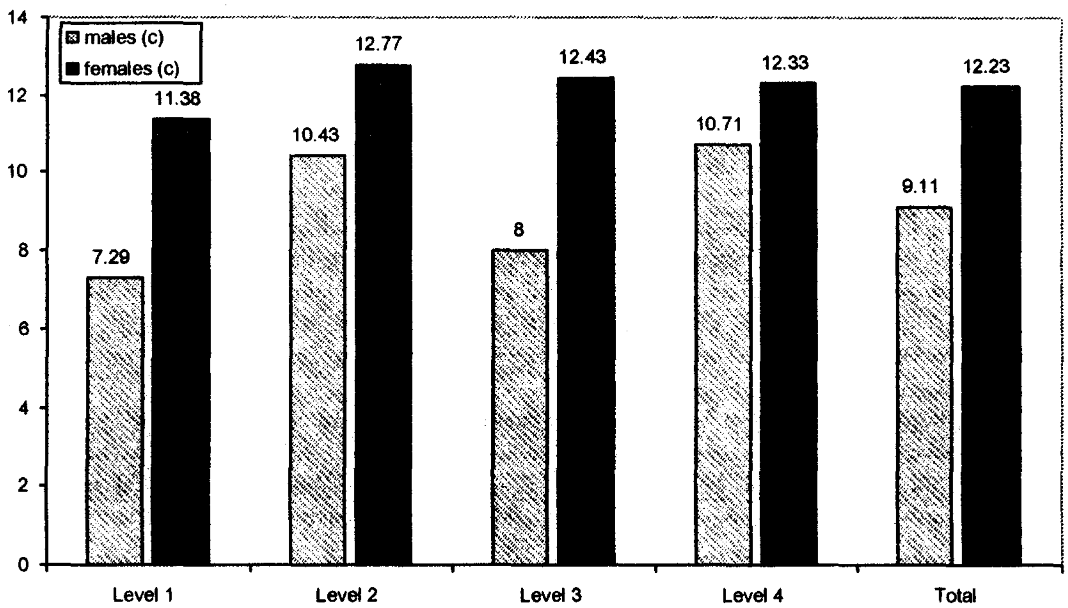

Figure 2: Average number of valid words generated beginning with the letter c.

In this study females produced significantly more words than males both for the letter *p* and for the letter *c*. These results are similar to those found when asking subjects to generate words in their native language (*e.g.*, Capitani *et al.*, 1998). Because the words generally learned in these classes are high frequency words, it is unlikely that the effect found is caused by difficulty in accessing words in the participants' first language. In addition, all students were exposed to the same Spanish words, and the amount of exposure was similar for all students. Thus it is unlikely that the effect found was driven by exposure. Instead, the results provide evidence that the gender effect on the letter fluency task is not caused solely by vocabulary knowledge but seems to, at least partially, be influenced by more general factors such as how inappropriate words are inhibited or otherwise avoided[5]. This interpretation is supported by findings that individuals commit errors of intrusion and repetition on this task, with repetitions being more common than intrusions (*e.g.*, Auriacombe *et al.*, 2001)[6]. These two types of errors are both related to functions commonly associated with aspects of executive functioning. Among the many functions attributed to executive functioning, the two relevant here are inhibition and working memory. Intrusions can be seen as a failure to inhibit associated words that are irrelevant for the task at hand. Repetition, on the other hand, although possibly related to difficulty in inhibition, is most likely related to limits in working memory. The reason for this is that, in order to avoid repetition, one needs not only to search for appropriate words, inhibiting related words that start with different letters, but also have to maintain in working memory the words that fit that criteria which have already been generated.

In addition to the above evidence, which suggests that the task relies largely on executive functions (as indicated by the errors typically made), there is also evidence that executive functions do show gender effects. A number of studies have found that females perform better on tasks that test executive functions (*e.g.*, Kelly and Britton, 1996; Warrick and Naglieri, 1993). These two types of converging evidence support the interpretation of the results of the first study.

There are two concerns with the above data, which merited a second study. First, it is possible in the above task that, although the Spanish words the participants were exposed to were the same, the words that the participants actually learned were not. In this sense, the results may simply indicate that females are better at learning (retaining, acessing, etc.) words they are exposed to than males. The second concern was that there was a large range and variance in age of the participants (SD 6.3, range: 16-53), which may have independent effects on the results (e.g. Capitani *et al.*, 1998). Although in some levels males are older than females and in others females are older than males, in the two most advanced levels, males are older than females. Since the later levels produced the most words and thus had the biggest effect on the average score, it would be consistent with the findings of Capitani *et al.* (1998) for the males to have produce less words simply because they are older[7].

