Phonological Memory and the Acquisition of Grammar in Child L2 Learners

Josje Verhagen, Paul Leseman, & Marielle Messer

Utrecht University

Correspondence concerning this article should be addressed to:

Josje Verhagen

Utrecht University

Postbus 80140, Utrecht 3508 TC, the Netherlands.

E-mail: J.Verhagen@uu.nl

Abstract

Previous studies show that second language (L2) learners with large phonological memory spans outperform learners with smaller memory spans on tests of L2 grammar. The current study investigated the relationship between phonological memory and L2 grammar in more detail than has been done earlier. Specifically, we asked how phonological memory relates to specific L2 grammar skills, after controlling for L2 vocabulary, and using different phonological memory tasks. Participants were 36 Turkish child learners of Dutch and 34 Dutch first-language (L1) children. All participants completed a Dutch narrative task to assess their production of subject-verb agreement, auxiliaries, and verb placement, and a Dutch vocabulary test. Phonological memory was measured through serial recall of Dutch words, high-probability nonwords, and low-probability nonwords. The results show weak correlations between phonological memory and grammar in the L1 group due to ceiling effects. For the L2 group, moderate to strong correlations between phonological memory and grammar were found. Regression analyses showed that word recall significantly predicted all three L2 grammar skills, above and beyond vocabulary. These findings indicate that the ability to temporarily store L2 material in phonological memory is important for L2 grammar learning, but that specifics of the memory tasks also play a role.

Keywords: child L2 acquisition, phonological memory, (non)word recall, grammar, vocabulary

Phonological Memory and the Acquisition of Grammar in Child L2 Learners

Numerous studies have shown that phonological memory, or verbal short-term memory, is involved in word learning such that children with large phonological memory spans are better word learners than children with small memory spans, both in first (L1) and second language (L2) learning (Atkins & Baddeley, 1998; Cheung, 1996; Gathercole, 2006; Masoura & Gathercole, 2005). Evidence is accumulating that phonological memory is also related to L1 and L2 grammar learning (Adams & Gathercole, 1996; 2000; Blake, Austin, Cannon, Lisus, & Vaughan, 1994 for L1; French & O'Brien 2008; Paradis, 2011; Service & Kohonen, 1995 for L2). For L2 acquisition, studies show that children with good phonological memory skills perform better on tests of L2 grammar than children with poorer memory skills. In most studies, however, effects of phonological memory on grammar are mediated by L2 vocabulary skill (Engel de Abreu & Gathercole, 2012; French, 2006; Service & Kohonen, 1995; but see French & O'Brien, 2008).

In this study, we ask how phonological memory is associated with the acquisition of specific grammatical structures in Turkish child L2 learners of Dutch. Our aim is threefold. First, we examine whether in these children differences in phonological memory are related to the acquisition of three grammatical sub-skills (i.e., subject-verb agreement, auxiliary verbs, and verb placement). Second, we investigate if effects of phonological memory are independent of vocabulary knowledge. Finally, we examine to what extent relationships between phonological memory and grammatical skills depend on the type of phonological memory task used. To this aim, we include three measures of phonological memory that vary in the extent to which performance on the task is dependent on existing L2 knowledge. By investigating the relationships between phonological memory and grammar in more detail than has been done in earlier studies, we hope to shed more light on when phonological memory and L2 grammar are associated and when they are not.

Review of Literature

In previous studies, it has been assumed that phonological memory and grammar learning are related because learners with well-developed memory spans are better able to create long-term linguistic representations than learners with smaller memory spans (Ellis & Sinclair, 1996; Speidel, 1989, 1993). While empirical support for a relationship is increasing, there is still relatively little research on the role of phonological memory in (child) L2 acquisition. Acquisition studies adopting a generative framework, for example, have typically left little room for the potential impact of individual factors on grammar learning, such as differences in working memory skill.

Evidence that phonological memory is related to L2 grammar learning comes from two types of studies. First, artificial and foreign language learning studies have found associations between phonological memory and grammatical abilities, both in children (Daneman & Case, 1981) and in adults (Williams & Lovatt, 2003). Daneman and Case (1981) taught novel sentences describing interactions between bug-like creatures to English-speaking two- to six-year-olds. Their results showed that word span correlated significantly with children's production and comprehension of the novel sentences. In a study with adults, Williams and Lovatt (2003) found that phonological memory was related to the ability to learn determiner-noun agreement in an unknown language (Italian) as well as an artificial language. To explain this, they suggest that phonological memory is needed to generalize patterns in familiar items to new sequences. This is very similar to the proposal made by Ellis and Sinclair (1996) who argued that the more often language structures are rehearsed in phonological memory, the more likely it is that they are learned and generalized (see also Martin & Ellis, 2012).

A second source of evidence for an association between phonological memory and L2 grammar learning comes from child L2 classroom studies (Engel de Abreu & Gathercole, 2012; French, 2006; Kormos & Sáfár, 2008; Service, 1992; Service & Kohonen, 1995). Service (1992) and Service and Kohonen (1995) examined the relation between phonological memory and scores on a variety of tests requiring knowledge of grammar in Finnish school-aged children learning English. In both studies, phonological memory, assessed with nonword repetition, was a significant predictor of children's test scores a few years later. However, the effects of phonological memory were explained by increases in children's vocabulary knowledge. Similar results were obtained by French (2006), who found that effects of phonological memory at time 1 on English grammar at time 2 in French-speaking children learning English were mediated by L2 vocabulary knowledge. Likewise, Engel de Abreu and Gathercole (2012) observed that relationships between phonological memory and grammatical proficiency in Luxembourgian children's first, second, and third language were mediated by vocabulary knowledge.

Different results were obtained by French and O'Brien (2008), however. They found effects of phonological memory on L2 grammatical acquisition were independent of vocabulary skill. French and O'Brien examined data from eleven-year-old French-speaking children over a five-month intensive English program. As a measure of phonological memory, they not only administered children a nonword repetition task based on English, but also a repetition task containing Arabic words. The authors assumed that performance on the Arabic task would not be influenced by children's existing linguistic knowledge (about French and English), and they were interested in investigating the relationships with grammar learning for both tasks separately. Grammar was measured through a written test assessing a variety of grammatical structures that receive explicit instruction in English programs in Quebec, such as tense, morphological inflections, negation, and word order. Phonological memory was shown to predict a significant amount of variance in grammar in addition to the contribution made by vocabulary knowledge, regardless of whether English-like nonword repetition or Arabic word repetition was taken as a measure of phonological memory. These results differ from those of Engel de Abreu and Gathercole (2012), French (2006), Service (1992), and Service and Kohonen (1995), all of whom found that effects of phonological memory were mediated by vocabulary knowledge. French and O'Brien's finding that phonological memory and grammar are interrelated irrespective of vocabulary knowledge is important, as it shows that vocabulary is not the underlying factor. Vocabulary is likely to be correlated with grammar in L2 learning and it is also known to be related to phonological memory, especially when measured with tasks involving words or wordlike nonwords (Masoura & Gathercole, 2005).Consequently, it is important to ensure that any correlation between phonological memory and grammar is not due to their common association with vocabulary.

The studies on the relationship between phonological memory and L2 grammar reviewed above have measured L2 grammar through tests assessing a mixture of morphological and syntactic structures. To the best of our knowledge, only two studies have investigated how phonological memory relates to more specific L2 grammar skills. Firstly, Paradis (2011) examined the relation between phonological memory and child L2 learners' scores on a test assessing the production of bound and free morphemes third-person singular *–s*, simple past *–ed*, irregular past, copula *be*, and the auxiliaries *be* and *do*. She found that phonological memory predicted a significant amount of variance in children's production of tense and agreement morphemes. However, as vocabulary was not controlled, we do not know whether the effect of phonological memory was unmediated by lexical knowledge. Secondly, O'Brien, Segalowitz, Collentine, & Freed (2006) found that phonological memory predicted adult English learners' production of function morphemes such as third-person singular *–s* in a narrative task in Spanish. Like in Paradis (2011), however, vocabulary was not controlled in this study, so the alternative explanation cannot be excluded that the relationship between phonological memory and the production of function morphemes was driven by differences in vocabulary skill.

