Acta Psychologica 260 (2025) 105735

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

# Literacy boosts semantic fluency even in ecologically valid categories

Tan Arda Gedik [a,b,*], Vania De la Garza [a]

[a] FAU Erlangen-Nürnberg, Germany
[b] Bilkent University, Turkey

ARTICLE INFO

*Keywords:*
Semantic Fluency
Literacy
Reading Fluency
Turkish

ABSTRACT

Semantic fluency tasks assess the ability to generate words within a category under time constraints and are shaped by linguistic, cognitive, and experiential factors such as literacy. This study examined how reading fluency (a proxy for literacy) and education are associated with performance in adult Turkish speakers across three semantic categories of varying ecological validity: animals, fruits, and household objects. Forty-four participants (24 illiterate, 20 literate; aged 21–80) completed the task. Literate participants produced significantly more words than illiterate participants in all categories ($p < 0.001$), including those reflecting everyday experience (i.e., fruits, and household objects). Regression analyses revealed that reading fluency was the only significant predictor (final model: reading composite $\beta = 0.765$, $p < 0.0001$; $R^2 = 0.58$). These results suggest that literacy may broaden vocabulary and world knowledge and enhance retrieval efficiency, highlighting its role in semantic fluency even for ecologically familiar categories.

## 1. Introduction

Semantic fluency is more than a measure of verbal speed: it provides a window into how the brain stores and accesses meaning. Successful performance requires semantic memory, executive control, and lexical retrieval, making it widely used in neuropsychological assessment to detect cognitive impairments, track aging-related decline, and examine language processing (Strauss et al., 2006). The most common task asks participants to name as many animals as possible within 60 s (e.g., Brucki & Rocha, 2004; Mitrushina et al., 2005), but research has also tested fruits, objects, and other semantic categories across diverse populations, including individuals with varying literacy levels (Yasa Kostas et al., 2024). However, few studies have systematically compared multiple semantic categories, (see Table 1 below) making it unclear whether illiterate adults perform worse across all categories or only on those that are experientially less familiar.

UNESCO (2024) estimates that 754 million adults worldwide cannot read or write. Illiteracy is often defined as the inability to read and understand simple texts, whereas functional illiteracy refers to difficulty using literacy skills in daily life (UNESCO, 1978; Vágvölgyi et al., 2016). In this study, “illiterate” refers to adults with minimal formal schooling (≤1 year) and very limited decoding ability, though some had informal exposure to reading. Our sample thus reflects a continuum of literacy experiences, allowing us to examine literacy and education as related but distinct factors.

Several studies have compared semantic fluency in literate and illiterate adults, with results summarized in Tables 1. Across studies, literate adults generally outperform illiterate adults on animal fluency tasks, though the magnitude varies. Most studies focus on animals; fewer examine familiar daily-life categories such as fruits or supermarket items. As can be seen in Table 1, in familiar categories, the performance gap is smaller. For example, both Da Silva et al. (2004), and Nielsen and Waldemar (2016) found that illiterate speakers performed comparably to literate speakers on supermarket items (a more familiar category) but more poorly on animals. These findings suggest that performance may depend on the familiarity, or ecological validity, of the category.

Prior research also debates whether education or literacy drives these differences since illiterate speakers differ in education experience from literate speakers. Schooling may enhance exposure to written materials, broaden topic knowledge, and improve reasoning skills, all of which contribute to richer semantic networks and better fluency (Denervaud et al., 2021; Ritchie & Tucker-Drob, 2018; Sternberg, 1987). Conversely, literacy itself reshapes brain structure and function, particularly in language-related areas, and supports cognitive processes relevant to semantic fluency, such as working memory and lexical access (Dehaene et al., 2010; Huettig & Mishra, 2014). While disentangling the effects of literacy and education is challenging since they are correlated, regression models including both predictors can estimate their

* Corresponding author at: FAU Erlangen-Nürnberg, Germany.
*E-mail address:* tan.gedik@fau.de (T.A. Gedik).

https://doi.org/10.1016/j.actpsy.2025.105735
Received 17 January 2025; Received in revised form 4 October 2025; Accepted 6 October 2025
Available online 14 October 2025

**Table 1**
An overview of previous studies on semantic fluency from illiterate and literate speakers.

| Study – country | Semantic category | Illiterate Mean (SD) | Literate Mean (SD) |
|---|---|---|---|
| Brucki and Rocha (2004) – Brazil | Animals | 12.1 (3.0) | 13.6–17.2 (3.6–4.8) by education groups |
| Da Silva et al. (2004) – Brazil | Animals/ supermarket | 12.4 (4.2)/ 15.9 (3.0) | 13.6–17.2 (3.6–4.8) |
| Kosmidis et al. (2004) – Greece | Combined (animals, fruits, objects) | 30.6 (5.4) | 40.4–50.1 (6.8–9.0) |
| Manly et al. (1999) – USA | Animals | 11.8 (4.0) | 11.6 (3.1) |
| Mathuranath et al. (2003) – India | Animals | 6.0 (2.8) | 7.3–9.4 (2.8–3.9) |
| Nielsen and Waldemar (2016) –Denmark | Animals/ supermarket | 12.2 (3.9)/ 18.9 (5.5) | 16.9 (4.7)/19.9 (5.5) |
| Ostrosky-Solís et al. (1999) – Mexico | Animals | 12.7–15.2 (5.0–5.6) by age | 19.9–21.6 (5.4–5.8) |
| Rosselli et al. (1990) – Colombia | Animals | 9.7–11.3 | 17.8 |
| Youn et al. (2011) - Korea | Animals | 11.8 (3.1) | 14.1 (3.5) |

Note. Semantic category = Most studies used a 60-second animal fluency task unless otherwise stated. Kosmidis et al. report a combined score for all categories tested. Median values were reported in Nielsen and Waldemar instead of means.

independent contributions.

