



HAL
open science

Semantic and phonological associative norms of 145 French words for children aged 5–9 years: Free associations and forced-choice task

Amandine Hippolyte, Nicolas Ribeiro, Laure Ibernou, Nathalie Marec-Breton,
Christelle Declercq

► To cite this version:

Amandine Hippolyte, Nicolas Ribeiro, Laure Ibernou, Nathalie Marec-Breton, Christelle Declercq. Semantic and phonological associative norms of 145 French words for children aged 5–9 years: Free associations and forced-choice task. *First Language*, 2025, 10.1177/01427237251313567 . hal-04920464

HAL Id: hal-04920464

<https://u-picardie.hal.science/hal-04920464v1>

Submitted on 30 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

**Semantic and phonological associative norms of 145 French words for children aged 5 to 9 years:
free associations and forced-choice task**

Authors: Hippolyte, A.^{*a1}, Ribeiro, N.^a, Ibernon, L.^a, Marec-Breton, N.^b and Declercq, C.^c

Affiliations

^a Centre de Recherche en Psychologie: Cognition, Psychisme et Organisations (UR 7273), Université de Picardie Jules Verne, Amiens, France

^b Université Rennes 2, LP3C, EA 1285, Rennes, **France**

^c Université de Reims Champagne-Ardenne, C2S, Reims, France

Acknowledgments section

We would like to express our gratitude to all the children who took part in this study, as well as their parents, the schools and the teachers involved in this research. Finally, we would like to thank Dr Valérie Golly-Ledoux for her involvement in data collection.

Author contributions

Amandine Hippolyte: Investigation, Resources, Data curation, Writing – Original draft, Writing – Review and Editing **Nicolas Ribeiro:** Formal analysis, Data curation, Writing – Original draft, Writing – Review and Editing **Nathalie Marec-Breton:** Conceptualization, Methodology, Resources, Writing – Review and Editing **Laure Ibernon:** Conceptualization, Methodology, Resources, Writing – Review and Editing, Funding acquisition **Christelle Declercq:** Conceptualization, Methodology, Resources, Writing – Review and Editing, Project administration.

¹ Correspondence to: CRP-CPO - UR UPJV 7273, Université de Picardie Jules Verne, Pôle Campus Sud - Bâtiment C - 2ème étage - 1, Chemin du Thil - CS 52 501 - 80025 Amiens Cedex 1 – France. Fax : +33 3 22 82 74 08.

E-mail addresses: amandine.hippolyte1@gmail.com (A. Hippolyte), laure.ibernon@u-picardie.fr (L. Ibernon).

Statements and declarations

Ethical considerations

ID-RCB: 2022-A00341-42

Consent to participate

All legal representatives of the participants have provided their agreement to participate.

Consent for publication

Not applicable, as no personal information has been published.

Declaration of conflicting interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This work was supported by the “French National Research Agency (ANR)”.

Data availability statement

Data are available at the following link:

https://osf.io/ud3be/?view_only=3cb257fa6b9048ce8c1795ea87eaaaab.

Statement of AI Utilization

Spelling and grammar were verified using an AI language model (ChatGPT).

Abstract

This study aimed to establish normative data for 145 words using phonological and semantic association tasks with 242 French schoolchildren, ranging from ages 5 (Grande Section) to 8 (Cours Élémentaire 2), providing a fundamental resource for future research and educational planning.

The participants were engaged in two primary tasks: a Free Association task, entailing image recognition followed by phonological or semantic word association; and a Forced Choice task, involving choosing semantically or phonologically related images from options.

Overall, there were more associations produced for semantic associations than phonological ones; however, there was a greater variety of phonological associations than semantic ones. We found that 5-year-old children performed better in semantic tasks by choosing expected associates compared to phonological tasks. However, this difference was not present in older children. This suggests that phonological networks develop more gradually than semantic ones, with phonological awareness continuing to improve with age while semantic stability is reached earlier. The established norms offer a framework for educators, clinicians, and researchers in understanding and assessing the semantic and phonological associative strategy of children in early educational settings within the French context.

Keywords: Associative abilities, children, semantic associations, phonological associations, cognitive development, early education

1. Introduction

Establishing language norms in psycholinguistics is a fundamental approach, particularly concerning associative norms. These norms are often used with adult participants to investigate how knowledge affects performance in memory and linguistic tasks (e.g. Cramer, 1968; Nelson et al., 2004; Nelson et al., 1998). A key aspect of language is its distinct representations in memory, which evolve over time. In early development, the oral mental lexicon, composed of phonological and semantic representations, emerges first, as children are initially exposed to oral language (Quémart et al., 2011). As they grow and begin to interact with written language, orthographic representations subsequently develop. This progression highlights how critical it is to understand the mechanics of language, not only in adults but also in children as they develop these linguistic skills. Associative norms, which provide insights into these language representations, are obtained by presenting individual words to a large sample of participants. They are asked to spontaneously produce the first word that comes to mind (Bonin et al., 2013). The strength of association can be calculated by counting the number of participants in the total sample who produced a specific word. The diversity of associations can be estimated by counting the number of different (non-idiosyncratic) responses generated by the sample (Nelson et al., 2004). Although association norms are commonly used in adult research, they are less widespread in developmental research, especially with non-reading children. This gap suggests that further exploration is needed to understand how associative norms evolve during development, particularly in children who are not yet literate. In this study, our focus was solely on oral language representations, excluding orthographic representations, as we sought to investigate the early stages of language development prior to literacy acquisition.

A semantic association is the link between the meanings of two words. Children establish several types of semantic associations. Taxonomic relations (Blaye & Bonthoux, 2001; Fenson et al., 1988) link concepts within the same category, such as “chien” (“dog”) and “chat” (“cat”) whereas thematic/associative relations (Blaye & Bonthoux, 2001) are based on contexts or functions, for example, “chien” (“dog”) and “niche” (“doghouse”). Perceptual characteristics, such as shape and color, also influence associations (Fenson et al., 1988; Imai et al., 1994). Understanding these varied association types not only sheds light on the cognitive strategies utilized by different age groups but also provides a framework for constructing effective experimental tasks.

A phonological association is when two different words share phonemes. This sharing can occur at the syllable (e.g. **camarade-matin**), phoneme (e.g. **chat-chien**), onset (e.g. **table-tasse**) or rhyme (e.g. **chat-mat**) level. Therefore, the phonological network may differ from the established semantic network, as associates sharing identical phonemes may not share a semantic relationship.

The study of semantic and phonological associations is crucial in understanding children’s language development. By analysing these connections, we can gain insight into how children

construct links between words and how phonological and semantic processing evolves over time. For example, Arias-Trejo et al. (2022) conducted a study examining semantic and phonological organization in very young children aged 18, 21, and 24 months. They selected ninety-six concrete nouns familiar to toddlers in this age range from adapted communicative inventories and developed thirty-two triads consisting of prime, target, and distractor words. The lack of readily available materials for children of this age necessitated the creation of their own resources, but utilizing existing pairs would have simplified their task. Establishing a controlled database for this type of research could simplify future studies, allowing scientists to save time and better control the items used.