3. Study 2

The following study aims at normalizing the letter fluency score according to a measure of the vocabulary size of each participant. Although many studies carefully control for cognates to assure that participants are recognizing words through their second language (L2) instead of through their first language (L1), that was purposely not done in this study. The reason for this is that the goal of the recognition task was to normalize the letter fluency score. Because cognates were frequently produced in the letter fluency task, excluding these words could have a biased effect on the normalization[8]. In addition, because L1 was controlled across participants, all participants were presumably at an equal advantage for cognates. In addition, to control for age, the participants were selected from high school classes so as to greatly limit the range of the ages of the participants.

3.1. Participants

Participants were selected from an original group of 168 students attending a high school in the Midwest. Two classes were selected at each level (Spanish 1, 2, 3, and 4). Subjects gave informed consent to participate in the study. As in the first experiment, based on information given by the

[5] This is not to say that vocabulary size and exposure have no effect, just that there are other independent factors that contribute to the gender effect.

[6] Although errors are quite rare in the current task, this should not be surprising. Errors tend to occur in oral versions of this task where participants are not able to go back and verify or modify their answers.

[7] Although the relatively young age of the participants would most-likely rule out this possibility, it is desirable to verify that this is not the cause of the effect.

[8] Eliminating cognates would be especially problematic because there is no way to guarantee that the males and females would produce equal percentages of cognates in the letter fluency task.

subjects, data were excluded for any subjects who did not fulfill the following criteria: right handed, native speaker of English, English spoken at home, and Spanish first learned after age eight. As a result, data from 81 participants were used (47 females, 34 males). The ages of the participants ranged from 14 to 17 (SD = .92, mean = 15.28).

3.2. Procedure

The experiment was performed in classrooms at the high school. Participants performed the letter fluency task followed by a word recognition task. Participants were given two minutes to complete each task. As in Study 1, participants also filled out a personal background questionnaire. The stimulus for the first task was the same as in Study 1 except only the letter *p* was used. This task was completed in an identical manner as in Study 1.

The word recognition task was designed to normalize the letter fluency score by adjusting for the number of words that the participants either knew or recognized in order to account for individual variation in vocabulary size. For this task, stimuli were taken from the list of Spanish words that started with the letter *p* generated by participants in Study 1. The generated words were divided into the following syntactic categories: adjective, adverb, noun, preposition, and verb. Each word was ranked by the number of times it occurred across subjects. Twenty percent of the words from each syntactic category—half of which were ranked (based on the number of times each word was produced across subjects in Study 1) as most frequent and half of which were ranked as least frequent—were selected. In the case that more than one word shared the boundary ranking, the desired number of stimuli were randomly selected from the equal-ranking words. Participants were instructed to translate or give a definition for each word on the list. For those items that they were not able to give the meaning for, they were asked to indicate whether or not they recognized the word.[9]

3.3. Coding and Data Analysis

Data analysis for the word generation task was the same as in Study 1. For the recognition task, all translations and definitions were verified against a Spanish/English and monolingual Spanish dictionary for general accuracy. The translation or definition was considered correct if it approximated the correct translation / definition, although not necessarily with the same degree of detail. Thus, both "money" and "Spanish money" were considered valid for the word *peso*, which has one definition meaning "Mexican money." However, for a word to be correct, it was required that the translation or definition maintain the same syntactic class as the original word. Therefore, "personally" would not be a valid translation for *personal* (correctly translated as "personal"). In addition, all incorrect responses in the Meaning column were analyzed as members of the Recognize column because individuals who incorrectly thought they knew the definition of the word did, at a minimum, recognize the word. Each subject received three scores. The Recall score was the number of valid words generated in the letter fluency task, and the Translation score was the number of valid translations or definitions in the word recognition task. The Recognize score was the number of valid translations, definitions, or indications of recognition in the word recognition task. To determine if the female advantage was driven by vocabulary size, two adjusted scores were computed. The Translation-adjusted score consisted of the Recall score divided by the Translation score. The Recognition-adjusted score was computed by dividing the Recall score by the Recognize score. See Figures 3 and 4 for the mean Recall and mean Recognition-adjusted scores, respectively.

[9] Participants were presented with three options: recognize, know, and translate/define. Before this task was administered each column was defined and an example was given describing when to use each. Because of the very low use of the "know" column, the "know" and "recognize" columns were collapsed in all analyses.