A critical gap in the literature concerns category familiarity in semantic fluency tasks with illiterate and literate speakers. Few studies systematically test multiple semantic categories, leaving it unclear whether illiterate adults underperform across all categories or only on less familiar ones. To address these gaps, we systematically compared semantic fluency across three categories: animals, fruits, and household objects in age-matched literate and illiterate Turkish adults. To our knowledge, this is the first study to systematically compare semantic fluency across multiple familiar categories in illiterate and literate adults, while simultaneously accounting for literacy as a continuous predictor. Thus, literacy was modeled as a continuous predictor alongside years of education. Our aim was to assess group differences across categories varying in ecological validity (from least to most valid: animals > fruits > household objects). The choice of a Turkish population was entirely based on availability.

## 2. Method

### 2.1. Participants

We collected data from 24 illiterate (23 females, mean age = 51.5, SD = 11.66) and 20 literate (17 females, mean age = 43.95, SD = 14.93) adult native Turkish speakers. The illiterate participants were based in Ankara, Turkey, and at the time of the study, had been attending the adult education center for literacy courses for an average of 9.52 months (SD = 7.34). Some participants were repeating the course for the second or third time. The curriculum included literacy education along with subjects like math, basic history, and Turkish, spread across 80 teaching units, with each unit lasting 40 min. Upon successful completion of these courses, participants receive a degree equivalent to a primary school diploma in Turkey. The literacy center screened speakers in conjunction with health experts from local hospitals for various speech and communication impairments such as dyslexia, as well as cognitive impairments. Speakers with specific speech or cognitive impairments are not admitted to the literacy course where our data was collected. Thus, speakers with such conditions were excluded from the study. This study was reviewed and approved by Bilkent University Ethics Committee with the approval number: 2022_12_21_01, dated Dec 22, 2022.

An important note here is that some illiterate participants in our study could read some words per minute (see Section 3 for results), possibly because they learned to read in the literacy course, or maybe their family members had taught them previously. These participants would be considered functionally illiterate. But, although they can read words, they cannot comprehend longer texts, hence they attend the literacy classes. In addition to this, a large number of our participants were completely illiterate (i.e., they could not read any words in one minute, see Table 2). Thus, for ease of reference, we refer to this population as illiterate.

### 2.2. Materials

#### 2.2.1. The 1-Minute Word Reading Task

We utilized the Turkish 1-Minute Word Reading Task, which was specifically developed by Gedik (forthcoming) to accommodate both literate and illiterate individuals, recognizing that traditional literacy assessments, such as author recognition tests, are unsuitable for those without literacy skills. Similar tasks have been designed for other languages, including English and Greek (Deeney, 2010; Simos et al., 2013; Torgeson et al., 1999). Previous studies have shown a modest but statistically significant correlation between performance on these tasks and adult vocabulary size (Simos et al., 2013), as well as a moderate correlation between reading speed and text comprehension in children (Yıldız et al., 2014).

The words for each section (real and pseudo) were drawn from http://www.lexique.org/shiny/unipseudo/, which provides a ready-made collection of pseudowords tailored for each language. The task consists of 120 real words and 120 Turkish sounding, Turkish looking, pronouncable pseudowords (i.e., fake words), each presented in order of increasing length, with four letter-length categories for both real and pseudowords (real words: 5, 6, 7, 8 letters; pseudowords: 6, 7, 8, 9 letters). There were several reasons for the difference in length of letters between the two subsections. First, real words benefit from participants' existing lexical knowledge, allowing for faster and more automatic recognition. Even short real words can often be identified with minimal phonological decoding due to their frequency and familiarity. In contrast, pseudowords require full phonological decoding, as they lack lexical representations. If pseudowords were as short as the real words (e.g., 5 letters), they might be read too quickly or be mistaken for real words, increasing the risk of confusion or false positives. Second, orthotactic and phonotactic constraints in Turkish make short, plausible pseudowords difficult to construct without unintentionally resembling existing roots or common morphemes. Turkish's agglutinative morphology, in particular, increases the risk of overlap between short pseudowords and real language units. By using longer pseudowords, the design ensures items are clearly non-lexical while still remaining pronounceable and phonotactically valid. Finally, from a materials construction perspective, longer pseudowords offer more flexibility in design. They allow for greater variation while avoiding accidental lexical overlap, thereby preserving the integrity of the nonword condition.

Each section (real words and pseudowords) is timed for one minute. Participants are first given 20 extra words to practice reading aloud accurately. Afterward, they are instructed to read as many words as possible within one minute, with separate trials for real words and pseudowords. One point is awarded for each correctly read word, while a penalty of one point is deducted for each incorrect reading. If a participant self-corrects a misreading, the deducted point is restored. The minimum possible score in each section is 0, while the maximum possible score in each section is 120. We derive a centered (i.e., scaled) composite score of real and pseudo words and use that score in regression analyses, which we call reading composite. Centering used data from all participants.

**Table 2**
Descriptive statistics for the variables per group, standard deviation in parentheses, *t*-tests and effect sizes (Cohen's d).

| | Illiterate | | Literate | | Comparison illiterate vs literate | Cohen's d |
|---|---|---|---|---|---|---|
| | Mean (SD) | Range | Mean (SD) | Range | | |
| Age | 51.5 (12.66) | 21–66 | 43.95 (14.93) | 26–80 | t(37.49) = 1.79, p = 0.082 | 0.55 |
| Education | 0.79 (1.41) | 0–5 | 16.1 (3.38) | 11–24 | t(24.49) = −18.90*** | 0.55 |
| Animals | 13.79 (4.47) | 6–23 | 25 (3.61) | 18–33 | t(41.97) = −9.19*** | 2.73 |
| Fruits | 9.92 (3.13) | 4–18 | 14.9 (3.61) | 6–22 | t(37.97) = −4.84*** | 1.48 |
| Household objects | 15.29 (4.99) | 7–26 | 23.55 (6.25) | 7–32 | t(36.12) = −4.77*** | 1.48 |
| TotalWord | 39 (9.47) | 22–53 | 63.45 (9.80) | 40–75 | t(40.04) = −8.36*** | 2.53 |
| Reading composite | −1.65 (0.58) | −2.00–0.15 | 1.98 (0.90) | 0.23–3.45 | t(31.55) = −15.51*** | 4.87 |
| Real word (max = 120) | 12.29 (19.73) | 0–73 | 100 (17.07) | 71–120 | t(41.92) = −15.80*** | 4.72 |
| Pseudo word (max = 120) | 2.95 (5.74) | 0–20 | 59.95 (18.54) | 22–93 | t(22.04) = −13.26*** | 4.32 |

