The 10-Word Auditory Verbal Learning Test and Vocabulary Performance in 4-and 5-Year-Old Children

Kasperek, Ann; Kingma, Annet; de Aguiar, Vânia

Published in:
Journal of speech, language, and hearing research

DOI:
10.1044/2023_JSLHR-22-00706

ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

The 10-Word Auditory Verbal Learning Test and Vocabulary Performance in 4- and 5-Year-Old Children

Ann Kasperek,[a] Annet Kingma,[b] and Vânia de Aguiar[a,c]

[a] Center for Language and Cognition Groningen, University of Groningen, the Netherlands [b] Department of Pediatrics, University Medical Center Groningen, the Netherlands [c] Department of Radiation Oncology, University Medical Center Groningen, the Netherlands

ARTICLE INFO

Article History:
Received December 11, 2022
Revision received April 28, 2023
Accepted July 6, 2023

Editor-in-Chief: Julie A. Washington
Editor: Emily Lund

https://doi.org/10.1044/2023_JSLHR-22-00706

ABSTRACT

Background: Understanding the different factors that determine vocabulary development in young children is essential for the diagnosis and rehabilitation of language disorders in children. Language development is closely related to other cognitive processes such as auditory verbal learning and memory. This research focuses on the development of a novel auditory verbal learning test (AVLT) for 4- and 5-year-old children within the Dutch population. This new test is an adaptation of the common AVLT for both older children and adults, usually including a list of 15 words. Considering the lower attention span and limited executive functioning in young children, the word list of this new instrument is reduced to 10 words. Besides, a second recognition form has been developed to improve the ability to distinguish between possible underlying learning and memory deficits.
Method: Ninety-five preschool children (ages 4;0–5;12 [years;months]) were tested with this new AVLT 10-word test for kids (10WT-K), yielding different measures of verbal auditory memory. Forty-eight of 95 children received a recognition task with semantically unrelated items, and 47 of 95 received a recognition task with semantically related items. Three additional language skills were assessed to establish test validation: receptive and expressive vocabulary performance and nonword repetition. Outcome of the 10WT-K was related to scores on the language measures.
Results: Positive correlations were found between the total score of the 10WT-K and all three aforementioned language skills. We found no correlations between frequency of error types (intrusions and repetitions) and language measures. Furthermore, children who were administered the recognition list with semantically related items showed fewer correct answers and more false-positive and false-negative responses than children who received a recognition list with semantically unrelated items.
Conclusions: The 10WT-K for young children can be used to (a) measure different aspects of auditory verbal learning and memory, (b) clarify the nature of possible verbal learning difficulties, and (c) identify a possible nature of language disorders. The word recognition task tested with semantically related items provides a more accurate measurement of individual differences, namely, in distinguishing retrieval and storage abilities. The significant relation found between auditory verbal short-term memory capacity and vocabulary performance in preschool children is a first step toward establishing test validity.

Active word learning in children mostly seems an easy or even "automatic" process, although typically developing children show great variability in the speed and number of words they acquire during the preschool

Correspondence to Vânia de Aguiar : vania.de.aguiar@rug.nl. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

period (Klee & Harrison, 2001). Around the age of 2 years, the average amount of vocabulary knowledge corresponds to 200 words. Around the age of 3 years, the estimated average is 1,000 words, and at the end of the preschool period (5 years old), it is 3,000 words (Schaerlaekens, 2008). During the grades from kindergarten to Grade 2, the difference between children with small and large vocabularies continues to get larger, and this gap is seldomly eliminated by education alone (Stone, 2004).

The acquisition of receptive and expressive vocabulary, both in children and in adults, does not stand alone but shows interaction with other cognitive processes such as auditory verbal short-term memory (STM; Buchsbaum, 2016) but also long-term memory, controlled attention, and fluid intelligence (Gillam et al., 2021). Different assessments have been developed to specifically examine receptive and expressive vocabulary in young Dutch children, but these tests measure generally vocabulary size, rather than the learning process (Schlichting & Spelberg, 2012; Schlichting, 2004; Semel et al., 2012). Implementing verbal memory tests (both with known words and nonwords) in the assessment of language and developmental disorders can also provide a wealth of information on how familiar and novel verbal material is processed. By assessing this memory capacity before children start in elementary school, strengths and weaknesses in children's skills can be observed and intervention programs can be installed where needed. In appraising the literature on verbal memory assessment in young children, however, we found that assessment of word list learning, for the Dutch population, starts at the age of 6 or 7 years (Kemp, 2011; Kingma, 2021; Van den Burg & Kingma, 1999), with no norms available for younger children. This study addresses this gap. We developed and validated a novel auditory verbal learning instrument for 4- and 5-year-old children that may help clinicians in diagnosing and treating children with language impairment and/or deficits in verbal learning and memory.

From Memory to Word Learning in Typically Developing Children

Human memory can be described as an information-processing system where new memories are created based on three main mechanisms: encoding, storage, and retrieval. In modern cognitive psychology, *STM* is defined as "a cognitive faculty that enables humans and other complex animals to temporarily store, process and manipulate important pieces of information that are no longer readily available in the sensory environment" (Buchsbaum, 2016, p. 863). Whereas *phonological STM* refers to the simple storage of information in phonological form, it can be partially equated to the phonological loop described within Baddeley's working memory system (Baddeley, 2000), comprising four important elements. The phonological loop is responsible for storage and maintenance of information in a verbal form. This phonological loop consists of a phonological short-term store and a subvocal rehearsal process. Before 7 years of age, however, spontaneous rehearsal does not reliably occur. In younger children, the phonological loop therefore consists of the phonological store only (Gathercole et al., 2004). The phonological loop was initially termed the *articulatory loop* by Baddeley and Hitch (2019); later, they abandoned the term since they regarded articulation as the method of rehearsal rather than the storage medium itself. The visuospatial sketchpad is dedicated to the storage and maintenance of visuospatial information. The episodic buffer combines all verbal and visuospatial information into complex multimodal representations linking working memory to long-term memory (Repovs & Baddeley, 2006). The "central executive" is seen as a control system of limited attentional capacity that coordinates and controls these three slave systems.

Memories encoded can be translated onto the long-term memory system. This system has two components: Explicit/declarative memory refers to memories involving personal experiences (episodic memory) and factual information that we can consciously retrieve and intentionally articulate (semantic memory); implicit memory refers to perceptive and emotional unconscious memories that influence our behavior (Dew & Cabeza, 2011). Memory retrieval is the process of remembering information stored in long-term memory.

How does this memory system support word learning? Words are the core elements of language, represented in at least three components: meaning (or concept), lemma (grammatical characteristics), and word form (how the word sounds; e.g., Levelt et al., 1999). In order to learn a new word and assign a meaning to it, children have to complete at least the following three tasks: (a) perceive and segment a string of sounds (phonological form) from the stream of speech they hear, (b) hold this phonological form in STM while the lexicon is searched, and (c) identify a meaning and pair that to the new word form (Brackenbury & Pye, 2005). According to Martin and Gupta (2004) and Gupta and Tisdale (2009), the learning of phonological word forms depends on phonological STM and on the contribution of semantic memory. These links will be further detailed in the following paragraphs.

The first process named by Brackenbury and Pye (2005), perceiving and segmenting a string of sounds, relies mainly on processes of auditory perception and recognition of acoustic and prosodic patterns. These are critical to generate the input to Step 2 mentioned by the same authors. Here, the contribution of phonological STM to learning a new word is critical. As a string of sounds is heard, there is an automatic lexical search (Dell, 1997) during which the phonological information has to remain actively available in STM (Gupta & Tisdale, 2009; Martin & Gupta, 2004).

Importantly, forgetting of all or a part of the information kept in phonological STM during word learning means that the new lexical form cannot be stored correctly in the lexicon. In line with this, studies on the relation between vocabulary development and STM in preschool children between 4 and 6 years old (Majerus et al., 2006; Messer et al., 2010) have shown an important bidirectional influence between STM capacity and vocabulary development. A strong STM enables the child to form stable and precise long-term phonological representations. Conversely, as a child's vocabulary knowledge increases, the phonological representations of words become more fine-grained (more segmental rather than holistic in form), a process known as lexical restructuring (Metsala, 1999). This enhanced phonological awareness allows a more accurate encoding of new words in STM.

Continuing to Step 3 described by Brackenbury and Pie (2005), existing strings of sounds are identified as existing lexical (vocabulary) units and can be linked to their respective meanings, both of which are stored within semantic memory (Gupta & Tisdale, 2009; Martin & Gupta, 2004). However, strings of sounds that cannot be found in the phonological lexicon (novel words or nonwords) trigger the search for a meaning. Carey and Bartlett (1978) originally stated that the mapping process required to assign a heard word form to a meaning occurs in two phases: (a) the initial mapping of a linguistic label to a real-world entity (fast mapping) and (b) the subsequent retention and development of the initial representation (extended mapping; Vlach & Sandhofer, 2012). Fast mapping thus requires some kind of world knowledge (i.e., semantic memory), which will be linked to a new label. However, extended mapping enriches the semantic representation of the concept with critical distinctions between ordinate and subordinate entities, thus determining its correct category extension (Carey, 2010). Critically, in order to use either of these processes to learn words successfully, access to and retention in semantic memory are needed.