The first semantic and phonological association databases were published in English. The Edinburgh Associative Thesaurus created by Kiss et al. (1973) is an associative database consisting of words and their associates given by participants. The relation between the words is purely associative. The University of South Florida Free Association Norms is a database created in the USA (Nelson et al., 2004) to control the strength of semantic and phonological links between words in an experimental task. Participants were asked to spontaneously evoke the first word that came to mind in response to two distinct instructions: in one case, the words had to be associated semantically, while in the other they had to rhyme. This database presents both semantic and phonological associates. For phonological associates databases, the English Lexicon Project (Balota et al., 2007) focuses on the phonological similarity between words and includes detailed information on the phonological neighbors associated with each term. Although word association databases exist in several languages (e.g., Spanish: Fernandez et al., 2004; Portuguese: Comesaña et al., 2014), most tools are designed for English.

Databases of verbal associations in French are relatively rare (see Table 1). Ferrand and Alario (1998) were among the first to develop semantic association norms, including adult participants. Their study established valuable references for understanding semantic associations in French. Bonin et al. (2013) later expanded this research by proposing semantic association norms for 520 concrete words. Additionally, the French Lexicon Project (Ferrand et al., 2010) provides information on the phonological neighbors associated with each word. De La Haye (2003) developed verbal association norms for four groups of participants, including children aged 9 to 11 and adults. However, it is important to note that there are no specific norms for phonological associations in children.

Table 1. Summary of French Databases on Semantic and Phonological Association Norms.

Project	Age Group	Number of Words	Type of Words	Type of Instructions used for creation of database	Originality/Comments	
Bonin Association Norms (Bonin et al., 2013)	Adults	520 words	Concrete words	Verbal association	Free association task; participants were asked to write down legibly the first word that came to mind for each word presented in uppercase.	Contributes to expanding the available verbal association norms for French, while providing insights into their relationships with other psycholinguistic variables, which had not been systematically addressed before.
De La Haye Verbal Association Norms (De La Haye, 2003)	Children (9-11), Adults	240 words	Nouns, adjectives, verbs, and other common words, adapted to age groups	Verbal association	Free association task: participants wrote the first word that came to mind in response to a given written stimulus word	One of the rare associative norm databases for French-speaking children, offering valuable insights into age-related differences in word associations.
Ferrand et Alario Association Norms (Ferrand & Alario, 1998)	Adults	366 words	Object names	Verbal association	Free association task; participants were asked to write down the first word that came to mind for each word presented in alphabetical order	One of the first large-scale associative norm datasets published in French.
French Lexicon Project (Ferrand et al., 2010)	Adults	38,840 words and 38,840 nonwords	Words across various grammatical categories	Lexical decision task	No direct association task; data comes from lexical decision and naming tasks where phonological and orthographic neighbors are computed indirectly	A key resource for studying the lexical characteristics of French words, inspired by the English Lexicon Project.

Some of the cues provided by associative norms (size of set of associates, relative strength of associates, number of idiosyncratic responses, etc.) reflect how associative knowledge is organized and may therefore provide useful ways of studying developmental differences in knowledge organization and restructuring. Although infants show sensitivity to taxonomic pairs as early as 16-18 months (Arias-Trejo & Plunkett, 2013; Bauer & Mandler, 1989), the explicit formulation of taxonomic relationships seems to pose more challenges for children. This ability to articulate taxonomic structures generally emerges later in development, around the age of 9 years (Lucariello et al., 1992; Sell, 1992) and results from a gradual and extended process throughout childhood. This suggests a developmental distinction between early recognition of taxonomic relationships and the later, more mature ability to express these relationships verbally. Moreover, the acquisition of vocabulary seems to be linked to the level of association with previously known words (Vasanta & Sailaja, 1999). Therefore, like adults, associative norms can be a valuable resource for investigating organisation and restructuring of knowledge in relation to age.

The aim of this study was to produce an original French database containing norms pertaining to oral semantic and phonological associations. Both types of associations were collected from children aged 5 to 8 years through two tasks. The first was a Free Association task, following the most commonly used format in previous research aimed at building an association database. The second task was a Forced-Choice task in which children were asked to choose among three pictures, each representing a word, the one that best corresponded to a target word also illustrated by a picture. The free association task had the advantage of allowing the collection of children's spontaneous answers for each word, which are assumed to reflect the most direct and natural associations. However, while free association tasks provide valuable information about young children's associations, they have certain limitations due to the cognitive and linguistic demands they place on participants. The forced-choice task allowed these limitations to be overcome. Thus, the norms were developed with children through these two complementary tasks. To this purpose, we compiled sets comprising phonologically and semantically linked words alongside distractors devoid of any shared phonological or semantic information. The stimuli provided were illustrated for young children or individuals with neurodevelopmental difficulties who do not have access to written language and could be used by researchers for their material. We expected differences between groups and across tasks; however, due to the lack of research on non-reading French-speaking children, the direction of these differences remained unclear. Additionally, we expected more variability among younger children compared to older ones. Since this topic has not been studied before, and no norms exist for children regarding oral phonological and semantic associations, both areas needed to be investigated.

2. Methodology

2.1. Participants

A total of 242 children were recruited, spanning four grade levels from *Grande Section* to *Cours Élémentaire 2* (Males = 115; Females = 127; Mean age in months = 82.37). These children were educated in the French school system over a period of four years which are aimed at children of 5, 6, 7 and 8. Group characteristics are presented in table 2. Children were recruited from six schools in Rennes and Reims, France, including schools from both standard (4 schools) and priority educational networks (2 schools). This approach ensured representation from a range of economic backgrounds, as priority networks focus on supporting schools in more disadvantaged areas. Note that vocabulary size was not measured in this study.

Inclusion criteria required that participants did not have any identified special educational needs, had French as their mother tongue, and clearly understood the instructions, as verified by a test question before each task.

Table 2. Participant characteristics

Group	N	Mean age (in months)	SD	Range (in months)	Sex ratio (boys/girls)
5-year-old	84	65.12	4.10	58 – 73	43/41
6-year-old	43	77.56	4.68	70 – 85	18/25
7-year-old	58	88.69	4.19	81 – 98	33/25
8-year-old	57	101.30	4.98	94 - 117	21/36

A priori power analyses were conducted using G*Power 3.1.9.7 (Faul et al., 2007) to determine the required sample size for our primary analyses. For ANOVA tests, used to compare performance between semantic and phonological tasks across age groups, assuming a medium effect size ($w = 0.3$) with alpha set at .05 and desired power of .80, the analysis indicated a minimum required sample of 197 participants. For comparisons of proportions (e.g., expected, opposite, and distractor responses) between age groups and task types, we aimed for a larger sample to ensure sufficient power to detect potentially smaller effects. Consequently, we recruited 242 participants (approximately 60 per age group: 5-, 6-, 7-, and 8y.o.), allowing for robust analyses across all our research questions, including the examination of age-related differences in phonological and semantic associative abilities across different task types and response categories. This sample size provides adequate statistical power for our planned analyses while accounting for the complexity of our study design.