Note. Education = average of years spent in formal schooling, reading composite = composite score of the real and pseudowords read in one minute, TotalWord = the number of words summed across the three semantic categories. p-Values were Bonferroni corrected for multiple comparisons, the new value was p = 0.007.
*** $p < 0.001$.

### 2.2.2. Semantic fluency test

The test consisted of three categories and presented in the following order: animals, fruits, and household objects. The experimenter introduced the task as a game of speed, accuracy and time. Prior to testing, there was a trial phase with "weather conditions" in which the experimenter led: "*how many weather conditions can you name? Let us name a few together. There is sunny, rainy... what else can you name? Help me*" such that the participants were encouraged to generate words for the category 'weather conditions'. Participants received feedback for their correct or incorrect answers. If they provided an answer that was not relevant to "weather conditions" (i.e., the word *temperature*), then they were given feedback that it does not count because the word temperature is related but is not a weather condition. Once participants gave verbal confirmation in response to understanding the game, each category was presented with a 1-minute timer, and participants were instructed to name as many names as they could for that particular category. If they had questions prior to starting the timer regarding the category (i.e., does any animal count?) then they were answered. No questions were answered during the task. We provided no recommendations for how the participants should structure their word search and output so as to guarantee that whatever cognitive techniques they utilized were spontaneous. The experimenters voice recorded the responses of each participant, and coded the responses per category using the voice recordings. Two native Turkish speakers coded the data and had an agreement of 100 %. All responses were entered, whether the response semantically matched the category or not. However, we counted only the words that matched their respective semantic criterion and reported that here (e.g., *fruits responses: strawberry, apple, cucumber*, score = 2). Responses reported below were the semantically correct unique words. Tomato as a response in the fruit category was marked incorrect because in the Turkish culture and supermarkets tomatoes are sold under vegetables.

### 2.3. Procedure

Illiterate participants were tested individually in quiet, familiar rooms at adult education centers; literate participants were tested individually in a quiet university office. Sessions were voice recorded after informed consent: oral for illiterate participants, and written for literate participants. Given the vulnerable status of illiterate adults, their comfort was prioritized. The experimenter offered positive reinforcement, checked in after each task, and ensured participants wished to continue. Most were engaged and reported enjoying the study.

Demographic data were collected first (age, literacy duration, schooling, highest education level, prior literacy instruction), followed by the 1-Minute Word Reading Task and the semantic fluency task. Sessions lasted 10–15 min. Prompts were repeated neutrally upon request, and responses were recorded both digitally and by hand.

### 2.4. Data analysis

After preprocessing, data were imported into RStudio (2024) for analysis. Descriptive statistics were first obtained. We coded group as literate vs. illiterate and ran *t*-tests using t.test(), with effect sizes calculated via the lsr package (Navarro, 2015) and *p*-values adjusted for multiple comparisons.

We ran a linear regression with total words produced across all categories as the dependent variable. Predictors included age, years of formal education, and a composite reading score derived from real and pseudoword tasks. Despite the high correlation between education and reading ($r = 0.89$, $p < 0.001$), both were included to assess their unique contributions. Education was operationalized as years spent in formal schooling, and reading as a performance-based skill index. All predictors were scaled and centered for regression; descriptive statistics are reported as raw scores except for reading composite.

Non-significant effects were removed stepwise using log-likelihood tests (Crawley, 2010), retaining only significant predictors. Age was included as a covariate to control for cognitive decline. Model assumptions (multicollinearity, heteroscedasticity, and outliers) were tested using the performance package (Lüdecke et al., 2021); outliers were retained to reflect true individual differences. The final model met all assumptions. Data and R code are available at:

https://osf.io/yvd7s/?view_only=080f194e5f9d488496a339f0e9593a58

## 3. Results

Table 2 provides data on age, number of years spent in formal schooling, reading composite, and the performance on the semantic fluency task. As can be seen, the two groups differ significantly from each other on all aspects except for age. The average educational attainment of the illiterate speakers in this study was 0.79 years (roughly 7 months), while literate speakers had, on average, about 16 years of schooling. Literate speakers could read out loud 100 real words per minute correctly, and illiterate speakers could read roughly 12. Similarly, literate speakers outperform illiterate speakers in reading pseudo words (59 pseudowords per minute >2 pseudowords per minute).

Table 2 shows semantic fluency results by literacy group and semantic category. Literate speakers outperformed illiterate speakers across all three categories with large effect sizes. Individual variation was substantial, as reflected in the standard deviations and Fig. 1. Contrary to expectations, illiterate participants produced fewer items in all categories, though some overlap with literate performance occurred. Category patterns differed by group: illiterates produced most words for household objects, then animals and fruits, whereas literates produced most for animals, followed by household objects and fruits.