As noted by Gathercole (2006), the presence of a correlation between phonological memory and language learning is of limited theoretical (and practical) value, as long we do not have clear ideas as to why differences in phonological memory would affect language learning. For word learning, researchers have assumed that phonological memory affects the efficiency with which long-term memory phonological representations are created (Baddeley, Gathercole, & Papagno, 1998; Gathercole & Baddeley, 1990; Metsala, 1999). That is, individuals with good phonological memory skills are assumed to create more robust and stable phonological representations, which are needed for language learning, than individuals with poorer phonological memory skills. A similar proposal has been made for grammar learning by Speidel (1989, 1993) and Speidel and Herreshoff (1989). They proposed that phonological memory affects children's ability to imitate adult models of morphosyntactic constructions, and their ability to store these models in a long-term store of linguistic patterns. From this store, children can use templates made up of chunked constructions in spontaneous speech to support grammar learning. In connectionist views on language acquisition, it has also been assumed that learning syntax is similar to learning words. Ellis and Sinclair (1996), for example, argued that in order to put words in good order, they must be stored in phonological memory, just as phonemes in word learning. The more often they are kept in phonological memory, the better they are learned and the easier it will be to generalize rules from them.

The relationship between phonological memory and language learning is not unidirectional, however. Previous studies have shown that performance on nonword repetition tasks in children improves over time, as a function of growing language proficiency (Gathercole et al., 1992). Phonological memory is also superior for nonwords that are wordlike versus nonwords that are less wordlike (Gathercole, Willis, Emslie, & Baddeley, 1991) and for nonwords composed of high-frequent phoneme combinations versus nonwords composed of low-frequent phoneme combinations (Vitevitch & Luce, 2005). This evidence shows that phonological memory is not an independent ability needed for language learning, but that it is itself influenced by existing language knowledge. Such long-term language knowledge (about words and phonemes) has been assumed to be applied during phonological storage to reconstruct blurred or incomplete memory traces, a process that has been termed redintegration or pattern completion (Brown & Hulme, 1997; Thorn, Gathercole, & Frankish, 2005).

The degree to which long-term knowledge supports phonological storage not only depends on the language-likeness of the material to be stored, but also on the amount of linguistic knowledge learners have available in long-term memory. For child L2 learners, Messer, Leseman, Boom, & Mayo (2010) found that benefits of long-term phonotactic support were greatest in children's dominant language, even though benefits of long-term phonotactic knowledge were also found in the L2. Moreover, this study showed that relationships with L2 vocabulary differed depending on the type of phonological memory measure used. Specifically, phonological memory assessed through recall of L2-like nonwords was related more strongly to L2 vocabulary than phonological memory assessed through recall of less L2-like nonwords. Similarly, Parra, Hoff, and Core (2011) found that relationships between phonological memory and grammar were dependent on the type of phonological memory measure used in native Spanish-English bilingual children. In these children, English-like nonword repetition correlated significantly with English vocabulary and grammar, but not with Spanish vocabulary and grammar. Spanish-like nonword repetition, in contrast, correlated with Spanish vocabulary and grammar, but not with English vocabulary and grammar. These results suggest that children's benefit from long-term linguistic knowledge when performing phonological memory tasks is language-specific. The results in Messer et al. (2010) and Parra et al. (2011) contrast with those of French and O'Brien (2008), however, who found that Arabic repetition was a stronger predictor of L2 English grammar than English-like nonword repetition. One explanation of this contradictory pattern of results is that the

English-like repetition task in French and O'Brien did not tap support from long-term knowledge and thus measured the same skill as the Arabic task, at least at time 1, as is also noted by the authors (French & O'Brien, p. 480).

Summarizing the review of studies we have presented, previous results show that phonological memory is correlated with grammatical performance in L2 acquisition. To explain this, it has been proposed that the temporary storage of individual phrases and sentences provides a database of structures from which learners can generalize and abstract grammatical patterns (Ellis & Sinclair, 1996; Speidel, 1993). There are some contradictory findings across studies, however. First, while most studies find that effects of phonological memory on L2 grammar are mediated by L2 vocabulary knowledge (Engel de Abreu & Gathercole, 2012; French, 2006; Service, 1992; Service & Kohonen, 1995), a few studies have found direct effects of phonological memory on grammar (French & O'Brien, 2008; Martin & Ellis, 2012). Second, whereas some researchers found that the relationship between phonological memory and grammar or vocabulary in a given language was strongest if phonological memory was assessed through nonwords that conform to the phonotactic rules of that language (Messer et al., 2010; Parra et al., 2011), others found the opposite result (French & O'Brien, 2008).

These differences across studies may originate from differences in the type of grammatical structures studied and/or the use of different phonological memory tasks. In this study, we examine data from Turkish child learners of Dutch to investigate how phonological memory relates to the acquisition of specific L2 grammar skills, using multiple phonological memory tasks. We address three questions. First, we ask how phonological memory relates to the acquisition of three grammatical sub-skills in L2

Dutch: subject-verb agreement, auxiliary verbs, and verb placement. Second, we ask if relationships between phonological memory and L2 grammar skills still hold after differences in L2 vocabulary are controlled. Finally, we ask whether there are differential relationships between phonological memory and L2 grammar skills depending on how phonological memory is assessed. To investigate this last question, we include data from three types of phonological memory task that vary in the extent to which performance on the task relies on existing L2 knowledge. Answering these questions will help to obtain a more detailed picture of where phonological memory and grammar learning are associated and where they are not.

This Study

In this section, we present a number of considerations that motivated our design and methodological choices as well as our expectations for the data.

The context of choice for our investigation of the relationship between phonological memory and grammar learning involves children who were raised in minority-language families in which Turkish was the main language of communication. The majority language Dutch only played a very minor role via communicative contacts outside the home and television, and systematic exposure to Dutch only started when children entered preschools or kindergarten at the age of three or four. At that age, they were immersed in Dutch kindergarten where Dutch was the only language of instruction, and no explicit L2 training was provided. Given that children were overwhelmingly exposed to Turkish during their early years, they are referred to as L2 learners rather than bilinguals, following other studies on the same population (Blom, 2008; Orgassa, 2009;

Verhoeven, Steenge, Van Balkom, 2011). Dutch monolingual children were also included to allow us to compare task performance between the two groups.

Child L2 learners often have trouble with tense and agreement markers such as verbal inflections and auxiliary verbs, both in Dutch (Blom, 2008; Orgassa, 2009; Verhoeven et al., 2011) and in other languages such as English and German (Haznedar, 2001; Ionin & Wexler, 2002; Prévost, 2003). In Dutch, finite verbs occur in second position in declarative main clauses due to a verb-second rule. This is illustrated in (1), where the verb *kopen* 'buy' appears in second position and bears inflections for tense and agreement (third-person singular present, henceforth 3PL):

(1) De ouders kopen een fiets voor hun dochter

The parents buy.*3PL* a bike for their daughter

Nonfinite verbs are placed in sentence-final position in Dutch. This can be seen in (2), where the non-finite past participle (PP) appears at the end and the finite auxiliary verb *heeft* 'has' is in second position:

(2) Het meisje heeft de hele dag gefietst

The girl has.*3SG* the whole day bike.*PP*

Finally, subject-verb agreement on finite verbs is marked through suffixation in Dutch: first-person singular (1SG) is marked through a zero suffix (*koop* 'buy'), second- and third-person singular (2SG, 3SG) through '-t' (*koopt* 'buys'), and first-m second-, and third-person plural (1PL, 2PL, 3PL) through '-en' (*kopen* 'buy').

Blom (2008) investigated verb inflection and verb placement in the production of Turkish children whose systematic exposure to Dutch also started around age four. Her results showed that children aged 4;8 to 8;0 years produced on average 63% correct subject-verb agreement in an elicited-production task. Verb placement was relatively accurate, even though not fully mastered (88% correct). However, as most of the children in our study were considerably younger (i.e., four years), we expected to find less accurate behavior in the current sample. In relation to problems with auxiliary verbs, researchers have observed that Turkish (and Moroccan) child learners of Dutch may omit the auxiliary verbs *hebben* 'have' and *zijn* 'be' and produce bare past participles instead (Verhoeven et al., 2011).

Serial recall tasks are used in this study to assess phonological memory. In serial recall, participants repeat lists of items while the lists increase in length. Like other tests of phonological memory such as nonword repetition, nonword recall requires little phonological processing and shows short-term memory characteristics such as decreasing accuracy with increasing list length (Archibald & Gathercole, 2007). In fact, in a study comparing nonword repetition and serial nonword recall, Archibald and Gathercole (2007) found that output demands were considerably less for nonword recall than nonword repetition, presumably because the multisyllabic items in nonword repetition elicit more co-articulated speech gestures than the monosyllabic items in nonword recall.