All study procedures were in accordance with European legislation and ethical guidelines. Written informed consent was obtained from both parents and children prior to their participation in the study tasks.

2.2. Material

This study covered two associative tasks: one free association task and one forced choice task. These tasks had two modalities as the association criteria can be phonological or semantic. For both tasks, the material consisted of a total of 145 identical words randomized into two orders; each order was then subdivided into 6 lists of 25 words each. There was no repetition across the lists from order #1 and none across the lists of order #2. The lists from order #1 were preferentially used for the semantic task, while those from order #2 were used for the phonological task. Each child participated in both modalities, performing the semantic and phonological tasks in two sessions. The order of the tasks was counterbalanced across participants. Therefore, each child engaged in one Free Association task and one Forced Choice task, with conditions (semantic or phonological) counterbalanced. The words were spoken aloud, accompanied by an image. The images were displayed on a slideshow.

Word selection

Words were selected using the Manulex lexicon. Manulex is a lexical database that provides word occurrence frequencies, calculated from a corpus of 54 French textbooks (Lété et al., 2004). This database is based on lexical frequency within children's stories and is organized by French school grade levels and corresponding age groups. It includes lists of word frequencies for three age groups: CP (approximately 6 years), CE1 (7 years), and cycle 3 (CE2-CM2, 8-11 years). A fourth level aggregates the entire range of primary school textbooks for a comprehensive view of school vocabulary. For our study, we selected the CP list as it represents the youngest age group, corresponding to an age up to 6 years.

The Standard Frequency Index (SFI) is the frequency index used in the Manulex database to represent word frequency. The SFI is a logarithmic derivative of the U index that provides a quick estimation of a word's occurrence scale. It allows for a more intuitive understanding of how often a word appears in texts. The formula is: $SFI=10*(\log_{10}(U)+4)$.

The selection criteria were twofold: first, to maximize the word's frequency in **French** (with a minimum frequency example being "*mouette*" (seagull) with SFI = 34.87, a maximum frequency example being "*maison*" (house) with SFI = 72.03, and a mean frequency of 57.87 with a standard deviation of 7.18); and second, to ensure the availability of a distinctive image counterpart in the Larousse dictionary while considering the presence of phonological and semantic analogs. Only nouns were selected. Manulex was used in this research despite being a written language database, as there is currently no such database for spoken language.

For each word, a semantic associate, a phonological associate, and a distractor were selected. The semantic associate had to be linked to the target word either taxonomically (e.g. coquelicot/rose [poppy/rose]) or associatively (e.g. maison/escalier [house/stairs]) with approximately the same number of phonemes, be imageable in Larousse, and exclude any phonetic similarity. The phonological associate had to share at least one phoneme if not more, at the onset of the word, and have no semantic link with the target word. The distractor was chosen to ensure it had no strong phonological connection (no more than one shared phoneme, and none at the onset or rhyme) and no semantic association with the target. Each word also needed to be concrete and easily imageable. To ensure this, a picture was selected to represent each word.

Picture selection

Pictures were chosen from *Mon Tout Premier Dictionnaire Larousse* (2017), a French-language encyclopedic dictionary that illustrates more than 5,000 words and is specifically aimed at the present sample's age category.

Presentation medium

The stimuli were presented on a computer using a slideshow (fig. 1). Each slide contained either a single image or a set of three images, depending on the task. The images were displayed on a screen at a viewing distance of approximately 60 cm, ensuring all participants had a clear and unobstructed view. Each image was sized at 151 x 151 pixels to maintain consistency and clarity across all presentations. The experiment was conducted in a well-lit room to avoid any visual interference.

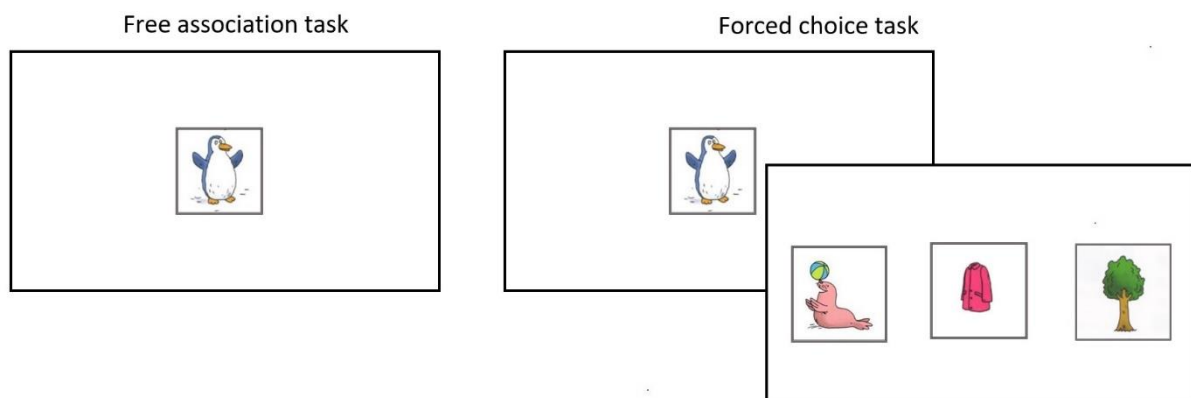


Figure. 1. Board example

Note. Manchot (penguin), phoque (seal), manteau (coat), arbre (tree).

Availability of materials

The first author of this study will provide the pictures and references for the copyrighted material to any researcher or educator upon request.

2.3. Procedure

There were two tasks: Free Association and Forced Choice. For each of these tasks, there were two modalities. In one, the criteria for production or choice were semantic, and in the other, the criteria were phonological. The Free Association task was always completed before the Forced Choice task. The total procedure took half an hour, as each condition needed fifteen minutes. Children were given approximately 15 seconds to respond for each item. If no response was provided, they were prompted once. If they still did not answer, the task moved on to the next item (table 3). For each of the two modalities, sessions were conducted individually in an unused room at the school. The student sat on a chair with the computer positioned in front of them, while the experimenter sat nearby, manually changing the stimuli. The experimenter recorded the student's responses by hand on a paper form.

Free Association

In the Free Association task, for each production of an associate, participants engaged in a two-step process involving image recognition followed immediately by the word association based on the cued word represented in the image:

1. Image Recognition: Participants first viewed an image and identified what it represented. Correct responses were praised, while incorrect ones were gently corrected.