Given that novel words with their respective phonological and semantic information are stored in vocabulary in an interactive network (McClelland et al., 1995), the number of nouns a children already have in their vocabulary affects their ability in comprehension and expression tasks. The larger and more robust the network, the more easily new memories are encoded and retained. Walker and McGregor (2013) emphasize that the vocabulary size has a causal effect on learning new words: A good word learner builds a large vocabulary, which, in turn, enables additional learning. The concept of "lexical restructuring" (Metsala, 1999) provides a clear rationale for how the developmental refining of phonological knowledge may support learning. When children first encounter new words, they are represented in a holistic manner at the lexical level. As their vocabulary grows, children develop phonotactic knowledge, which makes them more sensitive to the phonologically finer grained structures within words, and they adopt the phoneme as the basic unit of oral language processing. Even in preschool children, who have no literacy experience, phonological representations become more accurate and more segmented as they develop (Ainsworth et al., 2016), leading to more durable representations in the phonological loop.

Assessment of Vocabulary Knowledge and Verbal Memory

Assessment of vocabulary knowledge in a clinical or educational setting is typically done through static assessments, where children name images (Semel et al., 2012) or point to images that correspond to verbally presented words (e.g., Schlichting, 2004). While these tools are important to measure the vocabulary size and can give a good indication of a child's capacity compared to peers, they give little information about the child's potential learning ability or about which aspects of learning may be difficult. Furthermore, such assessments are influenced by linguistic and cultural biases, being less suitable for socially diverse settings, second-language learners, and bilingual individuals (Hasson et al., 2013). A different approach is dynamic assessment. Here, a child is given the opportunity to learn a certain behavior during a learning phase of the assessment. Subsequently, and usually within the same session, the product of that learning is tested. Because the learning process, and not prior knowledge, is the assessed function, this type of examination is less affected by cultural or linguistic biases.

Verbal learning is, by definition, a dynamic concept and thus is examined by using dynamic assessment. To understand the nature of novel word learning, different approaches exist. Yu and Ballard (2007), for example, developed a unified model of early word learning, more specifically a computational model that is able to discover spoken words from continuous speech and associate them with their perceptually grounded meanings. In the field of linguistics and language development, verbal learning is often studied using tasks where a novel word (e.g., *pilking*) is presented within a specific experimental manipulation (e.g., Arunachalam & Waxman, 2011). Often, the focus is on examining what information children use to derive initial (fast mapped) word meanings (e.g., syntactic and semantic contexts) or what variables affect the success in learning a new word (e.g., number and frequency of exposure; Carey, 2010). Only recently did these types of tasks transfer to clinical practice within the Quick Interactive Language Screener (Levine et al., 2020).

The latest version of the Wechsler Intelligence Scale for Children–Fifth Edition (WISC-V; Wechsler, 2014) provides a symbol translation test aimed to detect difficulties in visual–verbal associative memory (visual–verbal pairs have to be learned), measured in an immediate, delayed, and recognition trial. This task differs from a word list learning task. As described in the WISC-V Technical and Interpretive Manual (Wechsler, 2014), visual–verbal associative memory tasks similar to the symbol translation subtests are closely associated with reading decoding skills, word reading accuracy and fluency, text reading, and reading comprehension. They can support the nature of diagnosing learning disabilities but are less appropriate in defining the nature of developmental language disorders.

Another strategy to examine verbal learning is via an auditory verbal learning test (AVLT), an approach initially developed by Rey (1941, 1958) and adapted widely in different forms and languages (Vakil et al., 2012; Van den Burg & Kingma, 1999, 2021). In the Dutch adaptation of the AVLT for children aged 6–13 years, children are presented with 15 existing nouns, over five learning trials (Van den Burg & Kingma, 1999, 2021). Each time, the child is asked to recall as many words as possible. Following a 25-min interval, the child is again asked to complete a delayed recall trial and, after that, a recognition trial. During the recognition trial, the words learned are presented along with distractors, and the child indicates whether or not those words were part of the learning set. This task is well known and widely used in the Netherlands both in clinical practice and in research.

Novel word learning and verbal list learning share several processes; both tasks require the perception of a phonological form and the storage of that form in the phonological STM. Furthermore, both tasks prompt a search in the lexical-semantic system. In the context of verbal list learning, the lexical search leads to reactivation of the existing lexeme, which corresponds to the concept, making it easier to activate (Paradis, 2009). In the case of novel word learning, a lexical search does not identify existing lexemes directly, although when presented along with existing concepts, a semantic representation can be retrieved. The new lexeme can then be mapped onto the existing concept (Green, 1998).

Implementing an auditory-verbal memory test in the diagnostic process of young children with language problems can lead to a better understanding of encoding, storage, and retrieval abilities necessary for word learning (Archibald, 2017; Jackson et al., 2021). The administration of an AVLT provides a wealth of information (Lezak, 1995): immediate word span under overload conditions (Trial 1), final acquisition level (Trial 5), total acquisition and learning curve (Trials 1–5), delayed recall after 20–30 min, recognition out of a list with distractor items, number of repetitions, and number and types of intrusions.

Several word list tasks have been designed and validated for younger children, including normative data for immediate and delayed recall and recognition. The Californian Verbal Learning Test for Children (CVLT-C; Delis et al., 1994) is an examination of auditory and verbal learning for children between the ages of 5 years and 16 years 11 months. Goodman et al. (1999) presented normative data for 4-year-old children on the CVLT-C. The Wide Range Assessment of Memory and Learning (Sheslow & Adams, 1990) also includes a verbal list learning task starting at the age of 5 years. The Children's Memory Scale (Cohen, 1997) assesses auditory-verbal learning from ages 5 through 16 years. The three forementioned list learning tests all make use of 15 words or more in their list learning task (Kreutzer et al., 2011). They were developed for a large age span, which makes the test potentially less suitable for the youngest age groups. By reducing the number of words for the 10-word test for kids (10WT-K), this task is better suited to young children who have a lower attention span, among other yet immature cognitive functions, compared to older subjects. We also choose not to administer an interference list, given the limited executive functioning abilities in young children.

In this study, in comparison to the standard AVLT for children and adults, a new recognition format is added. In neuropsychological assessment, the recognition score can be used to distinguish between retrieval and storage limitations. Children who encounter difficulties in storage generally show an insufficient score on both delayed recall and recognition, whereas children who exhibit retrieval difficulties show a better score on recognition than on delayed recall. Kok and Kingma (2009) administered the recognition trial as part of the Dutch AVLT for children aged 6–13 years in 90 participants. The recognition lists included 15 distractor words with low association apart from the 15 stimulus words of the AVLT. An important finding of this study was the ceiling effect: 89% of the children obtained a high to maximal score (≥ 28/30). The 11% of the children who obtained low scores (≤ 27/30) also had low delayed recall scores, indicating storage problems. The ceiling effect for the recognition trial raised the question if this score is a valid measure to detect retrieval difficulties. These findings prompt the need to develop a recognition test with higher discrimination value. Lezak (1995) emphasizes that use of words (and not digits for example) in AVLTs introduces the possibility to study effects of word properties on verbal learning. Following this suggestion, we added a second recognition trial to the 10WT-K with a high degree of semantic association between words that may provide a more valid approach than the current test.

An additional approach to assessment of storage and manipulation of phonological information is using unknown words, also known as nonwords. A task such as nonword repetition has shown high diagnostic accuracy to discriminate both monolinguals and bilinguals with developmental language disorder from typically developing children (Boerma et al., 2015). Nonword repetition ability is the ability to repeat possible but nonoccurring word forms and, just like list learning, is a good measure of phonological STM capacity (Coady & Evans, 2008), hence of an important aspect of the ability to learn new words. Furthermore, because nonwords are designed in such way that the phonological structures included in stimuli abide to the phonotactic characteristics of the language (e.g., Boerma et al., 2015), performance can be facilitated by this knowledge, just like in list learning tasks. Importantly, phonotactic knowledge is an important predictor of vocabulary in monolingual and bilingual children (Messer et al., 2010).

For nonwords and existing words, the knowledge about the phonological structure of the native language appears to influence STM performance (Majerus et al., 2006). However, nonword repetition tests also differ from word list tests. First, in nonword repetition tests, only the recall immediately after each individual item is examined, and children are exposed to the nonwords only once. Furthermore, lexical characteristics of the word other than phonotactic structure (e.g., lexical frequency and age of acquisition) and semantic properties (imageability and semantic category) do not play a role in performance on nonword repetition tests. This way, their potential contribution to verbal learning cannot be measured. In the CVLT-C (Delis et al., 1994), for example, several words of one semantic category are included in order to measure clustering strategies. In our own proposed 10-word test, distractors in the recognition test include semantically related words in order to test recognition in a more complex setting.