Table 3 presents the results of the final, best-fitting regression model (including all participants), which identifies the ability to read words (as

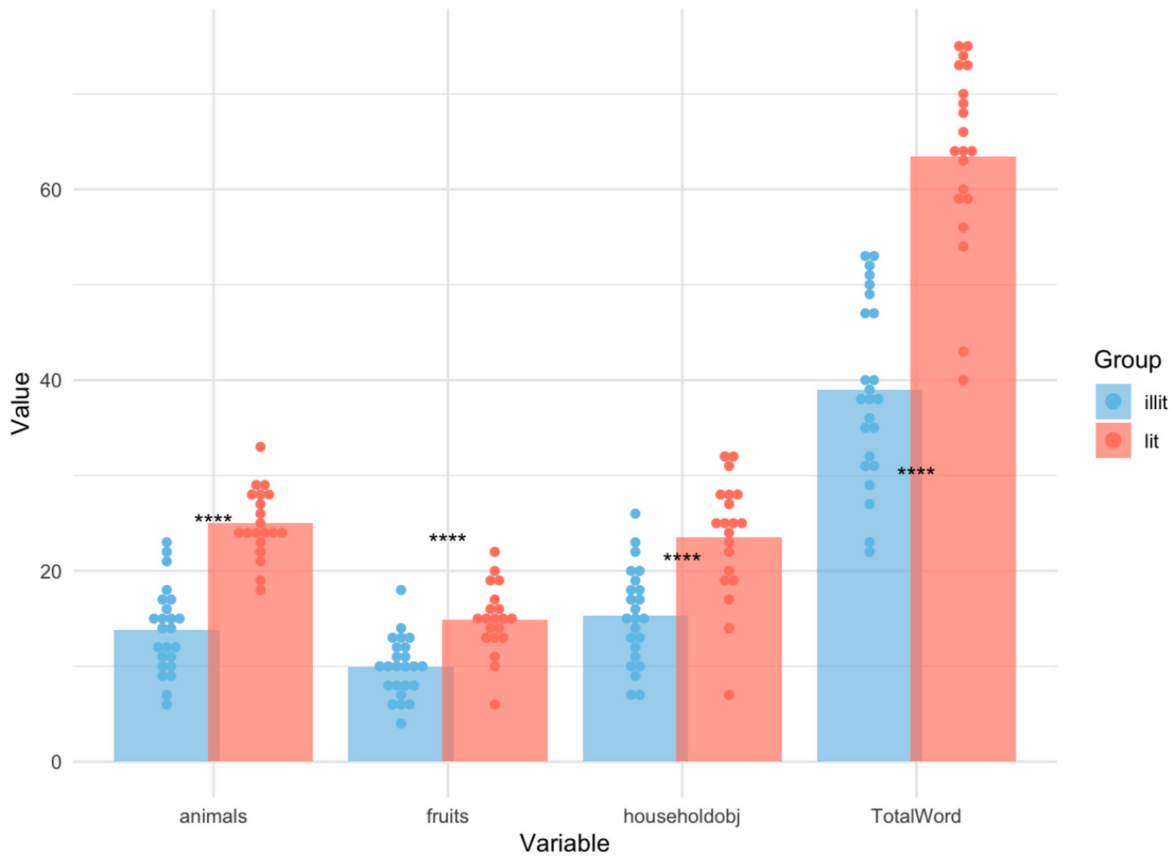


**Fig. 1.** Words produced per category per group.
Note. **** denotes a p < 0.001.

**Table 3**
Final best-fitting model.

| Variable | Estimate | Standard error | *t* value | Pr(>\|t\|) |
|---|---|---|---|---|
| Intercept | −0.000 | 0.098 | 0.000 | 1 |
| Reading composite | 0.765 | 0.092 | 7.717 | $p < 0.0001$ |
| Model $R^2$ | | | | 0.58 |

measured by the reading composite score) as the sole significant predictor of semantic fluency performance (outcome variable = Total-Words). The model explains 58 % of the variance in the data ($R^2 = 0.58$), which indicates a strong explanatory power and suggests that reading composite is a robust predictor of semantic fluency.

The predictor, reading composite, has a standardized estimate of 0.765, which signifies a large effect size. The associated t-value (7.717) and extremely low *p*-value ($p < 0.0001$) demonstrate the statistical significance of this variable. The intercept, which is effectively zero (Estimate = −0.000) with no statistical significance ($p = 1$), suggests that without any contribution from reading composite, the baseline semantic fluency performance is negligible. This reinforces the importance of reading ability as a key driver in the model. Finally, while the model accounts for over half the variance, the $R^2$ values can inflate after model simplification (Grömping, 2007), thus caution is needed when interpreting $R^2$ values. Fig. 2 visualizes the relationship between the performance across the three semantic categories and the overall words produced and reading composite scores per group.

## 4. Discussion

### 4.1. General overview

This study builds on prior work showing that literate adults outperform illiterate adults in semantic fluency tasks, extending these findings by demonstrating that this advantage persists even for familiar, ecologically valid categories (fruits and household objects). These results challenge the view that category familiarity mitigates literacy effects.

We found large individual differences and some overlap in word generation across animals, fruits, and household objects (Fig. 1), consistent with prior studies (Brucki & Rocha, 2004; Da Silva et al., 2004; Ostrosky-Solís et al., 2000; Youn et al., 2011). Our analyses focused on the number of words, so any claim about shared semantic knowledge remains tentative until item-level analyses are conducted.

Comparisons with earlier studies are limited to the animal category, the most commonly tested semantic criterion. Illiterate participants here averaged 13.79 animal names, slightly above the mean of previous studies (11.25), while literate participants averaged 25, far exceeding previous literate means (16.31). Contrary to our hypothesis and some prior findings (Brucki & Rocha, 2004; Da Silva et al., 2004; Kosmidis et al., 2004; Nielsen & Waldemar, 2016), the performance gap did not narrow for familiar categories. Although fruits and household objects showed more overlap between groups, our results suggest that performance depends strongly on literacy and its associated cognitive effects, even for ecologically valid categories.