2. Word Association: Participants were prompted to list words that came to mind when thinking about the term associated with the displayed image.

During the training session, participants received praise for any response during this step. However, if they failed to mention a predetermined "target" word (e.g., "bone" for a dog image), they were encouraged to explain their reasoning. If necessary, the examiner introduced the "target" word and clarified its semantic or phonological relationship to the term. Participants went through two training rounds to ensure they understood the task. Instructions were reiterated until comprehension was confirmed.

For the phonological Free Association, the translated instructions were: "We're going to play a game together. I will show you some images, and each time, you'll tell me what the image represents. Additionally, you'll let me know if you can think of a word that sounds similar to it." For the semantical Free Association, the translated instructions were "We're going to play a game together. I will show you some images, and each time, you'll tell me what the image represents, and then you'll share the word that comes to mind when you think of the word represented by the image."

Forced Choice

This task involves matching a primary image with one of three secondary images based on either phonological or semantic relationships, depending on the session type. The task was divided into several steps for each set of images:

1. Image Recognition: The participant was first asked to identify what a specific image (the primer) represents. Incorrect answers were gently corrected.

2. Identification of Additional Images: Three more images were presented one at a time. For each of these images, the participant was tasked with identifying the represented item. Expected responses were praised, and incorrect ones were corrected.

3. Matching: Finally, the participant was required to choose the best match among three secondary images for the primary image/word, based on either phonological or semantic relationships, depending on the current session type. The three images consisted of: an **expected response**, which matched the primary word based on the current session type association (e.g., phonological association during the phonological task or semantic association during the semantic task); an **opposite response**, which matched the primary word based on the non-current session type association (e.g., semantic association during a phonological task or phonological association during a semantic task); and a **distractor**, which did not match the primary word on either phonological or semantic levels.

Table 3. Recapitulative table of the experimental design

Type of task		Material	Training item
Semantic	Free	Target word with image	1. Dog 2. Eggplant
	Forced	Target word with image, then three choices of words to be associated with three other images	
Phonological	Free	Target word with image	1. Baby 2. Eggplant
	Forced	Target word with image, then three choices of words to be associated with three other images	

Data availability

All data and the untranslated protocol are available at the following link:

https://osf.io/ud3be/?view_only=3cb257fa6b9048ce8c1795ea87eaaaab.

Statistical analysis

Statistical analyses were conducted using the Python (3.8) and Jamovi (2.3.28) softwares. The significance threshold was set at 0.05, and all tests were two-tailed. We compared the performance

of four groups of children on two treatment modalities (phonological and semantic) using two types of tasks: a Free Association task and a Forced-Choice task.

For the Free Association tasks, we measured the average number of different responses produced per word for each treatment modality. A Student's t-test was used to compare these averages. "Do not know" responses were analyzed using a Kruskal-Wallis test to compare frequencies across the four age groups, and a Mann-Whitney test was applied to identify specific differences.

For the forced-choice tasks, a Wilcoxon signed-rank test was used to compare performance between the phonological and semantic modalities at each age level, for expected responses. Finally, pairwise comparisons were performed using a Mann-Whitney test to examine specific differences between age groups for each task and each type of response, with an additional correction for multiple tests.

3. Results

There are four groups spanning four class years. These groups will be referred to as 5y.o., 6y.o., 7y.o., and 8y.o. in this section.

3.1. Free Association

In the following analyses, we will label the word most frequently provided by children in response to a given word as the primary associate (as seen in De La Haye, 2003; Ferrand & Alario, 1998).

General analysis

The word with the highest frequency in Manulex is 'maison', which translates to 'house'. For the semantic associates, the main ones were 'cabane', which translates to 'hut' (3 occurrences), 'caravane', which translates to 'motorhome' (2 occurrences), 'fenêtre', which translates to 'window' (2 occurrences), and 'rentrer', which translates to 'getting in' (2 occurrences). Its main phonological associates were 'maisonnette' /mɛ.zɔ.net/ with 5 occurrences, 'mai' /mɛ/ (2 occurrences), and 'mami' /ma.mi/ (2 occurrences).

On average, the number of different words produced for semantic associates was 17.06 ($M = 17$, $SD = 3.80$), while for phonological associates, the average was 19.91 ($M = 19.91$, $SD = 3.92$). A t-test indicated a significant difference in the variety of words produced between semantic and phonological Free Association, $t(298) = 6.9$, $p < .001$, with a medium to large effect size ($d = 0.73$).

"Do not know" answers across age groups

A Kruskal-Wallis H test was conducted to assess differences in Do Not Know (DNK) ratios across the four age groups (5, 6, 7, and 8 years) for the **semantic modality** (fig. 2). The results indicated a statistically significant effect of age, $\chi^2(3) = 10.510$, $p = .015$. Post hoc pairwise comparisons with Bonferroni correction revealed that 5y.o. had significantly higher DNK ratios compared to 8y.o., $U = 3018$, $p_{adj} = .043$. No significant differences were observed between other

age pairs: 5 vs. 6 years ($p_{adj} = 1$), 5 vs. 7 years ($p_{adj} = .065$), 6 vs. 7 years ($p_{adj} = .844$), 6 vs. 8 years ($p_{adj} = .613$), or 7 vs. 8 years ($p_{adj} = 1$).

A Kruskal-Wallis H test for the **phonological modality** showed a strong age effect, $\chi^2(3) = 59.921, p < .001$ (fig. 2). Bonferroni-corrected pairwise comparisons indicated that 5y.o. had significantly higher DNK ratios than 6y.o. ($U = 2720.50, p_{adj} < .001$), 7y.o. ($U = 3818.50, p_{adj} < .001$), and 8y.o. ($U = 3892, p_{adj} < .001$). No significant differences were found between 6- and 7y.o. ($p_{adj} = .877$), 6- and 8y.o. ($p_{adj} = .090$), or 7- and 8y.o. ($p_{adj} = 1$).

A Wilcoxon signed-rank test showed significantly higher DNK ratios in the phonological task than in the semantic task, $Z = 3051.500, p < .001$, with a large effect size ($r = 0.796$), meaning that children were more likely to produce DNK responses in the phonological task.

Spearman's rank-order correlation coefficients were calculated to assess the relationship between age (in months) and DNK ratios for both modalities. For the semantic modality, a significant negative correlation was found, $r_s = -.173, p = .008$, indicating that DNK ratios decreased as age increased, albeit the correlation was weak. In the phonological modality, a significant negative correlation was observed, $r_s = -.475, p < .001$, suggesting a moderate association between increasing age and decreasing DNK ratios.

