Working memory training improves children’s syntactic ability but not vice versa: A randomized control trial

Ernesto Roque-Gutierrez, Paul Ibbotson

Open University, Milton Keynes MK7 6AA, UK

Article info

Article history:
Received 27 April 2022
Revised 20 September 2022
Available online xxxx

Keywords:
Language development
Working memory
Domain general
Cognitive transfer
Training

Abstract

We tested several hypotheses about the relation between syntax and working memory (WM). In a pretest/posttest randomized control trial, 104 native Cuban Spanish-speaking children (Mage = 7 years 2 months; 54 girls) took part in syntax training in their first language, syntax training in their second language, WM training, or no training (control). Compared with the control, children in the training conditions showed cognitive transfer from WM to syntax but not from syntax to WM. The result was most striking in the case of their first language, where WM training was as effective as language training in boosting syntactic performance. As well as establishing cognitive transfer at the group level, we also found that individual differences in WM performance, both at baseline and in training, predicted the extent to which children’s syntax improved. The directionality of transfer, the group-level and individual-level results, established in the context of a randomized control design, all point to a strong causal role for domain-general cognition in the processes of language acquisition.

© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
Introduction

Domain specificity

The status of human language as either an extensive recycler of general cognition (Bybee, 2010; Dąbrowska & Divjak, 2015, 2019; Goldberg, 2005; Ibbotson, 2020; Tomasello, 2003) or an encapsulated module with processes unique to itself (Chomsky, 1980; Fodor, 1983; Hauser et al., 2002; Pinker & Jackendoff, 2005) has been intensely contested. The debate is important because it speaks to what language is made of and how it is acquired. Both sides of the debate agree that language interacts with general cognition at some level; however, the extent of this interaction, the causal direction, the developmental periods in which it occurs, and the cognitive domains in which it happens, remain poorly understood.

The current study contributes to our understanding by testing whether transfer occurs between working memory (WM) and syntax. If transfer does not occur, then it suggests an impermeable boundary between these systems, which are informationally encapsulated, as a modular account would predict (Fodor, 1983). If transfer does occur, it would add to a growing body of research that shows deep integration between language development and general cognitive processes, including domain-general categorization processes and the acquisition of grammatical constructions (Goldberg, 2005; Ibbotson et al., 2012; Ibbotson & Tomasello, 2009), social cue use and grammatical subject acquisition (Ibbotson et al., 2013), inhibitory control and verb inflection (Ibbotson & Keavell-White, 2015), WM and the acquisition of passive constructions (Białecka-Pikul et al., 2016) executive function and the acquisition of parsing strategies and syntax (White et al., 2017; Woodard et al., 2016); and a parallel development of perceptual narrowing in both face and speech perception modalities (Krasotkina et al., 2021). Furthermore, because of our design—three treatment groups and a control—we are able to establish for the first time not just whether transfer is possible between these domains but also in which directions.

The broader theoretical implications of a positive transfer result are that the language-unique parts of language are in retreat to such an extent that it becomes more parsimonious to use the framework of general cognition to understand linguistic development rather than call on processes and structures unique to language (Ibbotson, 2020).

Cognitive transfer

The cognitive transfer methodology is well-suited to address the domain-general versus language-specific debate because it measures to what extent skills learned in one domain transfer to another. Cognitive transfer has proved both a subtle test of the interconnectivity of different components of the mind and a promising methodology for applied interventions. For example, in one of the biggest training studies to date with 17,648 children aged 6–8 years, Judd and Klingberg (2021) demonstrated positive transfer from spatial cognition training to mathematical learning (see also Bergman et al., 2011; Thorell et al., 2009).

With regard to WM specifically, the evidence for positive transfer is more mixed. Transfer has been observed in older adults (e.g., Borella et al., 2010), young adults (Jaeggi et al., 2008), children with attention-deficit/hyperactivity disorder (Klingberg et al., 2005); children with developmental language disorder (Delage et al., 2021; Stanford et al., 2019), and (most recently) children with autism spectrum disorder (ASD) (Delage et al., 2022). For example, Delage and colleagues (2022) assessed the impact of 12 h of WM training (simple and complex span tasks) across 8 weeks in 30 children with ASD aged 5 to 11 years. Results showed direct improvements on untrained WM tasks as well as transfer effects to syntax (root question accuracy and clitics), an effect that was still present 3 months after the posttest. Using a similar age group to the current study, Karbach et al. (2015) demonstrated cognitive transfer from WM training to reading but not to mathematics (see also Loosli et al., 2012). However, in a meta-analysis of 87 WM training studies with 145 experimental comparisons, Melby-Lervåg and Hulme (2016) concluded that “there is no evidence that working memory training convincingly produces effects that generalize to important real-world cognitive skills” (p. 525), and in a separate
meta-analysis of 24 WM training studies, Gathercole et al. (2019) reported no transfer between training on verbal short-term memory and visual–spatial WM or vice versa.