Research Questions and Predictions

In this study, we describe the development of a novel version of the existing Dutch AVLT for children aged 6–13 years. This novel 10WT-K, including 10 instead of 15 words, is aimed to study learning, memory, and recognition in 4- and 5-year-old children. With this verbal (list) learning test for young children, we aim to evaluate both linguistic and learning abilities. The initial learning provides a good index of phonological STM, while learning across five trials reflects initial learning and memory consolidation. By including a recognition test, possible deficits in storage and retrieval ability can be distinguished. Furthermore, we designed two versions of the recognition test, with and without semantically related distractors, in order to examine recognition within a more demanding setting, thereby aiming to avoid ceiling effects observed in the aforementioned Ducth AVLT for children aged 6–13 years (Kingma et al., 1993). The primary aim of this study is to investigate the concurrent validity of different measures of the 10WT-K with three language measures (receptive vocabulary, expressive vocabulary, and nonword repetition). We address the following questions:

1. Are quantitative measures of the 10WT-K related to expressive and receptive vocabulary knowledge and to nonword repetition in 4- and 5-year old children? Given the important role of phonological STM in the vocabulary development of typically developing children (Gathercole & Baddeley, 1989; Gupta & Tisdale, 2009), a significant correlation is to be expected between the total acquisition score and language measures.
2. Is the occurrence of certain error types (repetitions and intrusions) in the 10WT-K related to vocabulary knowledge in 4- and 5-year old children? We expect no correlations between these measures and language measures. Research by Van den Burg and Kingma (1999) revealed that the occurrence of certain error types (repetitions and intrusions) is not related to verbal learning and memory per se.

Furthermore, the two versions of the recognition task (with semantically related or unrelated distractors) are distinguished in order to improve the psychometric properties of this aspect of the 10WT-K, and thus the following question is addressed:

1. Is recognition with semantically related distractors more demanding than recognition with semantically unrelated distractors? We predict that the recognition in the presence of semantically related distractors is more difficult compared to the task with semantically unrelated distractors. This may provide a better measurement of recognition, avoiding ceiling effect reported in earlier studies.

Method

Participants

Ninety-five monolingual children, 47 boys and 48 girls, participated in this study (see Table 1 for characteristics). All subjects had no history of sensory deficits, language disorders, or other developmental disorders, as

Table 1. Participant characteristics.

List version AVLT	Age 4;0–4;6	Age 4;7–4;12	Age 5;0–5;6	Age 5;7–5;12
10A	12 (6 F, 6 M)	11 (6 F, 5 M)	12 (6 F, 6 M)	12 (6 F, 6 M)
10B	12 (6 F, 6 M)	12 (6 F, 6 M)	12 (6 F, 6 M)	12 (6 F, 6 M)
Total	24	23	24	24

Note. Age in years;months. AVLT = auditory verbal learning test; F = female; M = male.

confirmed by consulting the children's school information sheet. They were recruited in elementary schools in the province of East-Flanders, Belgium, by means of an information letter. Informed consent was given by all parents. The education level of the mother was used as reference for the socioeconomic background of the children (see Table 2). Parents were asked to fill out the Children's Communications Checklist-2–Dutch Version (CCC-2-NL) questionnaire (Geurts, 2007) by means of an online platform (Praktikon, n.d.). Eighty-two of 95 valid responses were received, five parents did not respond, and eight parents gave inconsistent responses to the CCC-2-NL questionnaire.

Materials

Two parallel versions of the 10WT-K were developed (Versions A and B) to exclude test–retest effects in a scheduled follow-up assessment. The two versions of the 10WT-K developed for this study are similar to the lists of Van den Burg and Kingma (1999, 2021), but the 10WT-K was specifically developed for this younger population. For further scientific information on the word lists, the authors should be contacted.

Word selection started with short-listing 141 words that children acquire before they start primary school from the following sources: *NCDI* (*MacArthur Communicative Development Inventories, Dutch version*; Zink & Lejaegere, 2002), *Taalachterstand en Taalverwerving* (Schlichting & de Koning, 1998), *Duizend-en-een-woorden* (Bacchine et al., 2005), and *Basislijst Amsterdamse Kleuters* (Mulder et al., 2009). The following criteria for word selection were applied: Both word lists should equally include both concrete and abstract nouns, words with short and long vowels and diphthongs, and words belonging to the following semantic categories: food, animals, nature, body parts, toys, places, persons, transport, abstract words, and items around the house. These criteria were also used in the development of the 15-word versions of Van den Burg and Kingma (1999, 2021). As the 10 words came from 10 different semantic categories, children could not benefit from semantic clustering strategies in the encoding, storage, and retrieval process.

Table 2. Education level of the mother.

Education level	Age 4;0–4;6	Age 4;7–4;12	Age 5;0–5;6	Age 5;7–5;12
Lower secondary education	0	0	2	1
Higher secondary education	7	7	4	8
Bachelor's or master's degree	16	16	18	15

Note. Age in years;months.

Furthermore, it was ensured that Lists A and B were matched in length in phonemes (Marian et al., 2012), frequency (Marian et al., 2012), phonological neighborhood (Marian et al., 2012), arousal (Moors et al., 2013), Age of Acquisition of a word (Moors et al., 2013), concreteness (Brysbaert et al., 2014), and semantic association (Grave et al., 2018). The means and standard deviations of the psycholinguistic variables in both word lists are presented in Table 3. Averages were compared by means of a t test. No significant differences were found between the two word lists for any of these variables ($p > .3$ for all comparisons).

In the recognition trial, the 10 words presented in the earlier five trials are presented again along with distractors. The child has to recognize the earlier words among the new distractors. One set of distractors consists of 10 words that are semantically similar to the 10 target words (List A/SR and List B/SR), while the other set of distractors consists of 10 words that are semantically different from the 10 target words (List A/NSR and List B/NSR). The words in the distractor lists are matched to the target learning items in both versions (A or B) in the following aspects: number of words with long vowels, number of words with diphthongs, length in phonemes, AoA, frequency, arousal, phonological neighborhood, and concreteness. Semantic relatedness between the set of distractors and the target words is similar across the A and B versions, as measured using semantic vectors (Grave et al., 2018; $p = .435$). Furthermore, between related and unrelated distractors, there is a significant difference in semantic relatedness: for both Lists A and B ($p = .000$).

Receptive vocabulary knowledge was assessed by means of the Receptive Vocabulary subtest of the WPPSI-III-NL (Wechsler Preschool and Primary Scale of Intelligence—Third Edition, in Dutch; Wechsler, 2010). The item "to pay" was removed as there were two possible answers

Table 3. Comparison of psycholinguistic variables between Word Lists A and B.

Word list	Length in phonemes *M* (*SD*)	Frequency *M* (*SD*)	Phonological neighborhood *M* (*SD*)	Arousal *M* (*SD*)	Age of Acquisition *M* (*SD*)	Concreteness *M* (*SD*)	Semantic association *M* (*SD*)
A	3.90 (1.20)	79.44 (145.55)	9.20 (8.30)	3.88 (1.00)	4.34 (0.63)	4.41 (0.84)	0.56 (0.04)
B	4.00 (1.15)	194.37 (430.80)	9.60 (11.84)	3.65 (0.94)	4.08 (0.49)	4.65 (0.33)	0.56 (0.05)

("money" and "a computer"); this led to confusion in our group of children. To measure expressive vocabulary, the Expressive Vocabulary subtest of the Clinical Evaluation of Language Fundamentals Preschool–Second Edition (CELF Preschool-2) was used (Semel et al., 2012). Nonword repetition ability was measured by means of the Proef Fonologisch Bewustzijn (PFB; Eelen, 2006). As a nonverbal task to fill in the interval between Trial 5 and the delayed recall trial 20–30 min after that, Beery's Developmental Test of Visual-Motor Integration (VMI) was administered (Beery et al., 2010).

Procedure

Data were collected in schools, during school hours in a quiet room with as little distractors as possible. Due to the coronavirus pandemic, a transparent screen was used between the examiner and the child. No face mask was worn as this could influence intelligibility. The children were selected at random from each class and were instructed not to speak about the tests on return to their class. Forty-seven children were tested with Version A of the 10WT-K, and 48 children were tested with Version B. The use of both versions was equally distributed between age and gender. We preserve the procedure and measures employed for the abovementioned Dutch AVLT for 6–13 years. First, the list of 10 nouns was presented over five learning trials, with immediate recall tested following each presentation. The words were presented by means of an audio recording with an interval of 1 s between the words. All answers were written down, including repetitions (correct words that are mentioned several times) and intrusions (words that are not in the target list).

After a short conversation with the child to make them feel at ease, the following instructions were given at the beginning of the task (Trial 1): "You are going to hear a list of words. Listen carefully because when I stop, you have to tell me as many words as you can remember. It doesn't matter in what order you say them, just tell me as many words as you can. This is a difficult task, so don't mind if you remember only a few words." The list was then readministered for Trials 2–5 with a second set of instructions presented before Trial 2: "You are going to hear the same list again. When I stop, you have to tell me as many words as you remember. You must also say the words you have mentioned before. It doesn't matter in what order you say them. Just say as many words as you can remember." Taking into account the age and attention span of the children, instructions were repeated once if needed. When the child asked if they already said a specific word, this was not considered a repetition.