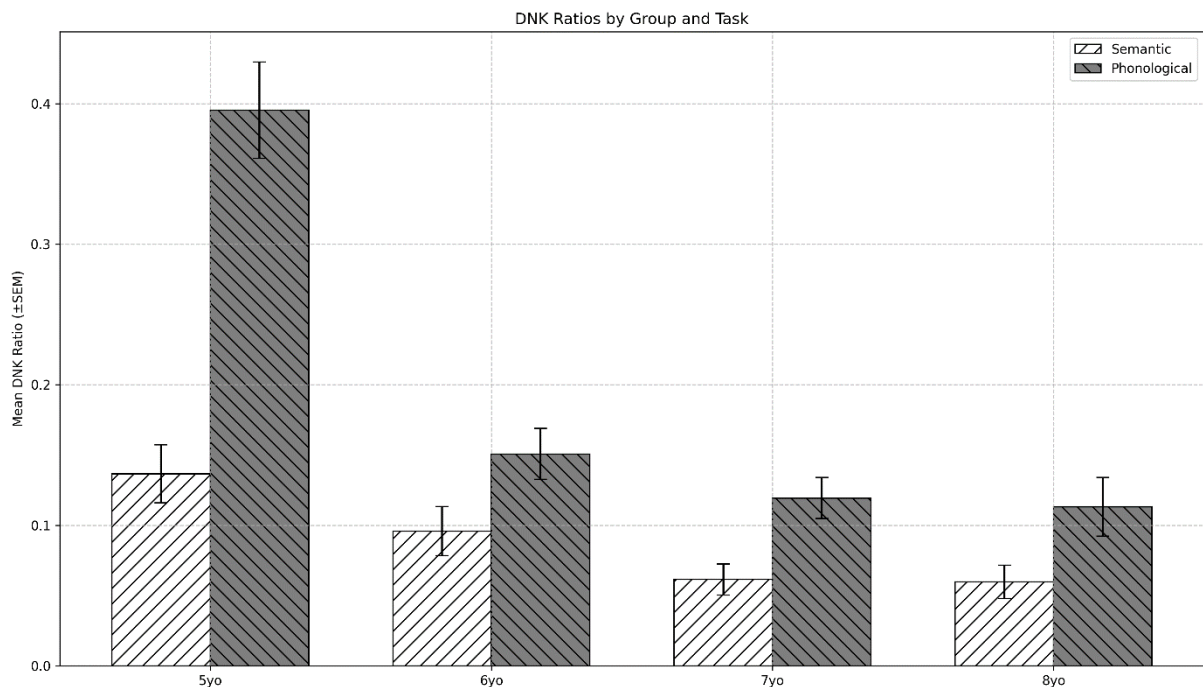


Figure 2. Analysis of DNK (“do not know”) Response Frequencies by Task Type Across Age Groups.

Note. The indicators of statistical significance are in the text below.

3.2. Forced Choice

The results are displayed in Figures 3a, 3b and 3c, which present the performance ratios computed as follows: the mean ratio of expected responses, the mean ratio of responses to the distractors and the mean ratio of opposite responses. "Opposite response" refers to occurrences where phonological associates were selected in a semantic task, and conversely, where semantic associates were selected in a phonological task.

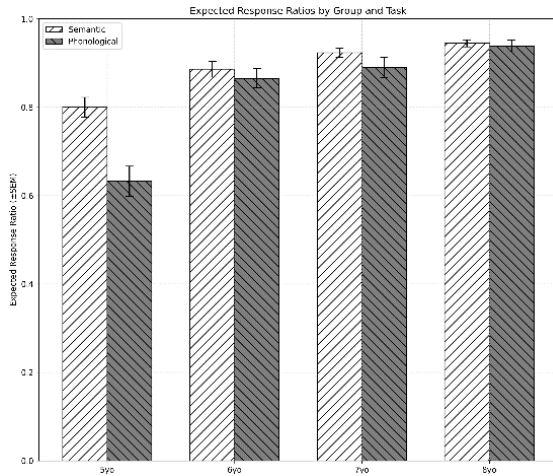


Fig 3.a

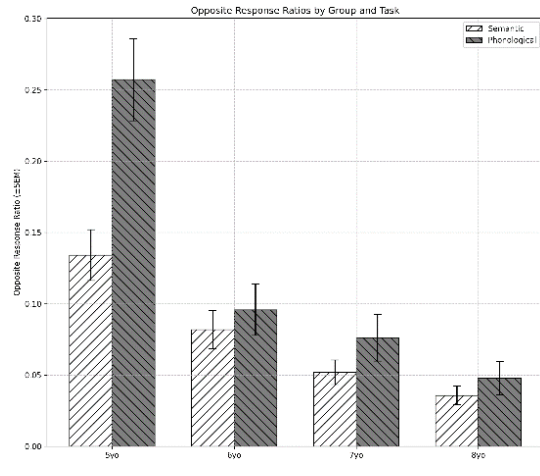


Fig 3.b

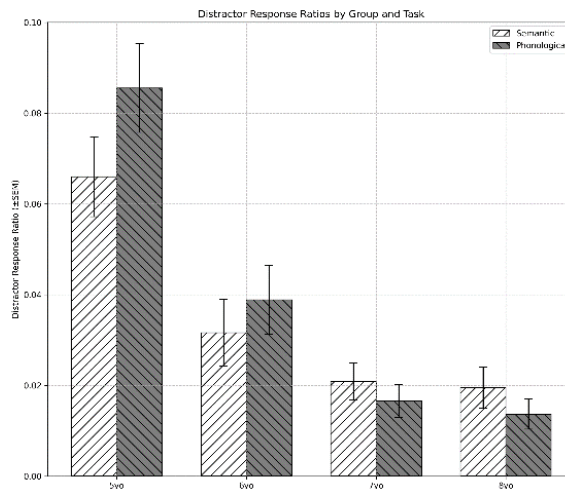


Fig 3.c

Figure 3.

Fig 3.a. Ratio of Performances for Task 2 (Expected Responses) for Semantic and Phonological Tasks for Each Age Group.

Fig 3.b. Ratio of Performances for Task 2 (Opposite Responses) for Semantic and Phonological Tasks for Each Age Group.

Fig 3.c. Ratio of Performances for Task 2 (Distractor Responses) for Semantic and Phonological Tasks for Each Age Group.

Note for all graphs. The indicators of statistical significance are in the text below.

Comparison of semantic and phonological tasks across age groups

For 5y.o., Wilcoxon signed-rank test showed significantly better performance on the semantic than phonological tasks for expected responses ($Z = 705.500, p < .001, r = 0.783$). For 6-, 7-, and 8y.o., no significant differences were found between tasks in expected responses (6y.o.: $Z = 235, p = .799, r = 0.403$; 7y.o.: $Z = 352.50, p = .60, r = 0.463$; 8y.o.: $Z = 411.500, p = .617, r = 0.545$).