3.4. Results and Discussion

A Welch's test were conducted on the raw letter fluency data. The analysis revealed that, as in Study 1, females generated significantly more words than males, t (78) = 2.23, p <.03. A Welch's test was conducted on an adjusted score based on the Recall/Translation scores. No significant difference was found. An additional Welch's test was conducted on an adjusted score based on the Recall/Recognition scores. The results indicated a marginally significant female advantage, t (79) = .78, p =.078.

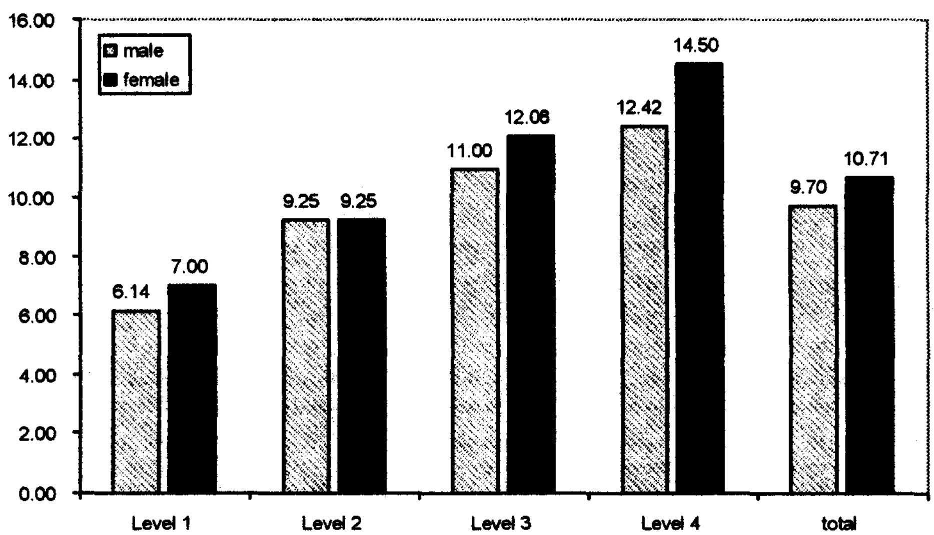

Figure 3: The average number of words generated during the letter fluency task, by class and gender.

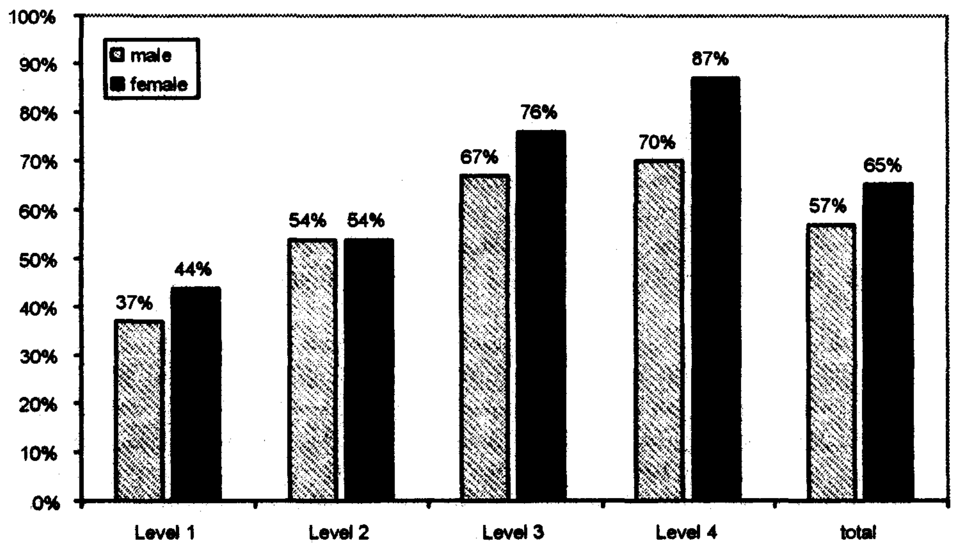

Figure 4: The Recognition-adjusted results of the letter fluency task, representing the number of words generated, divided by the total number of words recognized.

The results from the first experiment were replicated in this larger study, with the females generating more words than the males. Because the age range was much more restricted than in Study 1, it is unlikely that the results are driven by age effect. Importantly, the results show that, even when recognition vocabulary size is used to normalize the score on the letter fluency task, there is still a marginally significant female advantage. This suggests that the gender effect in the letter fluency task is not solely caused by exposure or the resulting vocabulary size.