Following the fifth trial, there was a 20- to 25-min interval in which two nonverbal tasks were administered. First, the VMI (Beery et al., 2010) was assessed. The VMI is a nonverbal assessment helping to identify deficits in visual perception, fine motor skills, and hand–eye coordination. The second nonverbal task was a "movement snack" (the child had to perform motor exercises that were depicted on cards, together with the examiner; TOP tot TEEN beweegspel, 2015). By adding this active task to our protocol, motivation and sustained attention were aided. Then, after this interval, and with no further presentations of the list, delayed recall was assessed. The child was not informed before that this task would occur. The following instruction was given: "A little while ago, I presented a series of words you had to remember. Can you tell me all the words that you still remember now?"

After the delayed recall, the recognition task was presented. The examiner read aloud 20 words (10 target, 10 distractors). The following instructions were given: "I am going to read aloud several words. If the word was in the list I read earlier, you can say 'yes,' if not, you can say 'no.' Listen very carefully because I will read the words only once." Following, three language tests were administered. The Receptive Vocabulary subtest of the WPPSI-III-NL measures an individual's ability to identify correct responses to spoken words, for instance, at a picture that represents the word spoken by the examiner. The Expressive Vocabulary subtest of the CELF Preschool-2-NL assesses the child's ability to name illustrations of people, objects, and actions. The PFB nonword repetition test consists of 10 nonwords that have to be repeated by the child. The words were randomly chosen from the standardization study of Belis et al. (1997), based on the Nonword Repetition Test—an adaptation of a test by Gathercole and Baddeley (1989) by De Jong (Department of Pedagogy, Vrije Universiteit Amsterdam; cited in Elen, 2004). The items consist of one-, two-, three-, and four-syllable words (Elen, 2004).

Scoring

Regarding the 10WT-K, the scoring procedure as described by Lezak and used by Kingma was adopted (Kingma, 2021; Van den Burg & Kingma, 1999). For each trial, all words were written down in the order in which they were recalled, including repetitions and intrusions. All correct answers received 1 point. The total score for each trial (Trials 1–5 and recall) and the sum of Scores 1–5 were written down on the score form. As for the recognition task, the scoring procedure as described by Kok and Kingma (2009) was followed by counting the number of correct answers (true positives and true negatives), the number of false negatives, and the number of false positives. The following scores were obtained:

- total score: the sum of all correct words a child recalled throughout the five trials;
- learning rate: the difference between the number of correct answers between Trial 5 and Trial 1;
- the number of words recalled after a 20-min interval;
- recognition score: the sum of correct answers in the recognition task (true positives and true negatives);
- recognition false-positive score: the number of false-positive answers, that is, words identified as present in the learned list, which were not presented; and
- recognition false-negative score: the number of false-negative answers, that is, words classified as absent in the learning list, when they were in fact presented.

The WPPSI-III-NL Receptive Vocabulary subtest, the CELF Preschool-2-NL Expressive Vocabulary subtest, and the PFB nonword repetition test were scored following the respective instruction manuals. In the nonword repetition task, phoneme substitutions that were consistent (and that were also noticed in spontaneous speech) were considered to be related to articulatory difficulties rather than difficulties with encoding. These substitutions were therefore not counted as errors.

Data Analysis

First, normality of data was examined by means of the Shapiro–Wilk test. Nonparametric tests were used to compare different measures in case of nonnormality. All analyses were done using IBM SPSS, Version 26.0 (IBM Corp., 2017). Spearman's rho correlations were used to answer the first and second research questions (correlations between quantitative and qualitative measures of the 10WT-K and the scores on the Receptive Vocabulary subtest of the WPPSI-III-NL, the Expressive Vocabulary subtest of the CELF Preschool-2, and the PFB nonword repetition task). To answer the third research question, the total score, the false positives, and the false negatives of the two groups (children tested with semantically related vs. semantically unrelated recognition trials) were calculated and compared by means of a Mann–Whitney *U* test.

Results

Descriptive Data

Figure 1 shows the distribution of the norm-referenced scores for the item "semantic development" in our participants, as evaluated by the CCC-2-NL. A high norm-referenced score corresponds to a higher rate of observed language difficulties by the parent (M = 10, SD = 3). In the group of respondents, the semantic development, as observed by their parents, can be described as "average" in 70 out of 82 children, "above average" in four out of 82 children, and "below average" in eight out of 82 children.

Table 4 gives an overview of the means, standard deviations, and medians of the relevant measures per participant group (receiving semantically related or semantically unrelated test versions) regarding the accuracy scores. Table 5 shows the error type data per participant group (intrusions and repetitions).

Research Question 1: Are Quantitative Measures of the New 10WT-K Related to Receptive and Expressive Vocabulary Knowledge and to Nonword Repetition in 4- and 5-Year-Old Children?

After applying a Bonferroni correction, α levels were established at .003 (.05/18). Results are presented in Table 6. Significant positive correlations were found between 10WT-K total score and CELF Preschool-2-NL

Figure 1. Distribution of norm-referenced scores: Children's Communications Checklist-2–Dutch Version (CCC-2-NL; Geurts, 2007), item semantic development.

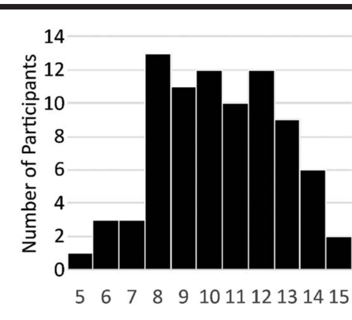


Table 4. Descriptive data of the 10-word test for kids relevant measures per participant group (scores).

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Total score	Recall score	Recognition score[a]	Expressive Vocabulary[b]	Receptive Vocabulary[c]	Nonword repetition[d]
List version	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**
List A/SR (*n* = 23)	3.09 (0.93)	4.30 (1.40)	4.87 (2.15)	5.52 (1.53)	5.78 (1.47)	23.57 (5.46)	4.65 (2.03)	8.87 (1.15)	26.52 (5.72)	26.87 (3.23)	5.52 (1.56)
List A/SNR (*n* = 24)	3.17 (1.18)	4.63 (1.25)	5.13 (1.42)	5.71 (1.79)	5.79 (1.94)	24.42 (6.14)	5.08 (1.50)	9.58 (0.57)	26.42 (4.43)	26.04 (3.72)	5.58 (1.11)
List B/SR (*n* = 24)	3.63 (1.32)	5.21 (1.47)	5.42 (1.96)	6.29 (1.51)	6.58 (1.61)	27.13 (6.30)	5.54 (1.61)	9.25 (0.97)	26.25 (5.23)	26.83 (3.89)	5.67 (1.14)
List B/SNR (*n* = 24)	3.29 (1.02)	4.38 (1.52)	5.29 (1.93)	5.92 (2.12)	6.58 (1.98)	25.46 (7.04)	5.54 (2.14)	9.38 (1.28)	26.33 (4.18)	26.29 (3.40)	5.46 (1.35)

Note. SR = semantically related; SNR = semantically nonrelated.

[a]Recognition score corresponds to true positives. [b]Expressive Vocabulary score from Clinical Evaluation of Language Fundamentals Preschool–Second Edition. [c]WPPSI-III-NL Receptive Vocabulary. [d]Proef Fonologisch Bewustzijn nonword repetition.

Table 5. Descriptive data of the 10-word test for kids relevant measures per participant group/list version (error types).

List version	Intrusions Trials 1–5	Intrusions Recall	Repetitions Trials 1–5	Repetitions Recall	False positives Recognition	False negatives Recognition
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
List A/SR (n = 23)	1.61 (3.54)	0.87 (1.23)	2.09 (2.24)	0.30 (0.62)	1.17 (1.52)	1.13 (1.15)
List A/SNR (n = 24)	1.71 (2.72)	0.58 (1.41)	1.88 (2.62)	0.21 (0.64)	0.50 (1.15)	0.42 (0.57)
List B/SR (n = 24)	1.67 (1.55)	0.50 (0.65)	2.54 (3.55)	0.58 (1.08)	0.79 (0.96)	0.75 (0.97)
List B/SNR (n = 24)	0.83 (1.07)	0.38 (0.75)	2.00 (3.38)	0.92 (2.71)	0.13 (0.44)	0.63 (1.28)

Note. SR = semantically related; SNR = semantically nonrelated.

Expressive Vocabulary, $r(94) = .517$, $p = .000$; between 10WT-K total score and WPPSI-III-NL Receptive Vocabulary, $r(94) = .322$, $p = .001$; and between 10WT-K total score and PFB nonword repetition, $r(94) = .302$, $p = .003$. In all cases, this indicates that higher scores for auditory verbal learning are associated with better receptive and expressive vocabulary capacities in 4- and 5-year-old children and a higher nonword repetition ability. Furthermore, we found significant positive correlations between 10WT-K recall score and PFB nonword repetition, $r(94) = .302$, $p = .003$, and between 10WT-K Recognition total score and PFB nonword repetition, $r(94) = .316$, $p = .002$, indicating the importance of phonological STM to support the encoding process in transferring information to long-term memory.