Semantic task

A Kruskal-Wallis H test showed a significant age effect in **expected responses** for the semantic task, $H(3) = 40.834, p < .001$ (fig. 3a). Bonferroni-corrected comparisons indicated that 5y.o. had significantly lower expected response ratios than 6y.o. ($U = 1184, p_{adj} = .015$), 7y.o. ($U = 1233.500, p_{adj} < .001$), and 8y.o. ($U = 1033, p_{adj} < .001$). No significant differences were found among the older groups.

For **opposite responses** in the semantic task, the Kruskal-Wallis H test showed a significant age effect, $H(3) = 34.349, p < .001$ (fig. 3b). Pairwise comparisons indicated that 5y.o. had significantly higher opposite response ratios than 7y.o. ($U = 3414, p_{adj} < .001$) and 8y.o. ($U = 3494, p_{adj} < .001$), and 6y.o. differed significantly from 8y.o. ($U = 1596.50, p_{adj} = .046$). No other significant differences were found among age groups.

The Kruskal-Wallis H test for **distractor responses** in the semantic task was significant, $H(3) = 36.552, p < .001$ (fig. 3c). Pairwise comparisons showed that 5y.o. selected distractors more frequently than 6y.o. ($U = 2408, p_{adj} = .003$), 7y.o. ($U = 3479.50, p_{adj} < .001$), and 8y.o. ($U = 3427, p_{adj} < .001$). No significant differences were found between the older age groups.

Phonological task

For the phonological task, a Kruskal-Wallis H test for **expected responses** was significant, $H(3) = 67.441, p < .001$. Comparisons showed that 5y.o. had lower expected response ratios than 6y.o. ($U = 719, p_{adj} < .001$), 7y.o. ($U = 936.50, p_{adj} < .001$), and 8y.o. ($U = 662.50, p_{adj} < .001$). Six-year-olds performed worse than 8y.o. ($U = 603.50, p_{adj} = .013$). No significant differences were found between 6- and 7y.o., or between 7- and 8y.o..

For **opposite responses**, the Kruskal-Wallis H test was significant, $H(3) = 53.370, p < .001$. Comparisons indicated that 5y.o. had higher ratios than 6y.o. ($U = 1942, p_{adj} = .003$), 7y.o. ($U = 3601.50, p_{adj} < .001$), and 8y.o. ($U = 3762.50, p_{adj} < .001$). No significant differences were found between 6- and 7y.o., 6- and 8y.o., or between 7- and 8y.o..

For **distractor responses**, the Kruskal-Wallis H test was significant, $H(3) = 62.463, p < .001$. Comparisons revealed that 5y.o. selected distractors more frequently than 6y.o. ($U = 1921.50, p_{adj} = .004$), 7y.o. ($U = 3749.50, p_{adj} < .001$), and 8y.o. ($U = 3705.50, p_{adj} < .001$). 6y.o. differed from

7y.o. ($U = 1272$, $p_{adj} = .05$) and 8y.o. ($U = 1266$, $p_{adj} = .03$). No significant differences were found between 7- and 8y.o..

Correlation between age (in months) and response ratios

Spearman's rank-order correlations revealed a positive relationship between age (in months) and expected responses in both tasks. In the **Semantic task** ($r_s = .362$, $p < .001$) and the **Phonological task** ($r_s = .477$, $p < .001$), expected responses increased with age.

4. Discussion

This study presents an original database of semantic and phonological associative norms for 145 French words specifically selected for children aged from 5 to 8 years. It addresses a notable gap in psycholinguistics, particularly within the context of the French language. Unlike existing associative norm databases which focus on adults (Bonin et al., 2013; Ferrand & Alario, 1998) or on typical reading children (De La Haye, 2003), our study provides insights into the developmental trajectory of associative ability in non-reading French children. The inclusion of both Free Association and Forced Choice tasks provides a rich dataset that captures different aspects of children's associative abilities. By incorporating both semantic and phonological associations, our database offers a more comprehensive view of children's linguistic processing. This dual approach allows researchers to examine how different aspects of language representation develop in parallel or diverge over time, contributing to a more nuanced understanding of language acquisition. The Free Association task offers insights into the spontaneous generation of associations, while the Forced Choice task allows for a more controlled examination of children's ability to recognize and select appropriate associations. De La Haye's investigation revealed minimal variance in primary associations across three age cohorts (9, 10, and 11 years). Subsequently, our research norms supplement this dataset by incorporating younger participants (aged 5 to 9 years). Furthermore, our work contributes to the existing databases created by De La Haye (2003), Tarrago (2005), and Ferrand and Alario (1998).

Our results reveal significant variability in associative abilities among children, both across age groups and between semantic and phonological tasks. This variability provides valuable insights into the developmental trajectory of linguistic processing. In the Free Association task, 5-year-olds had significantly higher DNK ratios than 8-year-olds in the semantic modality, with no differences among intermediate ages. The phonological task was more challenging overall, especially for 5-year-olds, who had higher DNK ratios than older groups. Across both modalities, DNK responses decreased with age, especially in the phonological task. Overall, producing semantic associates was easier for all children compared to phonological associates, with the phonological task being particularly challenging for the 5-year-olds. This suggests a developmental progression in both semantic and phonological associative

abilities, with phonological associations developing more gradually. In the Forced Choice task, performance improved significantly with age, particularly in the phonological task, where the improvement was more gradual across different age groups. This may indicate that the ability to recognize and select appropriate semantic associations develops more quickly than that of phonological associations.

The study also revealed interesting differences between semantic and phonological processing. In Free Association, there were generally fewer different associates produced for semantic associates than for phonological ones, suggesting that the phonological network may be less structured in children, for all ages combined. This unexpected finding suggests that phonological networks might be more flexible or less constrained than semantic networks in early childhood.

Therefore, it seems that phonological and semantic networks develop along distinct timelines and through different processes in children. These challenges for younger children for phonological association may stem from their ongoing development of phonological awareness skills (Carroll et al., 2003; Duncan et al., 2006; see Moats & Tolman, 2009, for a comprehensive understanding of how they develop across age; see Anthony & Francis, 2005; Content et al., 2001, for studies of phonological awareness in French-speaking children). In children who have not yet started reading, phonological awareness is still developing. They are also learning how sounds in language correspond to the letters they see on a page. Compared to children who have started reading instruction, their phonological awareness skills are less developed, and they are in the early stages of acquiring phonemic awareness—a critical skill for reading. In France, educational practices emphasize developing phonological awareness first to facilitate later reading acquisition (*Programme D'enseignement de L'école Maternelle*, 2021). Thus, in our task, their selection of different word associations might have arisen because the words shared a single phoneme, which weakened the connection between them and potentially made the task overly challenging for young children.