It is not entirely surprising that adjusting the generation score by the Translation score results in a non-significant result because more efficient searching of the lexicon may lead to better translation ability in a task where time is limited. In other words, the mechanisms that allow an individual to translate a word are likely similar to those that are required for the difficult task of generating words. Recognition of a word, on the other hand, is similar to priming and as such it is a very superficial lexicon searching problem that does not require accessing the meaning of a word. Because of this, recognition is a much more accurate measure of exposure than their ability to translate a word.

4. General Discussion

Study 1 and 2 both confirm a female advantage for the letter fluency task, even when exposure is controlled and the data are adjusted to account for vocabulary size. Since this task has a strong orthographic component, a possible concern is that the difference found may be the result of a difference in males or females in terms of orthographic skills. A number of studies have investigated this possibility (e.g. Hayes and Waller, 1994; Wallschlaeger and Hendricks, 1997). However, because Spanish has such a strong letter-sound correspondence (*e.g.*, Arries, 1994), this would most likely have only a minimal effect.

Another possible concern is that there may be a difference in males and females in their exposure to the underlying concepts through their L1. As mentioned previously, some researchers (e.g. Capitani *et al.*, 1999) have put forth the notion that, in the category fluency task, the number of words that are produced in a category is related to the exposure to that category. Because most of the words that the participants have been exposed to in their second language classes are primarily common words, both sexes have likely been amply exposed to all relevant words.

The results found beg the question of what more general differences between the genders drive this effect. To answer this, a closer look at the exact nature of the task is needed. Like most verbal tasks, the letter fluency task requires efficient searching of the lexicon; however, unlike other tasks, this type of searching is not used in normal speech. A number of theoretical approaches have suggested that normal accessing of a given word automatically activates a set of related words, in what is called spreading activation (e.g. McClelland and Rumelhart, 1981). Because semantically related words do not generally start with the same letter, these words must be inhibited. This notion is confirmed by intrusion errors, which have been found for verbal forms of the letter fluency task (*e.g.*, Auriacombe *et al.*, 2001). In addition, this task requires the inhibition of words already produced. Again, error data confirms this (*e.g.* Auriacombe *et al.*, 2001). This would indicate that, as suggested by Capitani *et al.* (1999), the differences in performance that cannot be explained through vocabulary size or exposure may be caused by a difference in executive functioning, particularly the ability to inhibit irrelevant responses.

Evidence for this proposal comes from studies comparing males and females using neuropsychological tests considered to measure executive function[10], such as the Stroop task (Franzon and Hugdahl, 1986; Golden, 1974; Mekarski, Cutmore, and Suboski, 1996). In this task, participants are shown a series of color names written in an ink color that is either congruent or incongruent with the word shown (*e.g.*, the word *red* written in blue ink) and are asked to name the color of the ink. Numerous studies (*e.g.*, Dunbar and MacLeod, 1984; Stroop, 1935) have found that participants are significantly slower for incongruent versus congruent conditions, indicating a difficulty to inhibit the conflicting information from the written word. Mekarsi *et al.* (1996) found a significant difference between reaction times for males and females, with females being consistently quicker than males. There was no evidence of a difference in accuracy between the two groups. This result echoes similar findings, such as those by Sarmany (1977), and Franzon and Hugdahl (1986). Thus, differences in executive functioning are a plausible explanation of the results found.

[10] Research suggests that executive function, controlled by the frontal lobe, plays a significant role in the inhibition of inappropriate responses (*cf.*, Goldberg, 2001, and references therein)

Both studies conducted here found that females produce significantly more words than males on a written version of the letter fluency task, even when accounting for overall vocabulary size. The studies cited above, as well as those by Kelly and Britton (1996) and Warrick and Naglieri (1993) indicate that the difference in performance between the two genders could plausibly be caused by differences in executive functioning.

References

Alameda, J. R. (1995). *Diccionario de frecuencias de las unidades lingüisticas del castellano*. Oviedo, Spain: Departamento de Psicologia, Universidad de Oviedo.

Arries, J. F. (1994). An Experimental Spanish Course for Learning Disabled Students. *Hispania, 77*, 110-117.

Auriacombe, S., Fabrigoule, C., Lafont, S., Amieva, H. (2001). Letter and category fluency in normal elderly participants: a populaion-based study. *Aging Neuropsychology and Cognition*, 8(2), 98-108.

Berninger, V. W. and Fuller, F. (1992). Gender differences in orthographic, verbal, and compositional fluency: Implications for assessing writing disabilities in primary grade children. *Journal of School Psychology*, 30(4), 363-382.