Research Question 2: Is the Occurrence of Certain Error Types (Repetitions and Intrusions) in the 10WT-K Related to Vocabulary Knowledge in 4- and 5-Year-Old Children?

After applying a Bonferroni correction, α levels were established at .004 (.05/12). Results are presented in Table 7. No significant correlations were found.

Research Question 3: Is Recognition With Semantically Related Distractors More Difficult and Therefore a More Valid Recognition Measure Than a Version With Semantically Unrelated Distractors?

The recognition total score of participants who received semantically unrelated words was significantly higher than the recognition total score of participants who received semantically related words, $U(n_{\text{unrelated}} = 48, n_{\text{related}} = 47) = 669.00$, $z = -3.60$, $p = .000$ (see Figure 2). Second, we found that the recognition false-positive ratio of participants who received semantically unrelated words was significantly lower than that of participants who received semantically related words, $U(n_{\text{unrelated}} = 48, n_{\text{related}} = 47) = 1{,}550.00$, $z = 3.75$, $p = .000$ (see Figure 3A). Finally, the recognition false-negative ratio of participants who received semantically unrelated words was lower than that of participants who received semantically related words, $U(n_{\text{unrelated}} = 48, n_{\text{related}} = 47) = 1{,}416.50$, $z = 2.40$, $p = .016$, as shown in Figure 3B. These results show that a recognition trial with semantically related distractors is more demanding compared to unrelated distractors in 4- and 5-year-old children.

Table 6. Correlations between quantitative measures of the 10-word test for kids (10WT-K), Clinical Evaluation of Language Fundamentals Preschool–Second Edition (CELF Preschool-2-NL) Expressive Vocabulary, Wechsler Preschool and Primary Scale of Intelligence—Third Edition (WPPSI-III-NL) Receptive Vocabulary, and Proef Fonologisch Bewustzijn (PFB) nonword repetition.

Measure	CELF Preschool-2-NL Expressive Vocabulary	WPPSI-III-NL Receptive Vocabulary	PFB nonword repetition
10WT-K total score	$r = .517^*$, $p = .000$	$r = .322^*$, $p = .001$	$r = .302^*$, $p = .003$
10WT-K learning rate	$r = .217$, $p = .035$	$r = .262$, $p = .010$	$r = .301$, $p = .003$
10WT-K recall	$r = .231$, $p = .024$	$r = .224$, $p = .023$	$r = .302^*$, $p = .003$
10WT-K recognition Total score	$r = .299^*$, $p = .003$	$r = .127$, $p = .219$	$r = .315^*$, $p = .002$
10WT-K recognition False positive score	$r = -.261$, $p = .011$	$r = -.189$, $p = .067$	$r = -.298$, $p = .003$
10WT-K recognition False negative score	$r = -.19$, $p = .064$	$r = .003$, $p = .970$	$r = -.214$, $p = .037$

*Correlation is significant at $p < .003$.

Table 7. Correlations between qualitative measures of the 10-word test for kids (10WT-K), Clinical Evaluation of Language Fundamentals Preschool–Second Edition (CELF Preschool-2-NL) Expressive Vocabulary, Wechsler Preschool and Primary Scale of Intelligence—Third Edition (WPPSI-III-NL) Receptive Vocabulary, and Proef Fonologisch Bewustzijn (PFB) nonword repetition.

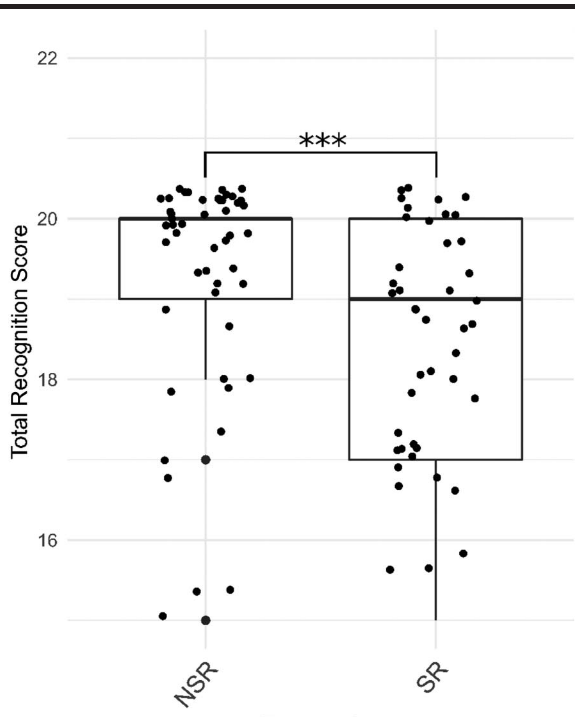
Measure	CELF Preschool-2-NL Expressive Vocabulary	WPPSI-III-NL Receptive Vocabulary	PFB nonword repetition
Intrusions Trials 1–5	$r = -.197$, $p = .055$	$r = -.162$, $p = .116$	$r = -.040$, $p = .698$
Intrusions recall	$r = -.089$, $p = .393$	$r = -.037$, $p = .721$	$r = -.117$, $p = .257$
Repetitions Trials 1–5	$r = -.120$, $p = .246$	$r = -.239$, $p = .019$	$r = -.063$, $p = .546$
Repetitions recall	$r = .142$, $p = .170$	$r = .158$, $p = .129$	$r = -.123$, $p = .237$

Note. The p values in the table are uncorrected. An adjusted alpha level was used (.004).

Discussion

Using an auditory verbal learning task specifically developed for 4- and 5-year-old children, significant correlations were found between auditory verbal learning and three language measures: receptive vocabulary, expressive vocabulary, and nonword repetition. Different from already existing verbal learning tests such as the CVLT-C (Delis et al., 1994), all 10 words in our parallel tasks came from a different semantic category, diminishing the opportunity to use semantic clustering strategies. This makes it possible to examine the child's learning ability in this auditory-verbal learning task with less possibility for the development of strategies that rely on other aspects of cognition. Furthermore, implementing a recognition trial with semantically related distractors can facilitate the detection of individual differences in this group of children. In the following paragraphs, we will further discuss the findings of our research, limitations, and future directions.

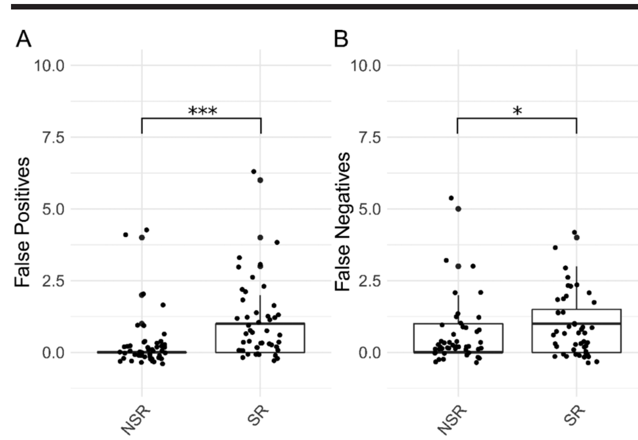
Figure 2. Distribution of recognition total score between list versions. NSR = recognition list with semantically nonrelated items; SR = recognition list with semantically related items.



Relation Between Auditory-Verbal Learning/ Memory and Vocabulary Achievement

The results of the first research question reveal significant positive correlations between receptive and expressive vocabulary knowledge and the total acquisition score on the verbal learning task, which is a measure of initial learning and consolidation. In the case of verbal list learning, we mainly measure the capacity of phonological STM. However, given the supraspan nature of the 10WT-K, attentional control provided by the working memory central executive is also activated (McCabe et al., 2010). We found no correlations between vocabulary knowledge and the score on the delayed recall task.

Figure 3. Distribution of recognition false positives and false negatives between list versions. NSR = recognition list with semantically nonrelated items; SR = recognition list with semantically related items.



The relationship we found between the auditory verbal learning process and vocabulary knowledge is known to be bidirectional. The better a new word can be rehearsed in phonological STM, the higher the probability that it will be transformed into a stable and precise long-term phonological representation (Gathercole & Baddeley, 1989). On the other hand, vocabulary growth itself is also a causal factor responsible for the developmental increase in verbal STM capacity (Martin, 2009; Metsala, 1999).

In young developing children, results of this study also show a positive correlation between nonword repetition (PFB score) and the total acquisition score of the 10WT-K reflecting the importance of phonological STM as a language learning device. Especially in young children, this phonological component is of great importance in developing their vocabulary (Baddeley, 2010). As the lexical representation of words in young children develop from a more holistic form to a more segmented form (lexical restructuring), phonological STM gains accuracy. Second, as children grow older, auditory verbal information becomes more semantically encoded because of the growth of their semantic network; they can associate new verbal material to already existing knowledge. Therefore, older children will be able to retain more auditory verbal material in their STM (Njiokiktjien, 2020). This growth in STM capacity is reflected in our descriptive data of the 10WT-K average total acquisition score, which shows an increase between 4- and 5-year-olds (from 20.92 to 29.29). The findings of correlations between the 10WT-K and the three aforementioned language measures support the construct validity of this new test to evaluate word learning mechanisms important for vocabulary development.