The participants in our study were presented with three serial recall tasks: one word recall task and two nonword recall tasks. In the nonword recall tasks, children repeated nonwords composed of high-frequent phoneme sequences (i.e., high-phonotactic probability nonwords) and nonwords composed of low-frequent phoneme sequences (i.e., low-phonotactic probability nonwords). Of these three tasks, word recall is supported most by long-term linguistic knowledge and thus is the most language-dependent measure. Previous research shows that even though it is generally seen as a measure of phonological storage because the task is subject to phonological effects (e.g., phonological errors) and serial position effects (Jefferies, Lambon Ralph, & Baddeley, 2004), some degree of semantic encoding may also take place, depending on properties of the stimuli used (e.g., word familiarity/concreteness) and task instructions (Campoy & Gathercole, 2008). High-probability nonword recall depends on linguistic knowledge to a lesser degree, as unfamiliar items (novel words) are used, but still is supported by long-term knowledge about phoneme combinations. Low-probability recall is least language-dependent, as the items contain infrequent phoneme combinations for which there is no or only very little support from long-term memory.

The following three questions are addressed:

1. How does phonological memory relate to the acquisition of subject-verb agreement, auxiliary verbs, and verb placement in child learners of Dutch?
2. If significant relationships are found, do they still hold after Dutch vocabulary is controlled?
3. Do relationships between phonological memory and grammar differ depending on the type of phonological memory task used, that is, on whether memory tasks are more or less dependent on long-term Dutch knowledge?

All three questions will be investigated in the L2 children but, where possible, the data of the L1 comparison group will also be included to see if results differ between the groups. However, it turned out that the monolingual children scored close to ceiling on the grammar measures, so investigations of the relationship between phonological memory related and grammar were not possible for these children.

As for the first question, we predict that phonological memory is related to the acquisition of all three grammatical structures. Given that phonological memory is assumed to be beneficial for creating a storehouse of linguistic structures that can be used for language production and rule generalization (Ellis & Sinclair, 1996; Speidel 1989, 1993), we expect that there will be significant relationships with all three grammatical structures. In addition, previous studies have found significant correlations between phonological memory and the production of agreement and auxiliary morphemes in L2 children (Paradis, 2011) as well as between phonological memory and the production of word order in children learning an artificial language (Daneman & Case, 1981). One possible outcome of the current study is that relationships with phonological memory are stronger for certain types of structure than for others. Since previous studies have used composite scores on grammar tests rather than focused on specific grammatical phenomena, we do not have any a priori predictions as to which subskills should be especially sensitive to differences in phonological memory. Therefore, the question of whether relationships with phonological memory will be stronger for certain grammar skills than for others (i.e., morphology vs. syntax or bound vs. free function morphemes) is exploratory.

Regarding our second question, no clear prediction is formulated. Most previous studies found that effects of phonological memory on L2 grammatical proficiency, operationalized as composite scores on L2 grammar tests, were mediated by L2 vocabulary, but there are some exceptions (French & O'Brien, 2008; Martin & Ellis, 2012). Previous studies looking at the L2 acquisition of function morphemes and word order did not control for vocabulary knowledge (O'Brien et al., 2006; Paradis, 2011), so it is an open question whether any possible effects in the current study will be independent of L2 vocabulary.

Regarding our final question, we predict that effects will be strongest for the most language-dependent measure (word recall) and least strong for the most language-independent measure (low-probability nonword recall). Specifically, based on previous studies showing language-specific relationships between phonological memory and vocabulary and/or grammar in L2 and bilingual children (Messer et al., 2010; Parra et al., 2011), we expect that Dutch word recall will be related to the acquisition of Dutch grammatical structures most strongly. Recall of Dutch high-probability nonwords will be less strongly related to the acquisition of Dutch grammar skills, and the weakest relationships will be with recall of Dutch low-probability nonwords, as the recall of such nonwords is relatively independent of existing L2 knowledge.

Method

Participants

Participants were 36 Turkish children who acquired Dutch as their L2. Children's mean age was 52 months (*SD* = 2.9, min-max = 49 - 66) and there were 22 boys and 14 girls. These children were a subset of the children studied in Messer et al. (2010). They were selected from this larger subset (*N* = 60) if they were predominantly addressed in Turkish by their parent(s) and if they had completed all the tasks reported on in the current study out of a larger test battery that also contained nonlinguistic tasks. One child was excluded as he produced fewer than five analyzable utterances in the production task, rendering his data not informative for the current purposes. A comparison group of 34 Dutch monolingual children was also included. These children also had a mean age of 52 months (*SD* = 2.1, min-max = 48 - 56) and there were 21 boys and 13 girls. These monolingual children also were a subset of the children studied in Messer et al. (*N* = 67), who were selected if they came from monolingual Dutch families and if they had completed all the tasks analyzed for this study.

In the L2 group, 27 children (75%) came from low/middle socioeconomic background (SES) families, defined as having parents with intermediate or vocational track as their highest attained educational level. In the L1 group, 13 children were from low/middle SES families (38%). This situation is representative for the two groups, as minority children with Turkish as their home language typically come from lower SES backgrounds than their Dutch monolingual peers in the Netherlands (Sociaal en Cultureel Planbureau [SCP], 2007). To control for this difference, SES was included as a covariate in the analyses comparing between-group performance.

All children were recruited through schools that had a moderate to high proportion of ethnic minority children. Children's primary caregivers were administered a screening questionnaire to assess whether they interacted with their child in Dutch (L1 group) or Turkish (L2 group). After they had agreed to participate in the project, caregivers completed another, detailed questionnaire about the type and frequency of language and literacy activities they performed with their child (cf. Scheele, Leseman, & Mayo, 2010). This questionnaire was administered in the form of an interview by trained research assistants. Assistants who were fluent in both Turkish and Dutch conducted the interviews with the L2 children's caregivers.

For the L2 group, the interview data indicated that the average amount of Turkish spoken at home (as opposed to Dutch) was 85% (*SD* = 18.3, min-max = 50 - 100). These data indicated that 13 families only spoke Turkish at home (36%), 19 families reported that one parent always spoke Turkish and the other parent spoke both Turkish and Dutch (52%), and five families reported that both parents spoke Turkish and Dutch (14%).

Nearly all of the L2 children had been exposed to Dutch via day care or preschools prior to kindergarten entrance (*N* = 34, 94% of the children), which they had attended on average for 3.5 half days per week (min-max = 2 - 5 half days, *SD* = 1 half day). No significant correlations were found between the language or memory measures and the percentage of Dutch spoken at home (*r*s between .09 and .22) or the number of days spent at Dutch preschools prior to kindergarten entry (*r*s between -.06 and .11). In the L1 group, 100% of all parental input was in Dutch. Children had been enrolled in kindergarten for on average four months (SD = 2 months, min-max = 1 - 8). The vast majority of children in this group also had attended day care or preschools for an average of four half days per week (*N* = 31, 91%). Informed consent was obtained for each child.

Procedure

Testing took place in a quiet room at children's schools and was done by trained research assistants who were fluent in Turkish and Dutch. There were two sessions which were on average one week apart. Each testing session lasted for approximately 75 minutes, including play breaks and tasks that were part of another study. The tasks were intermixed with other tasks and administered in a fixed sequence that aimed to optimally vary the task demands from one task to the next and avoid fatigue. The other tasks in the task battery tapped phonological memory in Turkish through word recall, and high- and low-probability nonword recall, and are reported on in Messer et al. (2010). The memory and language tasks reported on in this study were administered in the following order: Dutch vocabulary, word recall, low-probability nonword recall, narrative production (Day 1), and high-probability nonword recall (Day 2). To keep children motivated, they were rewarded with a small sticker after each task. All the sessions were videotaped.