### 4.2. Effects of literacy

In our regression analysis, reading composite, which may be drawing

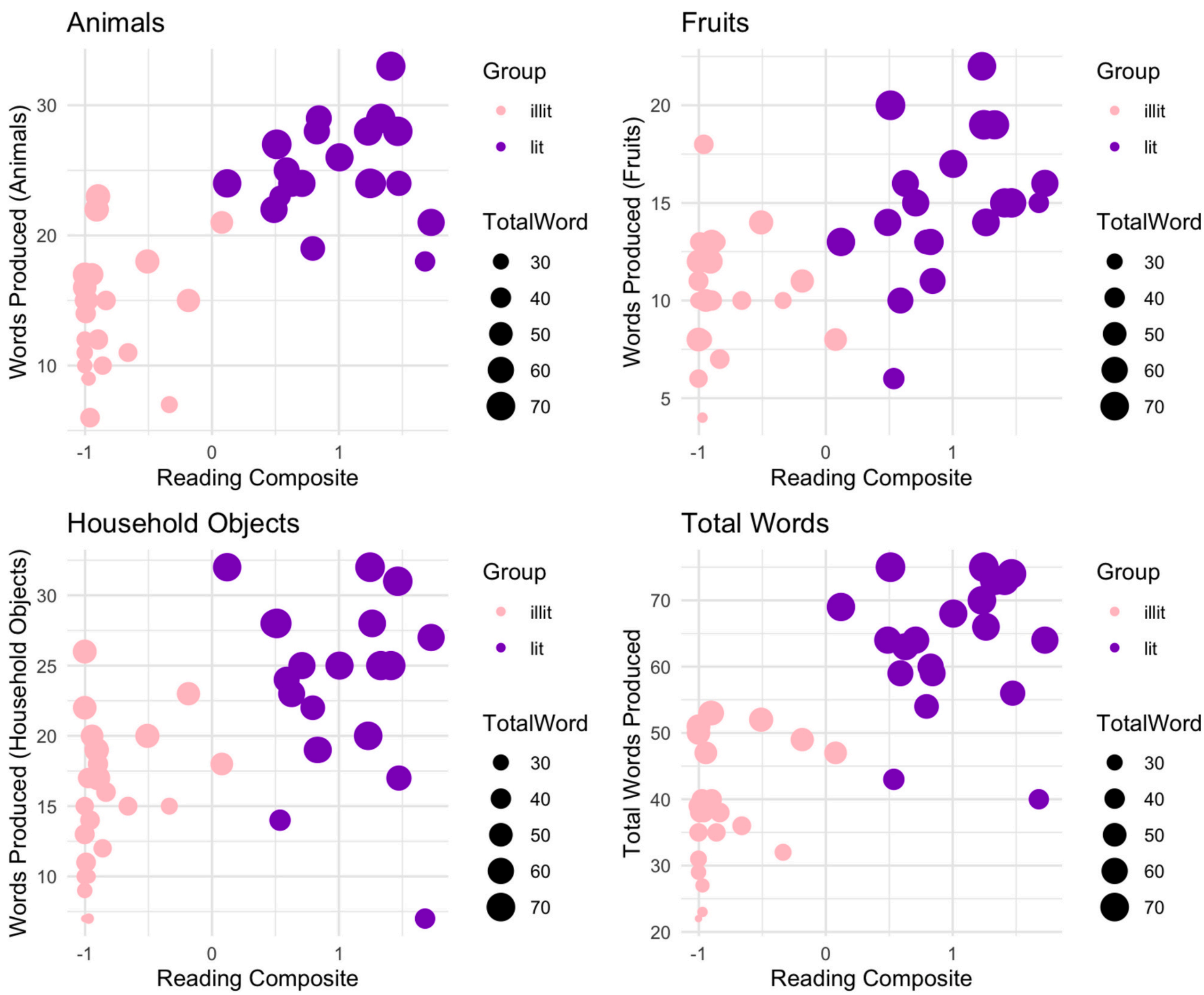


**Fig. 2.** Performance in the semantic fluency task per category and group plotted with reading composite.

on categorical differences (illiterate vs literate) in literacy experience in a continuous form, was the only significant predictor retained after model reduction. These results suggest that word-reading ability itself may play a key role, potentially reflecting broader cognitive effects of literacy. Although literacy (reading composite) was the strongest predictor of semantic fluency, education did not survive the final regression. This likely reflects the high correlation between literacy and education ($r = 0.89$) rather than a lack of effect. Formal schooling may still influence fluency indirectly by expanding vocabulary, exposure to language, and reasoning skills that support word retrieval, as briefly mentioned in the introduction. Literacy captures the proximal cognitive mechanisms, while education provides foundational experiences that enable their development. Future studies disentangling literacy and schooling could clarify their distinct and interactive contributions.

Our findings align with Kolinsky et al. (2014), who found that even rudimentary literacy skills among participants attending literacy classes were associated with higher semantic fluency scores. The group differences we observed across ecologically valid categories may reflect, first, literacy-related gains in general knowledge breadth, and second, more developed retrieval skills in literate individuals through reading, writing, and related practice, which are not mutually exclusive. We discuss these explanations below.

Literacy enriches vocabulary and expands the semantic network, allowing access to category exemplars beyond personal experience. Because written language is lexically richer than spoken language (Hayes & Ahrens, 1988), frequent reading exposes literate individuals to a wider range of items within each category (e.g., passion fruit, dragon fruit), enhancing both vocabulary and general world knowledge (Cunningham & Stanovich, 1998; see also Dąbrowska, 2009; Kosmidis et al., 2006). This advantage may help literate individuals outperform illiterate peers even in familiar categories, and may explain how illiterate speakers with rudimentary reading and writing skills produce more items than their peers who do not yet know how to read and write.