The positive correlation found between the ability of a child to repeat nonwords and the total score on the 10WT-K recall emphasizes that phonological STM capacity is important to support the encoding process in order to recall information correctly after a short period of time. Children who obtained a high score on the recognition task also performed better at repeating nonwords.

The Role of Error Types

The second research question focused on the occurrence of errors, more specifically intrusions and repetitions. No correlations were found between these error types and scores on expressive or receptive vocabulary. There is thus no evidence that these measures reflect processes important for language development, as previously shown by Van den Burg and Kingma (1999). Rather, these parameters are often interesting to study executive functioning (Kittler et al., 2006; Lezak, 1995). If it is indeed the case that intrusions and repetitions occur more frequently in children with poor executive abilities, the normative data of the 10WT-K can be helpful in the differential diagnosis of this clinical group.

An important remark to make here is that the occurrence of repetitions might be, to a certain extent, generated by the instructions given, as children heard the following message: “You must also say the words you have mentioned before.” Although this statement refers to words recalled in previous trials, this time reference may be unclear to some children. This is because the abstract sense of time starts to develop as from the age of 4–5 years and continues to develop until the age of 11 years (Njiokiktjien, 2020). Hence, some children may have interpreted this to mean that they could say words mentioned earlier during the same trial.

Recognition Trial: Toward a More Valid Approach

When we encode known verbal information (such as the 10 words in our word list), we predominantly encode the meaning/concept of the word (Cremona et al., 2020; Jefferies et al., 2006; Meltzer et al., 2016). This concept is defined by its relationship to other concepts in the semantic network, which give the concept its meaning (Vingerhoets & Lannoo, 1998). Difficulties between encoding and retrieval of these concepts can be observed by means of the recognition trial of an AVLT.

In the current Dutch AVLT for 6–13 years of age (Van den Burg & Kingma, 1999, 2021), children often perform at ceiling level on the recognition trial, making it difficult to discern between encoding and retrieval difficulties. Therefore, in this study, a different approach was explored by comparing two types of recognition tests: one list with items that are semantically related to the target words and one list with items that have no semantic relationship with the target words. Comparing the results of both lists revealed significant differences as well in total score (sum of correct answers) as in false-positive and false-negative ratio. Children who received a recognition task with semantically related items obtained a lower number of correct answers and showed more frequent false-positive and false-negative answers than children who received a recognition task with semantically unrelated items. As observed in the abovementioned Dutch AVLT for 6–13 years of age (Van den Burg & Kingma, 1999, 2021), most scores in the version with semantically unrelated items reached ceiling level. These results indicate that when a word has to be selected among semantically related distractors, there is greater activation of shared semantic features, posing greater demands.

By introducing a recognition task that is more demanding for children, a greater variability was seen in test results, making it possible to detect individual differences. Future studies should address the question if validity is higher with this approach. Also, it can be interesting to observe if these differences in performance between both recognition trials also exist in older children, between 6 and 12 years old.

When reviewing similar recognition tasks used for memory assessment in children such as the CVLT-C (Delis et al., 1994; Fahrer & Drozdick, 2020; Goodman et al., 1999), no ceiling effects in recognition tasks are reported in younger age groups. A possible explanation is that the recognition trial of the CVLT-C is more demanding due to the use of an interference list (List B). Hence, the total amount of words in their recognition list is longer compared to our recognition task. Furthermore, by calculating a "total recognition discriminability score," they combine both hits and false-positive scores, leading to a more accurate measure to examine whether the examinee could distinguish between target and nontarget words from List A (Farrer & Drozdick, 2020).

Limitations

Limitations to our study can be found in assessing the youngest group of children (4;0–4;6 [years;months]). As we administered the 10WT-K by means of an audio recording, some of these children experienced difficulties in auditory discrimination (e.g., park/ark/hark). Observing a speaker's articulatory patterns substantially improves the intelligibility of spoken speech; multisensory integration of speech inputs is important to effective communication (Ross et al., 2011). Although the use of an audio recording is necessary in terms of standardization, the option of reading the words out loud can be considered as this is a more natural way to present the words.

Second, a low sustained and/or focused attention sometimes interfered with performance throughout the group. Controlling for this factor in our group of participants could be an option, although attention is difficult to measure in this young age and would have to rely on questionnaires.

Furthermore, a remark can be made regarding the use of the receptive vocabulary items of the WPPSI-III-NL. We noticed that certain items are outdated, leading to errors that may be acceptable. For example, when a child has to select the target word "to pay," several children pointed at a person facing the computer, which is also a possible way to pay. We have removed this item in our test sample. As an alternative, we could choose a receptive and expressive vocabulary task that belongs to the same assessment battery, for example, the Schlichting Test for Language Reception and Production (Schlichting & Spelberg, 2012).

We note that the use of Bonferroni corrections may increase the possibility of beta error. Therefore, we advise caution in interpreting nonsignificant results.

Implications and Future Directions

First, this study contributes to the pediatric neuropsychological practice by developing a feasible auditory verbal learning task for young children. The vast majority of the participants understood the purpose of the task and cooperated well. Reliable results could be obtained in almost all children. Only one participant was excluded from the study due to a recently detected hearing disorder. Furthermore, this test could be used in older children who cannot pass the standard Dutch AVLT for children. Particularly in 6-year-old children, a list including 15 words occasionally seems too long. Although no normative data were acquired in 6 years olds part of the 10WT-K data can be used in a relative way in slightly older children, such as slope of the learning curve, difference between the score on learning Trial 5 and delayed recall, and the difference between the total score and the recognition score. Such relative use of normative data can also be considered in children with attention disorders or lower verbal intelligence.

This tool can be helpful in evaluating children with different neurological diseases, developmental disorders, or learning disabilities who often show deficits in auditory-verbal learning (Archibald & Gathercole, 2006; Kingma et al., 1993; Nagel et al., 2006; Riccio et al., 2007). By characterizing difficulties in phonological STM in children with language difficulties at a young age, more efficient manners of instruction can be implemented in educating these children.

Not only in clinical populations, but also in typically developing children, the 10WT-K can contribute to understanding the relation between language and memory functions. Longitudinal monitoring of the children in this study could shed light on the predictive value of the present 10WT-K regarding school results that are related to language skills at a later age (e.g., reading comprehension, critical listening, or spelling). Over the past years, early childhood education is increasingly faced with children speaking multiple languages, resulting in new challenges to ensure their inclusion and learning opportunities in the classroom (Langeloo et al., 2021). Future research could explore if this 10WT-K can be implemented in the longitudinal, dynamic assessment of multilingual children (Hasson et al., 2013).

Construct validity was established by correlating the total acquisition score with relevant language measures such as receptive and expressive vocabulary. Furthermore, by developing a recognition trial that is more demanding for children, differentiation between encoding and retrieval deficits will be easier to detect compared to when using existing instruments. Of course, in order to use this test in clinical settings, further research on validity and reliability is necessary.

Conclusions

By developing a 10-word AVLT for young children, an important contribution is made to the assessment of children who present with emergent language and/or learning problems. Language problems are often associated with academic, socioemotional, and psychiatric problems (Beeghly, 2006); therefore, there is a pressing need for early and accurate identification within this age group. Furthermore, the implementation of a recognition trial with semantically related distractors appeared to be more demanding for children of this age, leading to a potentially more valid approach compared to a recognition trial with semantically unrelated items only.

Author Contributions

Ann Kasperek: Conceptualization (Equal), Methodology (Equal), Formal analysis (Equal), Investigation (Lead), Data curation (Equal), Writing – Original draft (Lead), Visualization (Equal), Project Administration (Equal). **Vânia de Aguiar:** Conceptualization (Equal), Methodology (Equal), Formal analysis (Equal), Resources (Equal), Data curation (Equal), Writing – Review & Editing (Equal), Visualization (Equal), Supervision (Lead), Project Administration (Equal), Funding acquisition (Lead). **Annet Kingma:** Conceptualization (Equal), Methodology (Equal), Resources (Equal), Writing – Review & Editing (Equal), Visualization (Supporting), Supervision (Supporting).

Data Availability Statement

Data supporting the results of this research are available upon request from the authors.

Acknowledgments

This publication is part of the project Verb Processing and Verb Learning in Children With Paediatric Posterior Fossa Tumours (with project number VI.Vidi.201.003 of the research program NWO-Talentprogramma Vidi SGW 2020 which is financed by the Dutch Research Council [NWO]). This project is also supported by a research grant provided by Stichting Kinderneuropsychologie Noord Nederland for the VLYC project (Verbal Learning in Young Children). The authors would like to thank all children and their parents for their enthusiastic participation in this study. Special thanks go out to the teachers and coordinators of the schools for their willingness to gather all information and also for their genuine interest in the topic of this research and the theoretical and practical implications of it.