Phonological Memory Tasks

Serial word recall. This task was a Dutch adaptation of the serial word recall task in the Automated Working Memory Assessment (AWMA) battery (Alloway, 2007, cf. Messer et al., 2010 for the Dutch adaptation). In this task, children heard a sequence of words and had to recall each sequence in the correct order. The task was presented on a laptop, and pre-recorded sentences by native speakers were used for the presentation of the stimuli. The stimuli consisted of nouns, adjectives, adverbs, color names, and verbs (stems). To make sure that the stimuli were age-appropriate, native speakers had judged the words as not being too abstract for young children. Two practice trials were presented to familiarize children with the procedure. The task then started with a block of one item and presented children with blocks of increasing length, up to a block of seven items. Each block contained six trials. The scoring procedure of the AWMA was applied such that trials were scored as incorrect if one of the items was omitted, if the sequence of items was incorrect, or if an item was recalled incorrectly. If a child remembered the first four trials within a block correctly, (s)he automatically received a score of six and proceeded to the next block. Testing stopped after three incorrect recalls within one block. The scores could range from zero to 42.

Serial nonword recall. Two serial nonword recall tasks from the AWMA were adapted for Dutch (cf. Messer et al., 2010). Like the word recall task, they were administered on a laptop, and pre-recorded sentences by native speakers were used for the stimuli presentation. Two practice trials were presented. There were two sets of nonwords: nonwords with high transitional biphone probability and nonwords with low transitional biphone probability. These probability counts were based on a corpus of 42 children's books (cf. Messer, 2010) and checked against the lemma frequencies in the Dutch database CELEX (Baayen, Piepenbrock, & Gulikers, 1995). Biphone frequencies were calculated by summating the relative frequencies (per 10,000) of each word form in the corpus containing the biphone, and triphone frequency counts were used to correct for Dutch diphthongs (cf. Messer et al., 2010 for more details). A one-way ANOVA showed that the high-probability nonwords had significantly higher summated biphone frequencies than the low-probability nonwords based on CELEX lemma frequencies ($F(1,70) = 55.50$, $p < .001$, $\eta^2_p = .44$). The high- and low-phonotactic probability nonwords were also examined for phonotactic probability in Turkish to see whether Turkish children would have additional support from their L1, due to overlapping phonotactics between Dutch and Turkish. No significant difference in Turkish phonotactic probability between the Dutch high- and low-probability nonwords was found (see also Messer et al., 2010).

Dutch native speakers' likeliness ratings confirmed the difference in phonotactic probability between both sets on nonwords: A group of 16 native speakers gave significantly higher ratings to high-probability than low-probability items on a scale ranging from '0' (does not sound like a Dutch word at all)' to '5' (sounds a lot like a Dutch word) ($F(1,70) = 37.45$, $p < .001$, $\eta^2_p = .35$). Finally, length was checked to control for possible confounds (Vitevitch & Luce, 2005). The high-probability words were longer than the low-probability words ($F(1,70) = 5.70$, $p = .020$, $\eta^2_p = .08$), but as this difference would run counter to the predictions, it was not considered problematic. Table 1 summarizes the relevant properties of the items in the high- and low-probability nonword recall tasks.

[Insert Table 1 about here]

The procedure in the nonword tasks was the same as in the word recall task: Children repeated voice-recorded monosyllabic nonwords in lists of increasing length, starting with a block of one nonword up to a block of five nonwords (see Appendix for a list of the nonwords). Each block consisted of six trials. The original scoring procedure from the AWMA was applied such that a trial was rewarded with a score of one when none of the nonwords was omitted, the sequence was correct, and each nonword was recalled correctly. Each phoneme needed to be recalled correctly for a positive score, with the exception of consistently substituted phonemes resulting from articulation problems. With a total of six trials per block, the maximum score per block was 6. When the first four trials within a block were recalled correctly, the child automatically received a score of 6 and proceeded to the next block. Testing stopped after three incorrect recalls within one block. The scores could range from zero (first block) to 30 (fifth block), but none of the children exceeded the third block, so the maximum score was 18.

Vocabulary

The Dutch Test for Bilingualism (Verhoeven, Narain, Extra, Konak, & Zerrouk, 1995) was used to assess children's receptive vocabulary. This test has been designed for research into bilingual development and contains two language versions that can be considered equivalent. For the current study, only the Dutch version of the test was used. In this test, children heard a word and then chose one out of four line drawings presented on a laptop. Items involved nouns, verbs, adverbs, and numerals. The test started with a short practice session. To avoid fatigue, the original test was shortened by taking only the even items, resulting in 30 items. Scores were calculated as the percentage correct responses out of all responses for each child.

Narrative Production

To elicit production data from the children, children were read a story using an age-appropriate book with pictures and text of approximately 350 words. A hand puppet (Ernie from Sesame Street) was used as a playmate to enhance children's engagement in the task. The book had been translated from English to ensure that children would not be familiar with it. The story was about a cat that finds the new neighbors' kitten in her kitchen, tries to chase it away, but then ends up making friends with it. After having listened to the story, children were asked a series of questions to check their comprehension. Subsequently, they were asked to retell the story to Ernie, as Ernie had not been paying attention while the story was being read. Using the hand puppet, the experimenter signaled interest using only minimal responses (*hmm*, *ooh*) and non-verbal cues such as nodding.

Coding and Analysis

Native speakers checked the recordings of the memory tasks. In case of disagreement on more than half of the trials within a block, a third check by one of the investigators was decisive (less than 6% of the trials). The production data from the narrative task were transcribed and coded for subject-verb agreement, production of auxiliary verbs, and verb placement. Self-corrections, off-task utterances, and repetitions of the experimenter's speech were excluded for each child. The same was true of imperatives and elliptic utterances directly following an experimenter's elicitation question such *What is the cat doing?*, because such utterances are likely to elicit infinitives.

As for subject-verb agreement, it was coded whether there was correct agreement between the verb and the subject of the sentence. Specifically, thematic verbs that had correct verb inflections for person and number were coded as having correct agreement, whereas thematic verbs that did not have appropriate person and number inflections were coded as having incorrect agreement. An example of a sentence with incorrect agreement is *Loekie slapen* 'Loekie sleep', as the correct Dutch sentence would be *Loekie slaapt* 'Loekie sleeps'. Non-thematic verbs such as copula and auxiliary verbs were not analyzed, because such verbs were virtually always inflected correctly by the children and thus did not yield much variation across children, in line with earlier studies (Blom, 2008; Parodi, 2000). Scores were calculated as the percentage of utterances containing a correctly agreeing thematic verb out of all utterances containing a thematic verb.

As for the use of auxiliaries, we coded whether children produced the auxiliary verbs *zijn* 'be' and *hebben* 'have' in contexts that require such verbs. Example sentences without and with the auxiliary *zijn* are *Stoel gevallen* 'Chair fallen' and *De poes is al gevallen* 'The cat has already fallen'. Scores were computed as the percentage of utterances containing the auxiliary *hebben* or *zijn* out of all obligatory contexts.

Finally, regarding verb placement, we coded if children placed thematic verbs in second position in main clauses. Coding was restricted to sentences with a negator, adverbial, or object, because the syntactic position of the verb cannot be determined if there is no such element. More precisely, in a sentence like *Hij werkt* 'He works', we cannot tell whether the verb is in second or final position, but if there is an extra element such as an adverb (e.g., *Hij werkt vandaag* 'He works today'), the positioning of the verb becomes unambiguous. Utterances in which the verb preceded negation, an adverbial or an object were coded as having correct verb placement, whereas utterances in which the verb followed such elements were coded as having incorrect verb placement. Examples are *Loekie hier slapen* 'Loekie here sleep' (incorrect) and *Hij slaapt hier* 'He sleeps here' (correct). Again, non-thematic verbs were not included, because we know from earlier studies that learners do not have problems placing such verbs in a correct position (Parodi, 2000) and this was confirmed by our data in which nearly all non-thematic verbs were in second position. Scores were calculated as the percentage of utterances with correctly placed thematic verbs out of the total number of utterances containing thematic verbs and a negator, adverbial, or object.

Results

Scores on all assessments are presented in Table 2 for the L1 and L2 children. Mean percentages correct are provided for the language measures. However, as percentages based on few data are not very informative, Table 2 also presents information about the mean number of produced constructions for each structure, in parentheses; these numbers indicate how many utterances were analyzed (i.e., obligatory contexts). For example, for subject-verb agreement, 16.9 in the L2 group indicates that the L2 children on average produced 16.9 utterances that could be analyzed for subject-verb agreement (i.e., had a subject and a thematic verb) – 47.3% of which had a correctly agreeing verb.