Birdsong, D. and Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language*, 44, 235-249.

Capitani, E., Laiacona, M., and, Basso, A. (1998). Phonetically cued word-fluency, gender differences and aging: A reappraisal. *Cortex*. 34, 779-783.

Capitani, E., Laiacona, M., and Barbarotto, R. (1999). Gender affects word retrieval of certain categories in semantic fluency tasks. *Cortex*, 35, 273-278.

Chee, M. W. L., Caplan, D,, Soon, C. S., Sriram, N., Tan, E. W. L., Thiel, T., and Weekes, B. (1999). Processing of Visually Presented Sentences in Mandarin and English Studied with fMRI. *Neuron*, 23 (1), 127-137.

Dehaene, S.; Dupoux, E.; Mehler, J.; Cohen, L.; Paulesu, E.; Perani, D.; van de Moortele, P.-F., Lehéricy, S. and Bihan, D. L. (1997). Anatomical variability in the cortical representation of first and second language. *NeuroReport*, 8, 3809-3815.

Dunbar, K. and MacLeod, C. M. (1984). A horse race of a different color: Stroop interference patterns with transformed words. *Journal of Experimental Psychology: General*, 113, 501-517.

Franzon, M., and Hugdahl, K. (1986). Visual half-field presentations of incongruent color words: effects of gender and handedness. *Cortex*, 22, 433-445.

Goldberg, E. (2001). *The Executive Brain: Frontal Lobes and the Civilized Mind*. New York : Oxford University Press.

Golden, C. J. (1974). Sex differences in performance on the Stroop Color and Word Test. *Perceptual and Motor Skills*, 39, 1067-1070.

Hayes, Z. L. and Waller, G. T. (1994). Gender differences in adult readers: A process perspective. *Canadian Journal of Behavioruał Science*, 26 (3), 421-437.

Hedges, L. V. and Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, 269, 41-45.

Hines, M. (1990). Gonadal hormones and human cognitive development. In J. Balthazart (Ed.), Brain and behavior in vertebrates 1: Sexual differentiation, neuroanatomical aspects, neurotransmitters and neuropeptides (pp. 51-63), Basel, Switzerland: Karger.

Halpern, D. F., and Wright, T.M. (1996). A process-oriented model of cognitive sex differences [Special issue]. *Learning and Individual Differences*, 8, 3-24.

Johnson, J. S. and Newport, E. L. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21, 60-99.

Kelly, T. P. and Britton, P. G. (1996). Sex differences on an adaptation of the digit symbol subtest of the Wechsler Intelligence Scale for Children-III. *Perceptual and Motor Skills*, 83, 843-847.

Kim, K. H. S.; Relkin, N. R.; Lee, K.-M.; Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature*, 388, 171-174.

Kimura, D. and Hampson, E. (1994). Cognitive pattern in men and women is influenced by fluctuations in sex hormones. *Current Directions in Psychological Science*, 3, 57-61.

Laws, K. R. (1999). Gender affects naming latencies for living and nonliving things: implications for familiarity. *Cortex*, 35, 729-733.

Mekarski, J. E., Cutmore, T. R. H., and Suboski, W. (1996). Gender differences during processing of the stroop task. *Perceptual and Motor Skills*, 83, 563-568.

McClelland, J. and Rumelhart, D. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, 88, 375-407.

Montello, D. R., Lovelace, K.L., Golledge, R. G. and Self, C.M. (1999). Sex related differences and similarities in geographic and environmental spatial abilities. *Annals of the Association of American Geographers*, 515-533.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9, 97-113.

Phillips, L. H. (1999). Age and Individual Differences in Letter Fluency. *Developmental Neuropsychology*, 15(2), 249-267.

Sarmany, I. (1977). Different performance in Stroop's interference test from the aspect of personality and sex. *Studia Psychologica*, 19, 60-67.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643-662.

Tomer, R. and Levin, B. E. (1993). Differential effects of aging on two verbal fluency tasks. *Perceptual and Motor Skills*, 76, 465-466.

Wallschlaeger, M. and Hendricks, B. (1997). Gender differences in phonetic processing. *Current Psychology: Developmental, Learning, Personality, Social*, 16(2), 155-166.

Warrick, P. D. and Naglieri, J. A. (1993). Gender differences in planning, attention, simultaneous, and successive (PASS) cognitive processes. *Journal of Educational Psychology*, 85(4), 693-701.

Weber-Fox, C. M. and Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, 8:3, 231-256.