Various studies suggest that phonological priming emerges earlier than semantic priming, indicating a developmental precedence of phonological organization (Arias-Trejo & Plunkett, 2009, 2013; Arias-Trejo et al., 2022; Borovsky & Peters, 2019; Chow et al., 2019; Delle Luche et al., 2014; Mani & Plunkett, 2010, 2011; Rämä et al., 2013; Styles & Plunkett, 2009). Initially, infants establish phonologically-based organization, which aligns with evidence suggesting early phonological influences on word processing in infants as young as six months old (Becker et al., 2014). As vocabulary expands, reliance on both phonological and semantic information increases over time (Arias-Trejo et al., 2022). This suggests that lexical organization likely begins with phonological foundations before incorporating semantic elements as vocabulary complexity grows. Our findings indicate that while older children initially show stronger explicit skills in semantic tasks than in phonological ones at the age of 5, their performance becomes similar in both tasks as they mature, with variability and

differences across age groups decreasing over time. However, variability diminishes earlier in development for semantic tasks. This indicates that the developmental trajectories for semantic and phonological explicit skills diverge slightly. These results highlight the value of the database as a crucial tool for research and for exploring both semantic and phonological associations in development.

Limits and future research

This database has significant potential for both research and educational applications. Researchers can use these norms to design age-appropriate psycholinguistic experiments, while educators and clinicians can leverage this information to develop targeted interventions for language development or assessment tools for language disorders. With regard to other types of psycholinguistic norms, it is important to consider aspects such as imageability (while being aware that our study may be influenced by the presence of additional images), age of acquisition, emotional valence, etc., like Bonin et al. (2013) did in their study. It could potentially provide a more comprehensive understanding of lexical organization. The age of acquisition affects word processing and recall, as words learned earlier are more accessible (Izura et al., 2011; Palmer & Havelka, 2010). Furthermore, emotional valence, i.e., the emotional tone of a word, also influences responses, with positive or negative words eliciting different reactions compared to neutral words (Barriga-Paulino et al., 2022). Recognizing these factors provides a nuanced interpretation of participants' responses and addresses gaps in current research.

Our study has limitations due to the small number of participants in certain age groups, affecting the statistical robustness and broader applicability of our findings. This insufficiency prevented us from comparing free association task answers across different ages, making it impossible to observe changes in lexical networks as children age.

During the administration of the tasks, we observed that some children perceived semantic relationships between the distractors and the proposed words, suggesting an interaction between the phonological and semantic networks (Levy et al., 2021). Furthermore, unexpected associations in the forced-choice task could be attributed to such unintended semantic associations. This observation led us to consider the need to modify the associations and ensure appropriate material for researchers, checking the strength of semantic and phonological links for each word to prevent unintended overlap.

Our results suggest that the phonological network in children appears less structured than the semantic network, with a greater diversity of associations produced. This difference could be specific to children or related to the phonological modality itself. To date, no study has explored this question in adults; such a study would help determine whether this diversity of associations is specific to children or independent of age.

References

- Anthony, J. L., & Francis, D. J. (2005). Development of phonological awareness. *Current Directions in Psychological Science*, 14(5), 255–259. <https://doi.org/10.1111/J.0963-7214.2005.00376.X/FORMAT/EPUB>
- Arias-Trejo, N., Angulo-Chavira, A. Q., Avila-Varela, D. S., Chua-Rodriguez, F., & Mani, N. (2022). Developmental changes in phonological and semantic priming effects in Spanish-speaking toddlers. *Developmental Psychology*, 58(2), 236–251. <https://doi.org/10.1037/dev0001290>
- Arias-Trejo, N., & Plunkett, K. (2009). Lexical-semantic priming effects during infancy. *Philosophical Transactions of the Royal Society of London: Series B, Biological Sciences*, 364(1536), 3633–3647. <https://doi.org/10.1098/rstb.2009.0146>
- Arias-Trejo, N., & Plunkett, K. (2013). What's in a link: Associative and taxonomic priming effects in the infant lexicon. *Cognition*, 128(2), 214–227. <https://doi.org/10.1016/j.cognition.2013.03.008>
- Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, 39(3), 445–459. <https://doi.org/10.3758/BF03193014>
- Barriga-Paulino, C. I., Guerreiro, M., Faisca, L., & Reis, A. (2022). Does emotional valence modulate word recognition? A behavioral study manipulating frequency and arousal. *Acta Psychologica*, 223, 103484. <https://doi.org/10.1016/J.ACTPSY.2021.103484>
- Bauer, P. J., & Mandler, J. M. (1989). Taxonomies and triads: Conceptual organization in one- to two-year-olds. *Cognitive Psychology*, 21(2), 156–184. [https://doi.org/10.1016/0010-0285\(89\)90006-6](https://doi.org/10.1016/0010-0285(89)90006-6)
- Becker, A. B. C., Schild, U., & Friedrich, C. K. (2014). ERP correlates of word onset priming in infants and young children. *Developmental Cognitive Neuroscience*, 9, 44–55. <https://doi.org/10.1016/j.dcn.2013.12.004>
- Blaye, A., & Bonthoux, F. (2001). Thematic and taxonomic relations in preschoolers: The development of flexibility in categorization choices. *British Journal of Developmental Psychology*, 19(3), 395–412. <https://doi.org/10.1348/026151001166173>
- Bonin, P., Méot, A., Ferrand, L., & Bugaïška, A. (2013). Normes d'associations verbales pour 520 mots concrets et étude de leurs relations avec d'autres variables psycholinguistiques. *L'Année Psychologique*, 113(1), 63–92. <https://doi.org/10.3917/anpsy.131.0063>
- Borovsky, A., & Peters, R. E. (2019). Vocabulary size and structure affects real-time lexical recognition in 18-month-olds. *PLoS ONE*, 14(7), e0219290. <https://doi.org/10.1371/journal.pone.0219290>