[Insert Table 2 about here]

The L1 children performed close to ceiling on all grammar measures, and significantly outperformed the L2 children on subject-verb agreement ($F(1,69) = 70.10$, $p < .001$, $\eta^2_p = .51$), the production of auxiliaries ($F(1,66) = 1439.66$, $p < .001$, $\eta^2_p = .96$), and verb placement ($F(1,64) = 25.33$, $p < .001$, $\eta^2_p = .29$). Mean accuracy scores in the L2 group were around 50% on subject-verb agreement and verb placement, but there was much within-group variation, as indicated by the large standard deviations. The percentage of auxiliary realizations was very low in all children. As for vocabulary, we found, as expected, that the L1 children obtained significantly higher scores than the L2 children on the Dutch vocabulary test ($F(1,69) = 50.10$, $p < .001$, $\eta^2_p = .42$).

The L1 children also performed significantly better than the L2 children on word recall ($F(1,69) = 9.04$, $p = .004$, $\eta^2_p = .12$) and high-probability nonword recall ($F(1,69) =$ 6.53, $p = .013$, $\eta^2_p = .09$), but not on low-probability nonword recall ($F(1,69) = 2.15$, $p >$.

10, $\eta^2_p = .03$). These results fit well with the idea that long-term linguistic knowledge supports performance more in word recall than in high-probability nonword recall, and provides no or only very little support in low-probability nonword recall. The results in Table 2 also show that both groups of children performed better on high-probability nonword recall than on low-probability nonword recall. This difference was significant for the L1 group ($F(1,33) = 19.23$, $p < .001$, $\eta^2_p = .37$) and the L2 group ($F(1,35) = 16.85$, $p < .001$, $\eta^2_p = .33$ for L2 group). Repeating all analyses with SES as a covariate in the analyses yielded the same pattern of results: The L1 children significantly outperformed the L2 children on all tasks (all *p*s < .05), except low-probability nonword recall ($p > .10$)

Table 3 presents all bivariate correlations among vocabulary and the language and memory measures for the L1 and L2 children.

[Insert Table 3 about here]

For the L2 group, these data show that vocabulary was moderately and significantly correlated with word recall but not with the nonword measures. Vocabulary also correlated significantly with two of the three grammar measures (subject-verb agreement and verb placement). Word recall correlated significantly with all three grammar measures in this group. High-probability nonword recall correlated significantly with auxiliary use, and low-probability nonword recall correlated significantly with subject-verb agreement and verb placement. Correlations in the L1 group were weaker overall, and none of the correlations with the grammar measures was significant, probably due to the L1 children's homogeneous and very high scores. For vocabulary, we found significant correlations with verb placement, word recall, and high-probability nonword recall in this group. There was also a significant correlation between the two nonword recall measures in this group.

Table 4 shows the correlations between the grammar and memory measures with children's vocabulary scores partialled out for the two learner groups.

[Insert Table 4 about here]

Again, no significant correlations between grammar and the memory measures were found for the L1 group. For the L2 children, correlations between the memory tasks remained positive and significant, even when vocabulary was controlled (*r*s between .42 and .48). The correlations between grammar and phonological memory also remained largely the same when vocabulary was partialled out: All three grammar measures correlated significantly with word recall, (*r*s between .39 and .54), the strongest correlations being for subject-verb agreement ($r = .54$) and auxiliary use ($r = .50$). The three grammar measures also showed moderate correlations with low-probability nonword recall, reaching significance for two out of the three measures: subject-verb agreement ($r = .43$) and verb placement ($r = .35$). Neither subject-verb agreement nor verb placement correlated significantly with high-probability nonword recall. Auxiliary production, in contrast, correlated significantly with high-probability nonword recall ($r = .52$), but not with low-probability nonword recall.

The finding that low-probability nonword recall correlates more strongly with subject-verb agreement and verb placement than high-probability nonword recall is difficult to interpret. Why would recall of words and low-probability nonwords, but not recall of high-probability nonwords, be associated with the production of L2 grammatical structures? One possibility is that even though at the group level, there was an effect of phonotactic probability on L2 children's nonword recall, some children did not show an advantage for high-probability versus low-probability nonwords. More precisely, these children may not have benefited from phonotactic knowledge about Dutch in their recall of high-probability items due to their low Dutch proficiency, thereby not making the high- and low-probability nonword recall a valid distinction between tasks for these children. In fact, some children's insensitivity to Dutch phoneme distributions may have obscured the correlations between grammar and the nonword recall measures in the entire sample.

To investigate this possibility, we examined the L2 children's individual accuracy scores on the two nonword recall tasks, and found that seven children did not repeat high-probability nonwords more accurately than low-probability nonwords. As shown in Table 5, this subset of children also obtained relatively low scores on the word recall task as compared to the children who did show an effect of phonotactic probability on nonword recall.

[Insert Table 5 about here]

We then investigated how grammar scores related to phonological memory for only those L2 children who repeated high-probability nonwords more accurately than low-probability nonwords. The results of these analyses are presented in Table 6.

[Insert Table 6 about here]

In this more restricted sample, correlations with word recall are still highest and significant for all three grammar measures, also after Dutch vocabulary is controlled (*r*s between .44 and .56). Correlations with high-probability word recall are now significant for subject-verb agreement (r = .42) and moderate but non-significant for auxiliary use and verb placement (*r*s = .34 and .36). Correlations with low-probability nonword recall are weakest, with the exception of a moderate correlation with subject-verb agreement (r = .37).

To see how much of the variance in children's production of subject-verb agreement, auxiliaries, and verb placement was accounted for by the three phonological memory measures over and above vocabulary, three hierarchical multiple regression analyses were performed. In the first analysis, we investigated to what extent word recall, high-probability nonword recall, and low-probability nonword recall predicted children's use of correct subject-verb agreement, controlling for Dutch vocabulary. Then, we did two similar analyses with auxiliary use and verb placement scores as the dependent variables.

For all three analyses, the order of entry of the memory scores was based on the idea that these scores reflect varying degrees of support from long-term linguistic knowledge. The variable that is most semantically contaminated, as it were, by long-term lexical knowledge (word recall) was entered first, followed by the variable that lacked semantics, but tapped knowledge about frequent phoneme combinations (high-probability nonwords), and finally, the purest measure of phonological memory, for which no or only minor support is available from long-term knowledge (low-probability nonwords). The results, presented in Table 7, are based on the restricted sample of children who performed more accurately on high- than low-probability nonword recall.

[Insert Table 7 about here]

Vocabulary, entered first in the model, accounted for a significant amount of variance in children's production of correct subject-verb agreement (R^2 change = .16, p = .034) and verb placement (R^2 change = .14, p = .050), but not auxiliary use (R^2 change = .09, p > .10) Word recall, entered in the second step, accounted for a significant amount of additional variance in all three models (R^2 change = .26, p = .002 for agreement; R^2 change = .20, p = .013 for auxiliary use; R^2 change = .19, p = .022 for verb placement). When controlling for word recall, high-probability nonword recall, entered in the third step, explained some additional variance in both auxiliary use and verb placement but this was not significant. Finally, when adding low-probability nonword recall in a final step, no additional variance was explained in subject-verb agreement and verb placement, and for auxiliary use, the variance was even negative. Together, vocabulary and the three phonological memory measures accounted for 44%, 38%, and 36% of the variance in children's production of correct subject-verb agreement, auxiliaries, and correct verb placement, respectively.

The result that the nonword recall measures do not explain much variance after word recall may not be surprising given that word recall encompasses the phonological memory skills tapped by nonword recall. In order to see if the nonword recall measures explained significant variance if these were entered first, we reversed the order of entry of the memory variables in the model. So, low-probability nonword recall was entered first, followed by high-probability nonword recall, and in the last step, word recall. The results of these analyses are presented in Table 8.

[Insert Table 8 about here]

These results show that low-probability nonword recall is a marginally significant predictor of subject-verb agreement, but not of auxiliary use and verb placement, if it is entered together after Dutch vocabulary in the model (R^2 change = .12, p = .049). High-probability nonword recall is a significant predictor of auxiliary use, but not of subject-verb agreement and placement, if it is entered together after vocabulary and high-probability recall (R^2 change = .14, p = .040). In the final models, with all variables entered, word recall is the only significant predictor of children's grammar scores next to vocabulary, as we saw above.

Discussion

This study examined the relationships between phonological memory and the acquisition of three specific L2 grammar skills in Turkish child learners of Dutch. Data of Dutch monolingual children were also reported, but only served to compare performance on the tasks, since – due to ceiling performance on the grammar measures – relationships between phonological memory and grammar could not be investigated for the L1 data.