References

Ainsworth, S., Welbourne, S., & Hesketh, A. (2016). Lexical restructuring in preliterate children: Evidence from novel measures of phonological representation. *Applied Psycholinguistics, 37*(4), 997–1023. https://doi.org/10.1017/S0142716415000338

Archibald, L. M. (2017). Working memory and language learning: A review. *Child Language Teaching and Therapy, 33*(1), 5–17. https://doi.org/10.1177/0265659016654206

Archibald, L. M., & Gathercole, S. E. (2006). Short-term and working memory in specific language impairment. *International Journal of Language and Communication Disorders, 41*(6), 675–693. https://doi.org/10.1080/13682820500442602

Arunachalam, S., & Waxman, S. R. (2011). Grammatical form and semantic context in verb learning. *Language Learning and Development, 7*(3), 169–184. https://doi.org/10.1080/15475441.2011.573760

Bacchine, S., Boland, T., Hulsbeek, M., Pot, H., & Smits, M. (2005). *Duizend-en-een woorden* [A thousand and one words]. Stichting Leerplanontwikkeling.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417–423. https://doi.org/10.1016/s1364-6613(00)01538-2

Baddeley, A. (2010). Working memory. *Current Biology, 20*(4), R136–R140. https://doi.org/10.1016/j.cub.2009.12.014

Baddeley, A., & Hitch, G. J. (2019). The phonological loop as a buffer store: An update. *Cortex, 112,* 91–106. https://doi.org/10.1016/j.cortex.2018.05.015

Beeghly, M. (2006). Translational research on early language development: Current challenges and future directions. *Development and Psychopathology, 18*(3), 737–757. https://doi.org/10.1017/s0954579406060366

Beery, K. E., Buktenica, N. A., & Beery, N. A. (2010). *Developmental Test of Visual-Motor Integration–Sixth Edition*. Pearson.

Belis, A., Verbeeck K., Verloo, E. (1997). *Aanzet tot normering en vergelijkende studie van de Nonwoord repetitie test voor Antwerpse kinderen uit het tweede leerjaar* [Initiation to standardization and comparative study of the Nonword Rehearsal Test for Antwerp children from group 4] [Unpublished dissertation]. Katholieke Vlaamse Hogeschool, Afdeling Logopedie & Audiologie, Antwerpen.

Boerma, T., Chiat, S., Leseman, P., Timmermeister, M., Wijnen, F., & Blom, E. (2015). A quasi-universal nonword repetition task as a diagnostic tool for bilingual children learning Dutch as a second language. *Journal of Speech, Language, and Hearing Research, 58*(6), 1747–1760. https://doi.org/10.1044/2015_JSLHR-L-15-0058

Brackenbury, T., & Pye, C. (2005). Semantic deficits in children with language impairments: Issues for clinical assessment. *Language, Speech, and Hearing Services in Schools, 36*(1), 5–16. https://doi.org/10.1044/0161-1461(2005/002)

Brysbaert, M., Stevens, M., De Deyne, S., Voorspoels, W., & Storms, G. (2014). Norms of age of acquisition and concreteness for 30,000 Dutch words. *Acta Psychologica, 150,* 80–84. https://doi.org/10.1016/j.actpsy.2014.04.010

Buchsbaum, B. R. (2016). Chapter 69. Working memory and language. In G. Hickok & S. L. Small (Eds.), *Neurobiology of language* (pp. 863–875). Academic Press. https://doi.org/10.1016/b978-0-12-407794-2.00069-9

Carey, S. (2010). Beyond fast mapping. *Language Learning and Development, 6*(3), 184–205. https://doi.org/10.1080/15475441.2010.484379

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development, 15,* 17–29.

Coady, J. A., & Evans, J. L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *International Journal of Language & Communication Disorders, 43*(1), 1–40. https://doi.org/10.1080/13682820601116485

Cohen, M. J. (1997). *Children's Memory Scale: Administration manual*. The Psychological Corporation.

Cremona, S., Jobard, G., Zago, L., & Mellet, E. (2020). Word meaning contributes to free recall performance in supraspan verbal list-learning tests. *Frontiers in Psychology, 11,* Article 2043. https://doi.org/10.3389/fpsyg.2020.02043

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1994). *California Verbal Learning Test–Children's Version (CVLT-C)*. The Psychological Corporation.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review, 104*(4), 801–838. https://doi.org/10.1037/0033-295X.104.4.801

Dew, I. T. Z., & Cabeza, R. (2011). The porous boundaries between explicit and implicit memory: Behavioral and neural evidence. *Academy of Sciences, 1224*(1), 174–190. https://doi.org/10.1111/j.1749-6632.2010.05946.x

Eelen, R. (2006). *Proef Fonologisch Bewustzijn* [Assessment of Phonological Awareness]. VVL.

Elen, R. (2004). Ontwikkeling van het fonologisch bewustzijn bij kinderen (3;0–10;0) [Development of phonological awareness in children (3;0–10;0)]. *Stem-, Spraak-en Taalpathologie, 12*(1).

Farrer, T. J., & Drozdick, L. W. (2020). *Essentials of the California Verbal Learning Test: CVLT-C, CVLT-2, and CVLT3*. Wiley.

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language, 28*(2), 200–213. https://doi.org/10.1016/0749-596x(89)90044-2

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*(2), 177–190. https://doi.org/10.1037/0012-1649.40.2.177

Geurts, H. M. (2007). *CCC-2-NL: Children's Communications Checklist-2–Dutch Version*. Harcourt Test Publishers.

Gillam, R. B., Serang, S., Montgomery, J. W., & Evans, J. L. (2021). Cognitive processes related to memory capacity explain nearly all of the variance in language test performance in school-age children with and without developmental language disorder. *Frontiers in Psychology, 12,* Article 724356. https://doi.org/10.3389/fpsyg.2021.724356

Goodman, A. M., Delis, D. C., & Mattson, S. N. (1999). Normative data for 4-year-old children on the California Verbal Learning Test–Children's Version. *Clinical Neuropsychology, 13*(3), 274–282. https://doi.org/10.1076/clin.13.3.274.1748

Grave, E., Bojanowski, P., Gupta, P., Joulin, A., & Mikolov, T. (2018). Learning word vectors for 157 languages. *Proceedings of the International Conference on Language Resources and Evaluation (LREC 2018)*. https://fasttext.cc/docs/en/crawl-vectors.html

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*(2), 67–81. https://doi.org/10.1017/S1366728998000133

Gupta, P., & Tisdale, J. (2009). Does phonological short-term memory causally determine vocabulary learning? Toward a computational resolution of the debate. *Journal of Memory and Language, 61*(4), 481–502. https://doi.org/10.1016/j.jml.2009.08.001

Hasson, N., Camilleri, B., Jones, C., Smith, J., & Dodd, B. (2013). Discriminating disorder from difference using dynamic assessment with bilingual children. *Child Language Teaching and Therapy, 29*(1), 57–75. https://doi-org.proxy-ub.rug.nl/10.1177%2F0265659012459526

IBM Corp. (2017). *IBM SPSS Statistics for Windows, Version 26.0*.

Jackson, E., Leitão, S., Claessen, M., & Boyes, M. (2021). Word learning and verbal working memory in children with developmental language disorder. *Autism & Developmental Language Impairments, 6*. https://doi.org/10.1177/23969415211004109

Jefferies, E., Frankish, C. R., & Lambon Ralph, M. A. (2006). Lexical and semantic binding in verbal short-term memory. *Journal of Memory and Language, 54*(1), 81–98. https://doi.org/psycnet.apa.org/doi/10.1016/j.jml.2005.08.001

Kemp, S. L. (2011). NEPSY-ii. In J. S. Kreutzer, J. DeLuca, & B. Caplan (Eds.), *Encyclopedia of clinical neuropsychology*. Springer. https://doi.org/10.1007/978-0-387-79948-3_1575

Kingma, A. (2021). *Manual for the administration of the 15 words test for children*. Stichting Kinderneuropsychologie Noord-Nederland.

Kingma, A., Mooyaart, E. L., Kamps, W. A., Nieuwenhuizen, P., & Wilmink, J. T. (1993). Magnetic resonance imaging of the brain and neuropsychological evaluation in children treated for acute lymphoblastic leukemia at a young age. *The American Journal of Pediatric Hematology/Oncology, 15*(2), 231–238. https://doi.org/10.1097/00043426-199305000-00012

Kittler, P., Krinsky-McHale, S. J., & Devenny, D. A. (2006). Verbal intrusions precede memory decline in adults with Down syndrome. *Journal of Intellectual Disability Research, 50*(1), 1–10. https://doi.org/10.1111/j.1365-2788.2005.00715.x

Klee, T., & Harrison, C. (2001). *CDI words and sentences: Validity and preliminary norms for British English*. Paper presented at Child Language Seminar, University of Hertfordshire.

Kok, T. B., & Kingma, A. (2009). Herkenningsgeheugen bij kinderen [Recognition memory in children]. *Tijdschrift voor Neuropsychologie, 2,* 42–49.