The lack of consensus, even between meta-analyses, is partly attributable to the variation in WM subcomponents targeted in the studies (e.g., verbal, nonverbal, visuospatial, phonological), methodological variation (e.g., control group, no control), the tests employed (e.g., n-back, complex span, Test of Memory and Learning), differences between the trained and untrained arms of the studies (e.g., memory to memory “near” transfer vs memory to arithmetic “far” transfer), and populations (e.g., typical, atypical, children, adults). Regardless of whether the studies show evidence in favor of WM transfer or not, to our knowledge no study has used the combination of measures we employed here: purely nonlinguistic WM training in combination with a measure of syntactic ability. We chose this combination precisely to address the question of the domain specificity of language. Previous studies reporting correlations between WM and language performance have tended to employ WM measures with some linguistic component such as being asked to read aloud a sentence while simultaneously comprehending another, and so they cannot rule-out language-to-language transfer (e.g., Atkins & Baddeley, 1998; Daneman & Carpenter, 1980, 1983; Masson & Miller, 1983). Using a nonlinguistic measure of WM, as we did here with geometric shapes, means that any transfer that does occur between WM and syntax cannot be attributable to language-to-language transfer.

Evidence of cognitive transfer between WM and syntactic ability in the current study would suggest that they share a common functional hierarchy—that is, they call on similar cognitive resources to get the job done (e.g., Carroll, 1993). Following previous approaches (Lövdén et al., 2010; Noack et al., 2009, 2014), we adopted a cognitive taxonomy to qualify transfer distance according to the highest branch in the hierarchy that must be passed to connect the trained and transfer tasks (Fig. 1).

WM and language

We focused on WM because of its importance in underpinning a wide range of complex cognitive tasks and developmental outcomes for children, including reading comprehension (e.g., Cain, 2006; García-Madruga et al., 1997), reasoning (e.g., Conway et al., 2003; García-Madruga et al., 2007), arithmetic calculations (e.g., Deschuyteneer et al., 2006), mathematical problem solving (e.g., Passolunghi & Pazzaglia, 2004), academic performance (e.g., Alloway & Alloway, 2010), and fluid intelligence (e.g., Friedman et al., 2006). The essential underlying cognitive process that supports this range of behavior is one that enables active maintenance and regulation of a limited amount of task-relevant information (Baddeley & Logie, 1999). The n-back test of WM we employed here taps into this ability by asking

![Fig. 1. Proposed hierarchical relationship between our main measures. Both the nonlinguistic and linguistic components engage working memory (WM) in order to succeed at the task. Our research design allowed us to measure the potential spillover that results from training one arm of this hierarchy into the untrained other arm, mediated by the common cognitive function of WM.](image-url)
participants to decide whether a stimulus was previously presented in certain conditions, taxing WM to hold some information “in mind” (i.e., the sequence of stimuli) while simultaneously performing a different task (i.e., do the stimuli match?) (Kirchner, 1958).

A particular aspect of language that is thought to especially tax WM is agreement. Broadly speaking, agreement refers to a grammatical system where words or morphemes must coordinate their form with one another, for example, in person (e.g., “I am” vs. “he am”), number (e.g., “they is happy” vs. “they are happy”), gender (e masculine gato masculine vs. la feminine gato masculine), or case (e.g., “He is kissing her” vs. “him is kissing she”). This type of grammatical coordination is an obligatory part of many of the world’s languages, and thus successful acquisition represents an important developmental milestone.

Our proposal here is that WM is recruited for successful number agreement between subject and verb precisely when intervening lexical material forces long-distance coordination. For example, in (1), the plurality of the subject tareas needs to be held in mind while simultaneously processing the intervening prepositional phrases before agreeing in number with the downstream son (a full list of test sentences is available in Supplementary Material 1).