Previous studies on the relationship between phonological memory and L2 grammar have typically included general measures of grammatical proficiency, and effects of vocabulary were not always controlled. In the current study, we investigated whether: (1) there were significant relationships between phonological memory and L2 subject-verb agreement, auxiliaries and verb placement, (2) effects of phonological memory remained after controlling for Dutch vocabulary knowledge, and (3) relationships varied for three different measures of phonological memory. We predicted that there would be significant relationships for all three measures. Given the contradictory findings in earlier studies, it was an open question whether effects of phonological memory would be unmediated by vocabulary skill. Dutch word recall was predicted to show the strongest correlations, as this measure taps phonological storage of elements for which long-term L2 support is available.

For the L2 children, we found strong and significant correlations between word recall and children's production of grammatical structures, and more moderate (but often still significant) correlations for the two nonword recall measures. Regression analyses indicated that, together with vocabulary, the three memory measures explained a considerable amount of significant variance in children's use of subject-verb agreement, auxiliary use, and verb placement (35%, 38%, and 36%, respectively). After controlling for vocabulary, word recall turned out to be the strongest predictor of all three grammar measures. In reversed analyses, low-probability nonword recall also explained some significant variance, in subject-verb agreement, above and beyond vocabulary, and was a marginally significant predictor of verb placement. High-probability nonword recall was a significant predictor of auxiliary production when it was entered together with vocabulary and low-probability nonword recall.

The current results fit well earlier studies showing significant relationships between phonological memory and L2 grammar (French, 2006; French & O'Brien, 2008; O'Brien et al,. 2006; Paradis, 2011; Service, 1992; Service & Kohonen, 1995), and show that these relationships hold for various grammatical sub-skills involving the production of bound and free morphemes as well as word order. Overall, the pattern of results looked similar for the different sub-skills. This is in line with views that assume that phonological memory is important for learning and generalizing from grammatical structures (Ellis & Sinclair, 1996; Speidel, 1989, 1993) both at the morphological and syntactic level. However, in regression analyses, the two nonword recall tasks were differentially related to subject-verb agreement and auxiliary use. Future research could explore this finding further to see if specific grammar skills rely on the phonological storage of more or less L2-like elements.

These significant relationships between phonological memory and grammar remained after vocabulary was controlled. This supports results from French and O'Brien (2008), who found that effects of phonological memory on the acquisition of a range of grammatical structures were independent of vocabulary in French child learners of English. In fact, in our regression analyses, word recall turned out to be a stronger predictor than Dutch vocabulary for all three grammar measures. Again, this is very similar to the results in French and O'Brien, who found that phonological memory was a much stronger predictor of grammar than vocabulary. This overlap in results is striking, given that grammatical knowledge was assessed in a different way in both studies. French and O'Brien used a written, discrete-point test in which children had to select an answer from multiple responses, fill in missing responses, or rewrite sentences. The current study used a production task assessing children's ability to produce grammatical structures. Also, the children in French and O'Brien were classroom learners receiving explicit instruction in the L2, while the current L2 learners where minority children acquiring their L2 in a naturalistic setting.

Out of our phonological memory measures, we found that, as we had predicted, the most language-dependent measure (word recall) correlated most strongly with children's grammar scores. This fits well with results by Parra et al. (2011) showing language-specific relationships between phonological memory and grammar in native bilingual children. The results also align with findings by Messer et al. (2010) showing that relationships with L2 vocabulary in child L2 learners were stronger for high-probability than low-probability nonword recall.

Another way of interpreting our results concerns the characteristics of the word recall task. While we can rule out that the relationships between word recall and grammar were due to a common association with vocabulary, there are clear drawbacks to using word recall as a measure of phonological memory. As mentioned in the introduction, the task has generally been considered a task of phonological memory (Compoy & Baddeley, 2008), but it may also involve reliance on semantic strategies, as reflected in semantic errors and concrete nouns being recalled better than abstract ones (Caza & Belleville, 1999). If the current participants relied on semantic strategies, this is a serious threat to the idea of a relationship between phonological memory and grammar, as argued here. However, we think it is safe to conclude that the word recall task in the current study and sample assessed phonological memory, at least to some degree, for two reasons. First, our data showed that word recall was a significant predictor of all three grammar measures in the regression analyses, above vocabulary, which also was a significant predictor of two of the three grammar measures. Thus, both made independent and significant contributions to L2 grammar. Second, correlations between word recall and the nonword recall measures in the L2 group were stronger than between word recall and vocabulary, suggesting that word recall, at least in part, taps the same skill as the nonword measures. Third, previous studies have shown that phonological errors (e.g., *rut* instead of *rug*) are more common than semantic errors in serial word recall (Jefferies et al., 2004). Earlier work also indicates that phonological encoding is the predominant strategy and that only under certain conditions (i.e., slow stimulus presentation or specific task instructions), subjects switch to semantic encoding (Campoy & Baddeley, 2008; Logie, Della Sala, Laiacona, Chalmers, & Wynn, 1996). In the current study, an effort was made to ensure that the words were not too abstract or difficult for L2 children, which makes it unlikely that word familiarity played a major role. Yet, correlations in the L2 group showed a slightly different pattern from those in the L1 group, where performance on word recall correlated more strongly with vocabulary than with performance on the nonword recall tasks. This suggests that the L1 children may have depended more on semantic strategies than phonological ones when performing the task. However, while this difference between L1 and L2 children suggests a potential interesting area for further research, it is not considered problematic for the current study, as our focus was on associations between phonological memory and grammar in the L2 group.

The finding that recall of L2 words was the strongest predictor of children's grammar scores supports earlier findings of Parra et al. (2011) and Messer et al. (2010), but goes against the findings in French and O'Brien (2008) who found that Arabic word repetition was a stronger predictor than English (or L2)-like nonword repetition in French learners of English. One possible explanation of this contrast in findings is that the participants in French and O'Brien's study had not built up enough long-term L2 knowledge for benefits in repeating English-like stimuli to show up. Support for this idea comes from very high correlations between their English-like and Arabic task ($r = .94$ at time 1; $r = .88$ at time 2) suggesting that – as is also noted by the authors – the tasks may actually have assessed the same construct (cf. French & O'Brien, p. 480). Also, correlations between English-like repetition and L2 English grammar were stronger at the second than at the first measurement, which may suggest that, at least at the first measurement, children may have been of too low a proficiency in English to show an advantage in repeating English-like (vs. Arabic) stimuli.

In our analyses on the whole sample, stronger associations were found for low-probability nonword recall than for high-probability nonword recall for two out of the three grammar measures. The different outcomes for the more restricted sample suggest that different results may be obtained depending on whether nonword recall or repetition accuracy is influenced by phonotactic probability in the L2. French and O'Brien (2008) do not report whether the child L2 learners in their study actually showed an advantage for English-like nonwords as compared to Arabic words. Assuming that some of them did not, the question arises whether different results had been obtained in French and O'Brien's study if the analysis had been restricted to those children who responded more accurately to the English-like than to the Arabic task.

A related question is how relationships between phonological memory and grammar knowledge relate to long-term L2 knowledge (or L2 linguistic proficiency).

French and O'Brien (2008) found stronger correlations between phonological memory and the recall of English-like nonwords at the second measurement than at the first, when children had become more proficient in the L2. Other studies found that relationships between phonological memory and L2 grammar learning only held for more advanced L2 learners and not beginning learners (Kormos & Sáfár, 2008; O'Brien et al., 2006). O'Brien et al. (2006), for example, found effects of phonological memory, assessed through serial recognition of L1-like nonwords, on the production of function words for the more proficient learners in their sample, but not for less proficient learners. The authors propose that this may be due to beginning L2 learners being more dependent on their memory resources for lexical storage and retrieval. When these learners become more proficient, memory resources become available for the storage of morphemes that are poor in semantic content such as agreement affixes. Future research could address the role of L2 proficiency in relation to effects of phonological memory on specific grammar skills.