Fig. 2 shows a positive association between reading composite scores and semantic fluency across all categories, an effect that appears in both groups. Prior work suggests that strong oral vocabulary facilitates literacy acquisition (Lee, 2011), which in turn enhances cognition and language in a reciprocal cycle (Stanovich & Cunningham, 1992). Even in preschoolers, richer home literacy environments predict better phoneme and rhyme awareness (Foy & Mann, 2003). Exposure to written materials expands vocabulary, providing more items to retrieve within categories. If reading composite reflects larger vocabulary, its link to semantic fluency is theoretically plausible. Alternatively, it may capture effects of phonemic awareness. Other researchers Ardila et al. (2000) found that illiterate adults generated more words post-alphabetization since literacy may improve categorization of existing vocabulary (Foy & Mann, 2003; Tsegaye et al., 2011). These accounts are not mutually exclusive and may jointly explain the observed relationship.

A second explanation for the group differences is the neurological changes following literacy acquisition. Literacy reshapes brain structures involved in language and memory (Dehaene et al., 2010), potentially enhancing retrieval efficiency. Semantic fluency tasks demand not only knowledge but also rapid, precise retrieval under time pressure—skills that may be supported by literacy-related gains in working memory (Petersson et al., 2007) and retrieval mechanisms. Illiterate speakers may know the items but struggle to access them quickly, leading to cognitive overload. Literate individuals, with years of reading experience, may more efficiently organize and retrieve items, explaining their advantage even in familiar categories and the regression result linking alphabetization experience to fluency performance.

This study has several limitations. The sample size was small, the age range wide, and years of formal education showed little variance in the illiterate group, limiting analyses of its effects. Future work should recruit larger, more demographically balanced samples. To better understand why performance gaps persist despite the presumed ecological validity of certain categories, item-level analyses could reveal whether specific subcategories reduce group differences (Da Silva et al., 2004). Less time-constrained or recognition-based tasks may also uncover subtler effects of ecological validity. Finally, adopting more fine-grained semantic fluency measures (e.g., Da Silva et al., 2004) could yield deeper insights into group performance.

## 5. Conclusion

This study investigated the role of literacy, and ecological validity in semantic fluency tasks among Turkish illiterate and literate speakers. Our findings reveal significant differences in fluency performance across all tested categories, with literate speakers consistently outperforming their illiterate counterparts. Notably, the persistence of performance gaps in categories with presumed high ecological validity, such as fruits and household objects, challenges the assumption that experiential familiarity alone can mitigate disparities in semantic fluency.

Regression analyses further highlighted literacy as the most significant predictor of fluency performance, surpassing the contributions of formal education or age. This highlights the cognitive and linguistic advantages conferred by literacy, including enhanced semantic organization and retrieval efficiency. The findings align with prior research emphasizing the transformative impact of literacy on brain structures and cognitive functions, while also offering new insights into how literacy intersects with ecological validity in shaping linguistic outcomes.

## CRediT authorship contribution statement

**Tan Arda Gedik:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Vania De la Garza:** Writing – review & editing, Visualization, Validation, Methodology, Investigation.

## Declaration of competing interest

The authors declare no competing interests.

## Acknowledgements

We would like to sincerely thank Gülçin İrem Yıldırım for her help, dedication and excellence during data collection. We are also deeply grateful to Neslihan Yıldırım Turan for her help during data collection at the literacy center. Finally, we are forever grateful to our participants for taking part in our studies.

## Appendix. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.actpsy.2025.105735.

## Appendix A

**Table 1**
Full regression model.

| Variable | Estimate | SE | t-Value | p-Value |
|---|---|---|---|---|
| Intercept | 0 | 0.1 | 0 | 1 |
| Age | 0.04 | 0.11 | 0.33 | 0.743 |
| YearsofEdu | 0.23 | 0.27 | 0.84 | 0.404 |
| Reading composite | 0.57 | 0.28 | 2.07 | 0.045 |

**Table 2**
Pearson correlations for the entire data set.

| Variable | Age | YearsofEdu | ReadReal | Readpseudo | Reading composite | animals | fruits | householdobj | TotalWord |
|---|---|---|---|---|---|---|---|---|---|
| Age | 1 | −0.34* | −0.37* | −0.41** | −0.39** | −0.16 | −0.48** | −0.14 | −0.26 |
| YearsofEdu | −0.34* | 1 | 0.92*** | 0.92*** | 0.93*** | 0.75*** | 0.59*** | 0.56*** | 0.74*** |
| ReadReal | −0.37* | 0.92*** | 1 | 0.95*** | 0.99*** | 0.79*** | 0.62*** | 0.60*** | 0.78*** |
| Readpseudo | −0.41** | 0.92*** | 0.95*** | 1 | 0.99*** | 0.75*** | 0.63*** | 0.51*** | 0.73*** |
| Reading composite | −0.39** | 0.93*** | 0.99*** | 0.99*** | 1 | 0.78*** | 0.63*** | 0.56*** | 0.77*** |
| animals | −0.16 | 0.75*** | 0.79*** | 0.75*** | 0.78*** | 1 | 0.54*** | 0.69*** | 0.90*** |
| fruits | −0.48** | 0.59*** | 0.62*** | 0.63*** | 0.63*** | 0.54*** | 1 | 0.54*** | 0.75*** |
| householdobj | −0.14 | 0.56*** | 0.60*** | 0.51*** | 0.56*** | 0.69*** | 0.54*** | 1 | 0.90*** |
| TotalWord | −0.26 | 0.74*** | 0.78*** | 0.73*** | 0.77*** | 0.90*** | 0.75*** | 0.90*** | 1 |

* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

**Table 3**
Pearson correlations for the illiterate data set.

| Variable | Age | YearsofEdu | ReadReal | Readpseudo | Reading composite | animals | fruits | householdobj | TotalWord |
|---|---|---|---|---|---|---|---|---|---|
| Age | 1 | −0.27 | 0.01 | −0.09 | −0.02 | 0.16 | −0.23 | 0.18 | 0.1 |
| YearsofEdu | −0.27 | 1 | −0.09 | −0.15 | −0.11 | 0.15 | −0.07 | −0.08 | 0.01 |
| ReadReal | 0.01 | −0.09 | 1 | 0.95*** | 0.99*** | 0.23 | 0.08 | 0.35 | 0.32 |
| Readpseudo | −0.09 | −0.15 | 0.95*** | 1 | 0.97*** | 0.15 | 0.05 | 0.31 | 0.25 |
| Reading composite | −0.02 | −0.11 | 0.99*** | 0.97*** | 1 | 0.21 | 0.07 | 0.34 | 0.3 |
| animals | 0.16 | 0.15 | 0.23 | 0.15 | 0.21 | 1 | 0.04 | 0.59** | 0.80*** |
| fruits | −0.23 | −0.07 | 0.08 | 0.05 | 0.07 | 0.04 | 1 | 0.25 | 0.48* |
| householdobj | 0.18 | −0.08 | 0.35 | 0.31 | 0.34 | 0.59** | 0.25 | 1 | 0.89*** |
| TotalWord | 0.1 | 0.01 | 0.32 | 0.25 | 0.3 | 0.80*** | 0.48* | 0.89*** | 1 |

* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

**Table 4**
Pearson correlations for the literate data set.

| Variable | Age | YearsofEdu | ReadReal | Readpseudo | Reading composite | animals | fruits | householdobj | TotalWord |
|---|---|---|---|---|---|---|---|---|---|
| Age | 1 | −0.3 | −0.72*** | −0.56* | −0.65** | 0.02 | −0.58** | −0.09 | −0.27 |
| YearsofEdu | −0.3 | 1 | 0.63** | 0.47* | 0.55* | −0.37 | 0.15 | −0.03 | −0.09 |
| ReadReal | −0.72*** | 0.63** | 1 | 0.81*** | 0.92*** | 0.09 | 0.4 | −0.11 | 0.11 |
| Readpseudo | −0.56* | 0.47* | 0.81*** | 1 | 0.97*** | 0 | 0.32 | −0.25 | −0.04 |
| Reading composite | −0.65** | 0.55* | 0.92*** | 0.97*** | 1 | 0.04 | 0.37 | −0.2 | 0.02 |
| animals | 0.02 | −0.37 | 0.09 | 0 | 0.04 | 1 | 0.22 | 0.25 | 0.61** |
| fruits | −0.58** | 0.15 | 0.4 | 0.32 | 0.37 | 0.22 | 1 | 0.31 | 0.64** |
| householdobj | −0.09 | −0.03 | −0.11 | −0.25 | −0.2 | 0.25 | 0.31 | 1 | 0.84*** |
| TotalWord | −0.27 | −0.09 | 0.11 | −0.04 | 0.02 | 0.61** | 0.64** | 0.84*** | 1 |

* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

## Data availability

Data in the manuscript

## References


Ardila, A., Ostrosky-Solis, F., & Mendoza, V. U. (2000). Learning to read is much more than learning to read: A neuropsychologically based reading program. *Journal of the International Neuropsychological Society, 6*(7), 789–801. https://doi.org/10.1017/S1355617700677068

Brucki, S. M. D., & Rocha, M. S. G. (2004). Category fluency test: Effects of age, gender and education on total scores, clustering and switching in Brazilian Portuguese-speaking subjects. *Brazilian Journal of Medical and Biological Research, 37*(12), 1771–1777. https://doi.org/10.1590/S0100-879X2004001200002

Crawley, M. (2010). *R a language and environment for statistical computing: Reference index. R Foundation for Statistical Computing*.

Cunningham, A. E., & Stanovich, K. E. (1998). What reading does for the mind. In *, Vol. 4389. Abstracts of scientific articles* (pp. 128–130).

Da Silva, C. G., Petersson, K. M., Faísca, L., Ingvar, M., & Reis, A. (2004). The effects of literacy and education on the quantitative and qualitative aspects of semantic verbal fluency. *Journal of Clinical and Experimental Neuropsychology, 26*(2), 266–277. https://doi.org/10.1076/jcen.26.2.266.28089

Dąbrowska, E. (2009). Words as constructions. In V. Evans, & S. Pourcel (Eds.)*, Vol. 24. Human cognitive processing* (pp. 201–223). John Benjamins Publishing Company. https://doi.org/10.1075/hcp.24.16dab.

Deeney, J. (2010). The impact of timed word reading tasks on literacy assessment. *Literacy, 44*(1), 3–10.

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Filho, G. N., Jobert, A., … Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science, 330*(6009), 1359–1364.

Denervaud, S., Christensen, A. P., Kenett, Y. N., et al. (2021). Education shapes the structure of semantic memory and impacts creative thinking. *npj Science of Learning, 6*, Article 35. https://doi.org/10.1038/s41539-021-00113-8

Foy, J. G., & Mann, V. (2003). Home literacy environment and phonological awareness in preschool children: Differential effects for rhyme and phoneme awareness. *Applied PsychoLinguistics, 24*(1), 59–88. https://doi.org/10.1017/S0142716403000043 (Cambridge Core).

Gedik, T.A. forthcoming. Literacy at Work: Ultimate Native Language Attainment in Turkish. Boston/New York: DeGruyter.

Grömping, U. (2007). Estimators of relative importance in linear regression based on variance decomposition. *The American Statistician, 61*(2), 139–147. https://doi.org/10.1198/000313007X188252

Hayes, D. P., & Ahrens, M. G. (1988). Vocabulary simplification for children: A special case of 'motherese'? *Journal of Child Language, 15*(2), 395–410.

Huettig, F., & Mishra, R. K. (2014). How literacy acquisition affects the illiterate mind–a critical examination of theories and evidence. *Lang & Ling Compass, 8*(10), 401–427.