(1) Las tareas de la clase de Lengua Española de primer grado son para mañana

Gloss: The assignments for the first-grade Spanish class are for tomorrow.

Thus, in common with many other WM training programs, both tasks involve continuous updating of the to-be-remembered items in lists of unknown length. This approach is inspired by modern frameworks of sentence comprehension (often using long-distance dependencies like the ones used in this study) that claim that language processing, although it might operate on specialized representations, is nevertheless subject to general processing principles and constraints that govern other domains of memory (Bever, 1970; Lewis, 1996; Lewis et al., 2006; Marcus, 2006).

Second language learners

Finally, our study differentiates the effects of cognitive transfer among first language (L1) and second language (L2) processing. There are several reasons why we might expect children engaging either their L1 or L2 resources to differ in how they parse the linguistic elements in our study. These include prior evidence of L2 speakers’ greater reliance on shallow processing strategies driven by a heavier reliance on semantic and pragmatic information than L1 speakers (e.g., Clahsen & Felser, 2006); L2 speakers’ reduced ability to integrate syntactic and other information sources, including executive function and WM (Sorace, 2011); L2 speakers’ greater reliance on declarative memory in comparison with a more balanced use of both declarative and procedural memory in L1 speakers (Ullman, 2015); and L2 speakers’ greater susceptibility to similarity-based retrieval interference (Cunnings, 2017). In terms of any global cognitive advantage for L2 speakers, the bilingual advantage hypothesis remains a contested idea for executive function in general and even more so for WM specifically (Lehtonen et al., 2018; De Groot, 2013; Bialystok et al., 2004; Paap, 2019). We motivate the analysis of L1/L2 performance here not as a direct test of the bilingual advantage hypothesis or any other specific theory of

| Table 1 |
| Socioeconomic context of the participants’ schools |
| School | Britannia 1 | Britannia 2 |
| Municipality | 10 de Octubre | Centro Habana |
| Median salary of governmental jobs in Cuban pesos | 664 pcm | 610 pcm |
| Size sample | 42 (25 girls and 17 boys) | 62 (29 girls and 33 boys) |
| Mean age (years;months) | 7;2 | 7;2 |
sentence parsing but rather as a further condition that may differentiate transfer effects between two populations who differ in their familiarity with language and how they process it.

The current study: Hypotheses and predictions

The overall goal here was to understand more about the way in which WM and syntactic abilities interact in development and draw inferences about the domain specificity of language. Consequently, the current study explored four possible hypotheses about the way in which WM and syntactic training may affect performance: As a result of training, (1) WM improvements transfer to improved syntactic ability, (2) improved syntactic ability transfers to improved WM performance, (3) improvements in WM and syntax transfer in both directions, and (4) improvements in WM and syntax transfer in neither direction. Given the evidence reviewed regarding the relationship between WM and language, one might predict that participants show some level of transfer (Hypotheses 1–3); however, the equivocal literature on transfer in general makes no transfer (Hypothesis 4) entirely possible also. Whatever the outcome, we argue that if we observe transfer, the most parsimonious explanation is that there is a close functional mapping between the kinds of cognitive processes required to succeed in the WM task and the subject–verb agreement task.

Method

Participants

Of the 116 children initially contacted, 4 were excluded before the main phase of the experiment took place (1 declined and 3 missed the pretest); of those that entered the training phase, 6 did not complete (see Fig. 2 for breakdown across groups); and of those that completed the training phase, 2 missed the posttest. This resulted in 104 children available for final analysis (M_age = 7;2 [years; months], range = 6;4–7;9, SD = 5.33 months; 54 girls and 50 boys).

This age group was chosen because, first, WM is still undergoing significant development during this period (Spencer, 2020). At the same time, WM had developed to a point where it was robust enough to take part in the standardized, and cognitively demanding, memory training task. Second, language competency had developed to a point where we could be sure that participants could successfully engage with what was being asked of them yet still produce a significant amount of errors such that if cognitive transfer did take place, there would be enough room below ceiling performance for that to be detected. This performance level was established via a pilot study using a subset of the test sentences on a smaller group of children (N = 24, M_age = 7;0, range = 73–89 months) that resulted in a mean baseline target sentence accuracy of 43.06% for L1 and 52.78% for L2. None of the participants who took part in the pilot took part in the main study.