One factor that may have influenced our results is specific properties of the tasks used. First, a narrative production task was used to assess L2 grammatical proficiency, whereas other L2 tests may have been more appropriate: In a relatively free task such as the current one, learners may have avoided the production of structures they found difficult. Second, vocabulary was measured through a receptive task, grammar was assessed productively, and phonological memory was tested through tasks that rely on receptive as well as productive skills. Perhaps, correlations would have been stronger if only receptive or only productive measures had been used, as relationships may vary with task modality. In speech production, parts of the utterance have to be stored in memory while or before planning and articulating the other parts (Kormos & Sáfár, 2008). Listening involves less active planning and no articulatory processes, and writing allows for elaborate reflection and/or correction, either of which may rely on phonological memory as well as more complex memory skills. An indication that different parts of verbal working memory may be differentially related to task modality comes from Révész, (2008). Révész found that adult L2 learners with higher phonological memory skills improved more on an oral test, whereas learners with higher complex memory skills improved more on a written grammar test.

Conclusion

Taken together, the results of this study indicate that the acquisition of grammar by child L2 learners involves storage-mediated learning, just as the acquisition of words (Ellis & Sinclair, 1996; Speidel, 1989, 1993). Such relationships are found for the production of bound and free function morphemes as well as word order, and are independent of L2 vocabulary skill. However, they differ depending on the type of phonological memory task that is used and, to some extent, also on the grammar skills investigated. Future research involving larger samples, different tasks, and different L2 populations is needed to further examine the precise relationships between phonological memory and L2 grammar learning and thereby shed more light on the specific processes underlying phonological storage that could drive these relationships.

References

Adams, A. M., & Gathercole, S. E. (1996). Phonological working memory and spoken language development in young children. *Quarterly Journal of Experimental Psychology*, *49A*, 216-233.

Adams, A. M., & Gathercole, S. E. (2000). Limitations in working memory: Implications for language development. *International Journal of Language Communication Disorders*, *35*, 95-116.

Alloway, T. P. (2007). *Automated Working Memory Assessment*. London: Psychological Corporation.

Archibald, L. M. D., & Gathercole, S. E. (2007). Nonword repetition and serial recall: Equivalent measures of verbal short-term memory? *Applied Psycholinguistics*, *28*, 587-606.

Atkins, W. B., & Baddeley, A. D. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics*, *19*, 537–552.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database* (CD-ROM). Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

Baddeley, A.D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*, 158-173.

Blake, J., Austin, W., Cannon, M., Lisus, A., & Vaughan, A. (1994). The relationship between memory span and measures of imitative and spontaneous language complexity in preschool children. *International Journal of Behavioral Development*, *17*, 91–107.

Blom, E. (2008). Testing the domain-by-age model: Inflection and placement of Dutch verbs. In B. Haznedar & E. Gavruseva (Eds.), *Current trends in child second language acquisition: A generative perspective* (Language acquisition and language disorders, 46) (pp. 271-300). Amsterdam: Benjamins.

Brown, G. D. A., & Hulme, C. (1995). Modeling item length effects in memory span: No rehearsal needed? *Journal of Memory and Language*, *34*, 594–621.

Campoy, G., & Baddeley, A.D. (2008) Phonological and semantic strategies in immediate serial recall. *Memory*, *16*, 329-340.

Cheung, H. (1996). Nonword span as a unique predictor of second-language vocabulary learning. *Developmental Psychology*, *32*, 867–873.

Daneman, M., & Case, R. (1981). Syntactic form, semantic complexity, and short-term memory: Influences on children's acquisition of new linguistic structures. *Developmental Psychology*, *17*, 367-378.

Ellis, N. C., & Sinclair, S. (1996). Working memory in the acquisition of vocabulary and syntax. *Quarterly Journal of Experimental Psychology*, *49*, 234-250.

Engel de Abreu, P. M. J., & Gathercole, S. E. (2012) Executive and phonological processes in second-language acquisition. *Journal of Educational Psychology*, *104*, 974-986.

French, L. M. (2006). *Phonological working memory and second language acquisition: A developmental study of Francophone children learning English in Quebec*. Lewiston, NY: Edwin Mellen Press.

French, L. M., & O'Brien, I. (2008) Phonological memory and children's second language grammar learning. *Applied Psycholinguistics*, *29*, 463-487.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*, 513-543.

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language, 29*, 336-360.

Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D. (1991). The influences of number of syllables and wordlikeness on children's repetition of nonwords. *Applied Psycholinguistics*, *12*, 349-367.

Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology*, *28*, 887–898.

Haznedar, B. (2001). The acquisition of the IP system in child L2 acquisition. *Studies in Second Language Acquisition*, *23*, 1-39.

Ionin, T., & Wexler, K. (2002). Why is 'is' easier than '-s'?: Acquisition of tense/agreement morphology by child second language learners of English. *Second Language Research*, *18*, 95-136.

Jefferies E., Lambon Ralph, M. A., & Baddeley A. D. (2004). Automatic and controlled processing in sentence recall: The role of long-term and working memory. *Journal of Memory and Language*, *51*, 623-643.

Logie, R. H., Della Sala, S., Laiacona, M., Chalmers, P., & Wynn, V. (1996). Group aggregates and individual reliability: The case of verbal short-term memory. *Memory & Cognition*, *24*, 305-321.

Kormos, J., & Sáfár, A. (2008). Phonological short-term memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition*, *11*, 261-271.

Martin, K. I., & Ellis, N. C. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition*, *34*, 379–413.

Masoura, E. V., & Gathercole, S. E. (2005). Phonological short-term memory skills and new word learning in young Greek children. *Memory*, *13*, 422–429.

Messer, M. H., (2010). *Verbal short-term memory and vocabulary development in monolingual Dutch and bilingual Turkish-Dutch preschoolers*. Unpublished doctoral dissertation, Utrecht University, the Netherlands.

Messer, M. H., Leseman, P. P. M., 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*, 306-323.

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology*, *91*, 3-9.

O'Brien, I., Segalowitz, N., Collentine, J., & Freed, B. (2006). Phonological memory and lexical, narrative, and grammatical skills in second language oral production by adult learners. *Applied Psycholinguistics*, *27*, 377-402.

Orgassa, A. (2009). *Specific language impairment in a bilingual context: The acquisition of Dutch inflection by Turkish-Dutch learners.* Unpublished doctoral dissertation, University of Amsterdam.

Parodi, T. (2000). Finiteness and verb placement in second language acquisition. *Second Language Research*, *14*, 355-381.

Prévost, P. (2003). Truncation and missing surface inflection in initial L2 German. *Studies in Second Language Acquisition*, *25*, 65-97.

Révész, A. (2008). Working memory and the observed effectiveness of recasts on different L2 outcome measures. *Language Learning*, *62*, 93-132.

Scheele, A. F., Leseman, P. P. M., & Mayo, A. Y. (2010). The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics*, *31*, 117-140.

Service, E. (1992). Phonology, working memory, and foreign-language learning. *Quarterly Journal of Experimental Psychology*, *45A*, 21-50.

Service, E., & Kohonen, V. (1995). Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition? *Applied Psycholinguistics*, *16*, 155-172.

Sociaal en Cultureel Planbureau [SCP] (2007). *Jaarrapport integratie 2007 [Year report integration 2007]*. The Hague, the Netherlands.

Speidel, G. E. (1989). A biological basis for individual differences in learning to speak. In G. E. Speidel & K. E. Nelson (Eds.), *The many faces of imitation in language learning* (pp. 199-229). New York: Springer Verlag.

Speidel, G. E. (1993). Phonological short-term memory and individual differences in learning to speak: A bilingual case study. *First Language*, *13*, 69-91.

Speidel G. E., & Herreshoff, M. J. (1989). Imitation and the construction of long utterances. In G. E. Speidel & K. E. Nelson, (Eds.), *The many faces of imitation in language learning* (pp. 181-197). New York, Springer-Verlag.

Thorn, A. S. C., Gathercole, S. E., & Frankish, C. R. (2005). Redintegration and the benefits of long-term knowledge in verbal short-term memory: An evaluation of Schweickert's (1993) multinomial processing tree model. *Cognitive Psychology*, *50*, 133–158.

Van Kampen, J. (1997). First steps in WH-movement. Unpublished doctoral dissertation. Utrecht University.

Verhoeven, L., Narain, G., Extra, G., Konak, O. A., & Zerrouk, R. (1995). *Toets Tweetaligheid* [Test for Bilingualism]. Arnhem, the Netherlands: Cito.

Verhoeven, L., Steenge, J., & Van Balkom, H. (2011). Verb morphology as a clinical marker of specific language impairment: Evidence from first and second language learners. *Research in Developmental Disabilities, 32,* 1186-1193.