Kolinsky, R., Monteiro-Plantin, R. S., Mengarda, E. J., Grimm-Cabral, L., Scliar-Cabral, L., & Morais, J. (2014). How formal education and literacy impact on the content and structure of semantic categories. *Trends in Neuroscience and Education, 3*(3–4), 106–121. https://doi.org/10.1016/j.tine.2014.08.001

Kosmidis, M. H., Tsapkini, K., & Folia, V. (2006). Lexical processing in illiteracy: Effect of literacy or education? *Cortex, 42*(7), 1021–1027.

Kosmidis, M. H., Tsapkini, K., Folia, V., Vlahou, C. H., & Kiosseoglou, G. (2004). Semantic and phonological processing in illiteracy. *Journal of the International Neuropsychological Society, 10*(6), 818–827. https://doi.org/10.1017/S1355617704106036

Lee, J. (2011). Size matters: Early vocabulary as a predictor of language and literacy competence. *Applied PsychoLinguistics, 32*(1), 69–92.

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). Performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software, 6*(60), Article 3139. https://doi.org/10.21105/joss.03139

Manly, J. J., Jacobs, D. M., Sano, M., Bell, K., Merchant, C. A., Small, S. A., & Stern, Y. (1999). Effect of literacy on neuropsychological test performance in nondemented, education-matched elders. *Journal of the International Neuropsychological Society, 5*, 191–202. https://doi.org/10.1017/S135561779953302X

Mathuranath, P. S., George, A., Cherian, P. J., Alexander, A., Sarma, S. G., & Sarma, P. S. (2003). Effects of age, education and gender on verbal fluency. *Journal of Clinical and Experimental Neuropsychology, 25*(8), 1057–1064. https://doi.org/10.1076/jcen.25.8.1057.16736

Mitrushina, M., Boone, K. B., & Da Silva, C. G. (2005). In *Handbook of Normative data for neuropsycho-logical assessment* (2nd ed.). New York, NY: Oxford University Press.

Navarro, D. (2015). *Learning statistics with R: A tutorial for psychology students and other beginners. (Version*. University of New South Wales. https://learningstatisticswithr.com.

Nielsen, T. R., & Waldemar, G. (2016). Effects of literacy on semantic verbal fluency in an immigrant population. *Aging, Neuropsychology, and Cognition, 23*(5), 578–590.

Ostrosky-Solís, F., Ardila, A., & Rosselli, M. (1999). NEUROPSI: A brief neuropsychological test battery in Spanish with norms by age and educational level. *Journal of the International Neuropsychological Society, 5*(5), 413–433. https://doi.org/10.1017/S1355617799555045

Ostrosky-Solís, F., López-Arango, G., & Ardila, A. (2000). Sensitivity and specificity of the Mini-Mental State Examination in a Spanish-speaking population. *Applied Neuropsychology, 7*(1), 25–31. https://doi.org/10.1207/S15324826AN0701_4

Petersson, K. M., Silva, C., Castro-Caldas, A., Ingvar, M., & Reis, A. (2007). Literacy: A cultural influence on functional left-right differences in the inferior parietal cortex:

Literacy and inferior parietal cortex. *European Journal of Neuroscience, 26*(3), 791–799. https://doi.org/10.1111/j.1460-9568.2007.05701.x

Ritchie, S. J., & Tucker-Drob, E. M. (2018). How much does education improve intelligence? A meta-analysis. *Psychological Science, 29*(8), 1358–1369. https://doi.org/10.1177/0956797618774253

Rosselli, M., Ardila, A., & Rosas, P. (1990). Neuropsychological assessment in illiterates. *Brain and Cognition, 12*(2), 281–296. https://doi.org/10.1016/0278-2626(90)90020-O

Simos, P. G., Sideridis, G. D., Kasselimis, D., & Mouzaki, A. (2013). Reading fluency estimates of current intellectual function: Demographic factors and effects of type of stimuli. *Journal of the International Neuropsychological Society, 19*(3), 355–361.

Stanovich, K. E., & Cunningham, A. E. (1992). Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. *Memory & Cognition, 20*, 51–68.

Sternberg, R. J. (1987). Most vocabulary is learned from context. In M. G. McKeown, & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 89–106). Erlbaum.

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

Torgeson, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test review: Test of word reading efficiency (TOWRE)*. Austin, TX, USA: Inc.

Tsegaye, M. T., Bleser, R. D., & Iribarren, C. (2011). The effect of literacy on oral language processing: Implications for aphasia tests. *Clinical Linguistics & Phonetics, 25* (6–7), 628–639. https://doi.org/10.3109/02699206.2011.567348

UNESCO. (1978). *Towards a methodology for projecting rates of literacy and educational attainment* (Paris, France).

UNESCO. (2024). Literacy. https://www.unesco.org/en/literacy.

Vágvölgyi, R., Coldea, A., Dresler, T., Schrader, J., & Nuerk, H.-C. (2016). A Review about Functional Illiteracy: Definition, Cognitive, Linguistic, and Numerical Aspects. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.01617

Yasa Kostas, R., Kostas, K., MacPherson, S. E., & Wolters, M. K. (2024). Semantic verbal fluency in native speakers of Turkish: A systematic review of category use, scoring metrics and normative data in healthy individuals. *Journal of Clinical and Experimental Neuropsychology, 46*(3), 272–301.

Yıldız, M., Yildirim, K., Ates, S., Rasinski, T., Fitzgerald, S., & Zimmerman, B. (2014). The relationship between reading fluency and reading comprehension in fifth-grade Turkish students. *International Journal of School & Educational Psychology, 2*(1), 35–44.

Youn, J.-H., Siksou, M., Mackin, R. S., Choi, J.-S., Chey, J., & Lee, J.-Y. (2011). Differentiating illiteracy from Alzheimer's disease by using neuropsychological assessments. *International Psychogeriatrics, 23*(10), 1560–1568. https://doi.org/10.1017/S1041610211001347