- Carroll, J. M., Snowling, M. J., Stevenson, J., & Hulme, C. (2003). The development of phonological awareness in preschool children. *Developmental Psychology, 39*(5), 913–923.
<https://doi.org/10.1037/0012-1649.39.5.913>
- Chow, J., Aimola Davies, A. M., Fuentes, L. J., & Plunkett, K. (2019). The vocabulary spurt predicts the emergence of backward semantic inhibition in 18-month-old toddlers. *Developmental Science, 22*(2), e12754. <https://doi.org/10.1111/desc.12754>
- Comesaña, M., Fraga, I., Moreira, A. J., Frade, C. S., & Soares, A. P. (2014). Free associate norms for 139 European Portuguese words for children from different age groups. *Behavior Research Methods, 46*(2), 564–574. <https://doi.org/10.3758/S13428-013-0388-0/FIGURES/3>
- Content, A., Kearns, R. K., & Frauenfelder, U. H. (2001). Boundaries versus Onsets in Syllabic Segmentation. *Journal of Memory and Language, 45*(2), 177–199.
<https://doi.org/10.1006/jmla.2000.2775>
- Cramer, P. (1968). Mediated priming of polysemous stimuli. *Journal of Experimental Psychology, 78*(1), 137–144. <https://doi.org/10.1037/h0026171>
- De La Haye, F. (2003). Normes d'associations verbales chez des enfants de 9, 10 et 11 ans et des adultes. *L'année Psychologique, 103*(1), 109–130. <https://doi.org/10.3406/psy.2003.29627>
- Delle Luche, C., Durrant, S., Floccia, C., & Plunkett, K. (2014). Implicit meaning in 18-month-old toddlers. *Developmental Science, 17*(6), 948–955. <https://doi.org/10.1111/desc.12164>
- Duncan, L. G., Colé, P., Seymour, P. H. K., & Magnan, A. (2006). Differing sequences of metaphonological development in French and English. *Journal of Child Language, 33*(2), 369–399. <https://doi.org/10.1017/S030500090600732X>
- Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*Power 3 : A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. <https://doi.org/10.3758/bf03193146>
- Fenson, L., Cameron, M. S., & Kennedy, M. (1988). Role of perceptual and conceptual similarity in category matching at age two years. *Child Development, 59*(4), 897.
<https://doi.org/10.2307/1130257>
- Fernandez, A., Diez, E., Alonso, M. A., & Beato, M. S. (2004). Free-association norms for the Spanish names of the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments, & Computers, 36*(3), 577–583. <https://doi.org/10.3758/BF03195604>
- Ferrand, L., & Alario, F.-X. (1998). Normes d'associations verbales pour 366 noms d'objets concrets. *L'année Psychologique, 98*(4), 659–709. <https://doi.org/10.3406/psy.1998.28564>
- Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., Augustinova, M., & Pallier, C. (2010). The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840

- pseudowords. *Behavior Research Methods*, 42(2), 488–496.
<https://doi.org/10.3758/BRM.42.2.488>
- Imai, M., Gentner, D., & Uchida, N. (1994). Children’s theories of word meaning: The role of shape similarity in early acquisition. *Cognitive Development*, 9(1), 45–75.
[https://doi.org/10.1016/0885-2014\(94\)90019-1](https://doi.org/10.1016/0885-2014(94)90019-1)
- Izura, C., Pérez, M. A., Agallou, E., Wright, V. C., Marín, J., Stadthagen-González, H., & Ellis, A. W. (2011). Age/order of acquisition effects and the cumulative learning of foreign words: A word training study. *Journal of Memory and Language*, 64(1), 32–58.
<https://doi.org/10.1016/J.JML.2010.09.002>
- Kiss, G. R., Armstrong, C., Milroy, R., & Piper, J. (1973). An associative thesaurus of English and its computer analysis. In A. J. Aitken & R. W. Bailey (Eds.), *The Computer and Literary Studies* (Edinburgh, pp. 153–165).
- Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). Manulex: A grade-level lexical database from French elementary-school readers. *Behavior Research Methods, Instruments, & Computers*, 36, 156-166.
- Levy, O., Kenett, Y. N., Oxenberg, O., Castro, N., De Deyne, S., Vitevitch, M. S., & Havlin, S. (2021). Unveiling the nature of interaction between semantics and phonology in lexical access based on multilayer networks. *Scientific Reports*, 11(1), 14479. <https://doi.org/10.1038/s41598-021-93925-y>
- Lucariello, J., Kyratzis, A., & Nelson, K. (1992). Taxonomic knowledge: What kind and when? *Child Development*, 63(4), 978–998. <https://doi.org/10.1111/J.1467-8624.1992.TB01676.X>
- Mani, N., & Plunkett, K. (2010). In the infant’s mind’s ear: Evidence for implicit naming in 18-month-olds. *Psychological Science*, 21(7), 908–913. <https://doi.org/10.1177/0956797610373371>
- Mani, N., & Plunkett, K. (2011). Phonological priming and cohort effects in toddlers. *Cognition*, 121(2), 196–206. <https://doi.org/10.1016/j.cognition.2011.06.013>
- Moats, L., & Tolman, C. (2009). Excerpted from *Language Essentials for Teachers of Reading and Spelling (LETRS): The Speech Sounds of English: Phonetics, Phonology, and Phoneme Awareness (Module 2)*. Boston: Sopris West.
<https://www.readingrockets.org/topics/developmental-milestones/articles/development-phonological-skills>
- Mon tout premier dictionnaire Larousse*. (2017).
- Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida word association, rhyme, and word fragment norms. <http://w3.usf.edu/FreeAssociation/>

- Nelson, D. L., McKinney, V. M., Gee, N. R., & Janczura, G. A. (1998). Interpreting the influence of implicitly activated memories on recall and recognition. *Psychological Review*, *105*(2), 299–324. <https://doi.org/10.1037/0033-295X.105.2.299>
- Palmer, S. D., & Havelka, J. (2010). Age of acquisition effects in vocabulary learning. *Acta Psychologica*, *135*(3), 310–315. <https://doi.org/10.1016/J.ACTPSY.2010.08.002>
- Programme d'enseignement de l'école maternelle*. (2021). Ministère de L'Education Nationale et de la Jeunesse. Retrieved October 17, 2024, from <https://eduscol.education.fr/83/j-enseigne-au-cycle-1>
- Quémart, P., Janiot, M., Casalis, S., Petrova, A., & Maïonchi-Pino, N. (2011). Développement de la reconnaissance des mots écrits et représentations mentales : l'approche de la psycholinguistique cognitive. In J.-M. Roy, B. Putois, N. Maionchi-Pin, F.-X. Pénicaud, N. Baul, & V. Chambon (Eds.), *Peut-on se passer de représentations en sciences cognitives ?* (pp. 47–57). De Boeck Supérieur. <https://doi.org/10.3917/dbu.putoi.2011.01.0047>
- Rämä, P., Sirri, L., & Serres, J. (2013). Development of lexical-semantic language system: N400 priming effect for spoken words in 18- and 24- month old children. *Brain and Language*, *125*(1), 1–10. <https://doi.org/10.1016/j.bandl.2013.01.009>
- Sell, M. A. (1992). The development of children's knowledge structures: events, slots, and taxonomies. *Journal of Child Language*, *19*(3), 659–676. <https://doi.org/10.1017/S0305000900011612>
- Styles, S. J., & Plunkett, K. (2009). How do infants build a semantic system? *Language and Cognition*, *1*(1), 1–24. <https://doi.org/10.1515/LANGCOG.2009.001>
- Tarrago, R., Martin, S., De La Haye, F., & Brouillet, D. (2005). Normes d'associations verbales chez des sujets âgés. *Revue Européenne de Psychologie Appliquée*, *55*(4), 245–253. <https://doi.org/10.1016/j.erap.2005.05.001>
- Vasanta, D., & Sailaja, P. (1999). Making sense of compound nouns: A study of word relatedness in Telugu. *Journal of Psycholinguistic Research*, *28*(4), 331–346. <https://doi.org/10.1023/A:1023284930839/METRICS>