Vitevitch, M. S., & Luce, P. A. (2005). Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of Memory and Language, 52*, 193–204.

Williams, J. N., & Lovatt, P. (2003). Phonological memory and rule learning. *Language Learning, 53*, 67-121.

Table 1

Characteristics of the High- and Low-Phonotactic Probability (PP) Items in the Nonword Recall Tasks

	High-PP	Low-PP
Phonotactic probability (mean summated biphone frequency)	20.1	3.3
Likeliness rating (mean rating on a scale from 1 to 5)	3.6	2.6
Length (mean number of phonemes)	4.3	3.9

Note. These data are based on the first three blocks, since no child passed the third block (N = 36 items).

Table 2

Descriptive Statistics for the L1 and L2 Children

	L1 children				L2 children			
	Mean	(Nr obl. contexts)	*SD*	*N*	Mean	(Nr obl. contexts)	*SD*	*N*
Language (% correct)								
Subject-verb agreement	97.0	(29.8)	4.9	34	47.3	(16.9)	35.0	36
Auxiliaries	96.0	(10.2)	12.6	31	5.9	(6.3)	6.1	36
Verb placement	91.6	(18.5)	8.6	29	55.2	(12.0)	38.1	36
Dutch vocabulary	70.2	-	13.0	34	49.3	-	11.7	36
Phon. memory (nr. correct)								
Word recall	13.3	-	4.7	34	10.2	-	3.6	36
High-PP nonword recall	4.2	-	1.8	34	3.2	-	1.6	36
Low-PP nonword recall	2.7	-	1.7	34	2.1	-	1.3	36

Note. The scores for auxiliaries and verb placement in the L1 group are based on a smaller sample, because some children produced few analyzable utterances, due to their frequent use of the 'dummy' auxiliary *gaan* 'going to' (see also Van Kampen, 1997). (Nr obl. Contexts) = mean number of utterances analyzed; *% correct* = mean percentage correct; *nr. correct* = mean number correct; PP = Phonotactic Probability.

Table 3

Zero-order Correlations between Vocabulary, Language and Phonological Memory for the L1 and L2 Children (N = 34 and N = 36)

	1	2	3	4	5	6	7
1. Vocabulary	-	.23	-.23	.37*	.36*	.38*	.25
2. SV-agreement	.32*	-	-.09	.21	.22	-.26	-.05
3. Auxiliaries	.26	.38*	-	-.22	-.14	.06	-.06
4. Verb placement	.31*	.73**	.25	-	.05	.16	.16
5. Word recall	.34*	.51**	.48**	.39*	-	.23	.29
6. High-PP nwr	.28	.19	.50**	.12	.45**	-	.34*
7. Low-PP nwr	.14	.39*	.18	.32*	.41*	.48**	-

* $p < .05$, ** $p < .01$ *Note*. Correlation coefficients for the L1 group are shown in the upper triangle; correlation coefficients for the L2 group are shown in the lower triangle. PP = Phonotactic Probability; nwr = nonword recall.

Table 4

Partial correlations between Grammar and Phonological Memory (with Vocabulary Partialled Out) for the L1 and L2 Children (N = 34 and N = 36)

	1	2	3	4	5	6
1. SV-agreement	-	-.00	.19	-.02	-.28	-.21
2. Auxiliaries	.32*	-	-.21	-.23	-.16	.10
3. Verb placement	.71**	.47**	-	-.19	-.08	-.03
4. Word recall	.54**	.50**	.39*	-	.33*	.06
5. High-PP nwr	.19	.52**	.12	.45**	-	.34*
6. Low-PP nwr	.43*	.20	.35*	.42*	.48**	-

* $p < .05$, ** $p < .01$ *Note*. Correlation coefficients for the L1 group are shown in the upper triangle; correlation coefficients for the L2 group are shown in the lower triangle. PP = Phonotactic Probability; nwr = nonword recall.

Table 5

Descriptive Statistics for Memory Tasks for L2 Children Depending on Whether They Show a Phonotactic Probability (PP) Effect

	L2 children not showing probability effect			L2 children showing probability effect		
	Mean	SD	N	Mean	SD	N
Word recall	8.7	2.7	7	10.6	3.8	29
High-PP nonword recall	1.0	1.2	7	3.7	1.2	29
Low-PP nonword recall	2.3	1.4	7	2.1	1.4	29

Table 6

Correlations between Language and Phonological Memory for L2 Children Showing a Phonotactic Probability Effect (N = 29)

	1	2	3	4	5	6
1. SV-agreement	-	.44*	.67**	.53**	.32	.39
2. Auxiliaries	.37	-	.34*	.46*	.40*	.16
3. Verb placement	.66**	.29	-	.44*	.35	.36
4. Word recall	.56**	.46*	.44*	-	.40*	.43*
5. High-PP nwr	.34	.42*	.36	.40*	-	.66**
6. Low-PP nwr	.37	.19	.29	.43*	.66**	-

* $p < .05$, ** $p < .01$ *Note*. Zero-order correlation coefficients are shown in the upper triangle; correlation coefficients with Dutch vocabulary partialled out are shown in the lower triangle. PP = Phonotactic Probability; nwr = nonword recall.

Table 7

Results of Hierarchical Regression Analyses: Predicting Grammar from Phonological Memory, after Controlling for Vocabulary (N = 29)

	SV-agreement		Auxiliary verbs		Verb placement	
	B	ΔR^2	*B*	ΔR^2	β	ΔR^2
Step 1		.16*		.09		.14*
Dutch vocabulary	.39*		.30		.38*	
Step 2		.26**		.20*		.19*
Dutch vocabulary	.37*		.27		.38*	
Word recall	.51**		.44*		.44*	
Step 3		.01		.06		.03
Dutch vocabulary	.37*		.28		.38*	
Word recall	.46**		.34*		.37*	
High-PP nonword recall	.13		.26		.20	
Step 4		.01		.04		.00
Dutch vocabulary	.38*		.26		.38*	
Word recall	.44*		.39*		.35*	
High-PP nonword recall	.07		.42		.17	
Low-PP nonword recall	.11		-.26		.06	
R^2 total		.44**		.38*		.36*

* $p < .05$, ** $p < .01$. PP = Phonotactic Probability.

Table 8

Results of Hierarchical Regression Analyses: Predicting Grammar from Phonological Memory, after Controlling for Vocabulary – Reversed Order (N= 29)

	SV-agreement		Auxiliary verbs		Verb placement	
	B	ΔR^2	*β*	ΔR^2	*β*	ΔR^2
Step 1		.16*		.09		.14*
Dutch vocabulary	.39*		.30		.38*	
Step 2		.12*		.03		.09
Dutch vocabulary	.42*		.31		.39*	
Low-PP nonword recall	.34*		.18		.31	
Step 3		.01		.14*		.03
Dutch vocabulary	.41*		.29		.38*	
Low-PP nonword recall	.24		-.15		.16	
High-PP nonword recall	.16		.50*		.24	
Step 4		.15*		.12*		.10*
Dutch vocabulary	.38*		.26		.38*	
Low-PP nonword recall	.11		-.26		.06	
High-PP nonword recall	.07		.42		.17	
Word recall	.44*		.39*		.35*	
R^2 total		.44**		.38*		.36*

* $p < .05$, ** $p < .01$. PP = Phonotactic Probability.

Appendix

Stimuli items in the nonword recall tasks

	Low-probability	High-probability
Block 1		
1	Jimf	Zwag
2	Dwup	Grops
3	Pjoef	Zils
4	Fosk	Brof
5	Pifp	Traa
6	Faup	Gleg
Block 2		
1	Pjosr Fnup	Grigt Zwop
2	Fuup Pjif	Spraam Kwig
3	Vub Puif	Zifs Bropt
4	Fjaip Dzub	Greel Knit
5	Fip Posf	Knog Glin
6	Pgup Dwuuf	Ziks Glof
Block 3		
1	Mwup Fjif Njos	Zup Kjif Fjui
2	Ims Fwup Pjai	Zilg Brong Tris
3	Bnup Osf Fjeum	Snins Glirg Ceng
4	Fwut Gjuip Fimk	Fling Brops Zwis

5	Djai Pwut Fibs	Vlop Snilg Kwin
6	Zup Kjif Fjui	Zwit Snint Dromp

Note. Because no child exceeded Block 3, only the items of the first three blocks are listed.