Kreutzer, J. S., Caplan, B., & DeLuca, J. (2011). *Encyclopedia of clinical neuropsychology*. Springer. https://doi.org/10.1007/978-0-387-79948-3

Langeloo, A., Lara, M. M., Deunk, M. I., LoCasale-Crouch, J., & Strijbos, J. W. (2021). Profiles of learning opportunities of multilingual and monolingual children in kindergarten. *European Journal of Psychology of Education, 36*(2), 379–404. https://doi.org/10.1007/s10212-020-00487-0

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1–75. https://doi.org/10.1017/s0140525x99001776

Levine, D., Pace, A., Luo, R., Hirsh-Pasek, K., Michnick Golinkoff, R., de Villiers, J., Iglesias, A., & Wilson, M. S. (2020). Evaluating socioeconomic gaps in preschoolers' vocabulary, syntax and language process skills with the Quick Interactive Language Screener (QUILS). *Early Childhood Research Quarterly, 50,* 114–128. https://doi.org/10.1016/j.ecresq.2018.11.006

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). Oxford University Press. https://doi.org/10.1017/s1355617700000758

Majerus, S., Poncelet, M., Greffe, C., & Van der Linden, M. (2006). Relations between vocabulary development and verbal short-term memory: The relative importance of short-term memory for serial order and item information. *Journal of Experimental Child Psychology, 92,* 95–119. https://doi.org/10.1016/j.jecp.2005.07.005

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-linguistic easy access resource for phonological and orthographic neighbourhood densities. *PLOS ONE*, *7*(8), e43230. https://clearpond.northwestern.edu

Martin, N. (2009). The roles of semantic and phonological processing in short-term memory and learning: Evidence from aphasia. In A. Thorn & M. Page (Eds.), *Interactions between short-term and long-term memory in the verbal domain* (pp. 220–243). Psychology Press. https://doi.org/10.4324/9780203938966

Martin, N., & Gupta, P. (2004). Exploring the relationship between word processing and verbal short-term memory: Evidence from associations and dissociations. *Cognitive Neuropsychology, 21*(2–4), 213–228. https://doi.org/10.1080/02643290342000447

McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology, 24*(2), 222–243. https://doi.org/10.1037/a0017619

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*(3), 419–457. https://doi.org/10.1037/0033-295x.102.3.419

Meltzer, J. A., Rose, N. S., Deschamps, T., Leigh, R. C., Panamsky, L., Silberberg, A., Madani, N., & Links, K. A. (2016). Semantic and phonological contributions to short-term repetition and long-term cued sentence recall. *Memory and Cognition, 44*(2), 307–329. https://doi.org/10.3758/s13421-015-0554-y

Messer, M. H., Leseman, P. P. P., Boom J., & Mayo A. Y. (2010). Phonotactic probability effect in nonword recall and its relationship with vocabulary in monolingual and bilingual preschoolers. *Journal of Experimental Child Psychology, 105*(4), 306–323. https://doi.org/10.1016/j.jecp.2009.12.006

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology, 91*(1), 3–19. https://doi.org/10.1037/0022-0663.91.1.3

Moors, A., De Houwer, J., Hermans, D., Wanmaker, S., van Schie, K., Van Harmelen, A. L., De Schryver, M., De Winne, J., & Brysbaert, M. (2013). Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. *Behavior Research Methods, 45*(1), 169–177. https://doi.org/10.3758/s13428-012-0243-8

Mulder, F., Timman, Y., & Verhallen, S. (2009). *Handreiking bij de basiswoordenlijst Amsterdamse kleuters (BAK)* [Guide to the basic vocabulary list for Amsterdam toddlers]. ITTA.

Nagel, B. J., Delis, D. C., Palmer, S. L., Reeves, C., Gajjar, A., & Mulhern, R. K. (2006). Early patterns of verbal memory impairment in children treated for medulloblastoma. *Neuropsychology, 20*(1), 105–112. https://doi.org/10.1037/0894-4105.20.1.105

Njiokiktjien, C. (2020). *Gedragsneurologie Van Het kind* [Behavioral neurology in children]. Suyi.

Paradis, M. (2009). *Declarative and procedural determinants of second languages*. John Benjamins. https://doi.org/10.1075/sibil.40

Praktikon. (n.d.). *Bergop platform*. https://www.bergop.info

Repovs, G., & Baddeley, A. (2006). The multi-component model of working memory: Explorations in experimental cognitive psychology. *Neuroscience, 139*(1), 5–21. https://doi.org/10.1016/j.neuroscience.2005.12.061

Rey, A. (1941). L'examen psychologique dans les cas d'encéphalopathie traumatique (les problems) [The psychological examination in cases of traumatic encepholopathy (problems)]. *Archives de Psychologie, 28,* 215–285.

Rey, A. (1958). *L'examen Clinique En Psychologie* [The Clinical Examination in Psychology]. Presses Universitaires de France.

Riccio, C. A., Cash, D. L., & Cohen, M. J. (2007). Learning and memory performance of children with specific language impairment (SLI). *Applied Neuropsychology, 14*(4), 255–261. https://doi.org/10.1080/09084280701719203

Ross, L. A., Molholm, S., Blanco, D., Gomez-Ramirez, M., Saint-Amour, D., & Foxe, J. J. (2011). The development of multisensory speech perception continues into the late childhood years. *European Journal of Neuroscience, 33*(12), 2329–2337. https://doi.org/10.1111/j.1460-9568.2011.07685.x

Saffran, N. M. E. M. (1997). Language and auditory–verbal short-term memory impairments: Evidence for common underlying processes. *Cognitive Neuropsychology, 14*(5), 641–682. https://doi.org/10.1080/026432997381402

Schaerlaekens, A. M. (2008). *De taalontwikkeling van het kind. Een oriëntatie in het Nederlandstalig onderzoek* [The language development of the child. An orientation in Dutch-language research]. Wolters-Noordhoff.

Schlichting, J. E. P. T., & Spelberg, H. C. L. (2012). *Schlichting Test voor Taalproductie-II: voor Nederland en Vlaanderen* [Schlichting Test for Expressive Vocabulary, second edition for the Netherlands and Flanders]. Bohn Stafleu van Loghum.

Schlichting, L. (2004). *PPVT-III-NL | Peabody Picture Vocabulary Test-III-NL*. Pearson.

Schlichting, L., & de Koning, G. (1998). *TenT, taalachterstand en taalverwerving. Een programma voor kinderen met een grammaticale en lexicale achterstand* [A therapy program for children with grammatical and lexical difficulties]. Harcourt Test Publishers.

Semel, E., Wiig, E., & Secord, W. A. (2012). *CELF Preschool-2-NL. Nederlandse bewerking door de Jong, J* [CELF Preschool-2-NL. Dutch adaptation by de Jong, J.]. Pearson.

Sheslow, D., & Adams, W. (1990). *Wide Range Assessment of Memory and Learning*. Jastak.

Stone, C. A. (2004). *Handbook of language and literacy: Development and disorders*. Guilford Press.

TOP tot TEEN beweegspel [Head-to-toe movement game]**.** (2015). Dubbelzes.

Vakil, E., Blachstein, H., Wertman-Elad, R., & Greenstein, Y. (2012). Verbal learning and memory as measured by the Rey-Auditory Verbal Learning Test: ADHD with and without learning disabilities. *Child Neuropsychology, 18*(5), 449–466. https://doi.org/10.1080/09297049.2011.613816

Van den Burg, W., & Kingma, A. (1999). Performance of 225 Dutch school children on Rey's Auditory Verbal Learning

Test (AVLT): Parallel test–retest reliabilities with an interval of 3 months and normative data. *Archives of Clinical Neuropsychology, 14*(6), 545–559. https://doi.org/10.1016/s0887-6177(98)00042-0

Vingerhoets, G., & Lannoo, E. (1998). *Handboek neuropsychologie* [Handbook of neuropsychology]. Acco.

Vlach, H. A., & Sandhofer, C. M. (2012). Fast mapping across time: Memory processes support children's retention of learned words. *Frontiers in Psychology, 3,* 46–46. https://doi.org/10.3389/fpsyg.2012.00046

Walker, E. A., & McGregor, K. K. (2013). Word learning processes in children with cochlear implants. *Journal of Speech, Language, and Hearing Research, 56*(2), 375–387. https://doi.org/10.1044/1092-4388(2012/11-0343)

Wechsler, D. (2010). *WPPSI-III-NL. Nederlandse bewerking door Hendriksen, J. & Hurks, P* [WPPSI-III-NL Dutch adaptation by Hendriksen, J. & Hurks, P.]. Pearson.

Wechsler, D. (2014). *The Wechsler Intelligence Scale for Children–Fifth Edition*. The Psychological Corporation.

Yu, C., & Ballard, D. H. (2007). A unified model of early word learning: Integrating statistical and social cues. *Neurocomputing, 70*(13–15), 2149–2165. https://doi.org/10.1016/J.NEUCOM.2006.01.034

Zink, I., & Lejaegere, M. (2002). *N-CDI's: Lijsten voor communicatieve ontwikkeling. Aanpassing en hernormering van de MacArthur CDI's van Fenson et al. Acco* [Lists for communicative development. Adjustment and re-standardization of the MacArthur CDI's from Fenson et al.].