Participants were recruited from two private English language schools: Britannia 1 (n = 42; M_age = 7;2; 25 girls and 17 boys) and Britannia 2 (n = 62; M_age = 7;2; 29 girls and 33 boys) located in two different municipalities of Havana, Cuba. All children aged 6 and 7 years enrolled at the time were considered eligible to be part of the study. At these language schools, children attended two 60-min English classes per week in addition to their weekly or twice-weekly English classes at their regular school. Although exposure to English has become more widespread in the daily lives of Cubans, mostly through American music, films, and TV shows and especially after the collapse of the Soviet Union during the early 1990s, the level of English proficiency throughout the country is generally low. Participants’ proficiency in English was estimated at the upper end of the pre-A1 level based on the Common European Framework of Reference for Languages and the latest descriptors of language competences for learners in the 7- to 10-year age range (Szabo, 2018). This means that participants can recognize and use simple principles of word order in short statements and can understand simple statements and questions delivered face to face.

Both private language schools are part of the same academic network under the same management, so they follow similar pedagogical approaches and have shared principles, and their teachers and students often collaborate on classes and projects. Tuition fees are standard for both schools at
250 Cuban pesos per month, which means that access is limited to those parents who can afford this (Table 1).

Upon giving their consent, parents were briefly interviewed about their child’s experience with English. Although the proficiency levels of all students were broadly comparable by virtue of their placement test scores in the language school, we wanted an additional measure of L2 exposure in case this affected the learning outcomes of the training in any way. To provide a rough estimate of participants’ exposure to English, parents provided answers to the following questions. Has the child received any English lessons in the past? When and for how long?; Aside from their normal classes, would you say your child is exposed to English at home? Could you think of examples of activities that involve English (e.g., music, video games, films, listening to a family member)? How many hours a week does your child spend engaged in activities that involve English?; Do you or any other members of the family communicate with the child in English? How frequently? (every day, 2–3 times a week, once a week, once a month, or never); In an average week, how many hours is your child exposed to someone who speaks to them in English? Based on the answers to these questions, the child’s expo-
sure to English was estimated in months and was used in the main analysis, along with gender and age, as a covariate.

The study was approved by the human research ethics committee at The Open University (Reference: HREC/3368/Roque-Gutierrez). Further details appear in Appendix A.

Design

To assess baseline performance, all participants took part in the same three pretests: L1 subject–verb agreement tasks, L2 subject–verb agreement tasks, and a WM test (more on the specifics of these is provided in the “Assessment and training stimuli” section below). Using randomized block assignment, participants were assigned to one of four groups based on how they scored in the pretest. By counterbalancing scores across groups on these key outcome measures, we could ensure that whatever participants achieved in their posttest results, the initial groups had comparable starting points. During the subsequent training phase, Group 1 received only subject–verb agreement training in Cuban Spanish, Group 2 received only subject–verb agreement training in English, Group 3 received only nonlinguistic WM training, and Group 4 received no training (control); see Fig. 2 for the overall design structure and the number of participants falling into each group. The same measures were administered at posttest as were administered at pretest to all groups, namely L1 subject–verb agreement tasks, L2 subject–verb agreement tasks, and a WM test.

Testing and training schedule

Data were collected from October 17 to November 28, 2019. Baseline pretest measurements were collected over a 3-day period for all children across the two participating schools. After block assignment, training began. The intensity and duration of the WM training was based on the median intervention hours per week established by Soveri et al. (2017) in a meta-analysis of 33 n-back studies: 6.67 h over 15 sessions. We then used this to match the duration of the linguistic training such that across all training groups (1–3) children received two 25-min sessions of training per week, delivered during normal school hours, for the first 2 weeks and three sessions of the same duration per week for the following 4 weeks, 16 (2*2+3*4) of 25 min spread out over 6 weeks for a total of 6.67 h per week. During this period, the control group received teaching as normal consisting of their regular classes twice a week with the same teacher with no extra linguistic or memory training. After a total of 6 weeks of training, the same posttest measures were administered to all children remaining in the study.

Assessment and training stimuli

L1 and L2 stimuli were subject to several controls that tried to match the languages as far as possible on as many variables as could significantly affect the results, including lexical frequency, local agreement and semantic plausibility, dependency length, and phonetic distinctiveness. The details of all stimuli controls can be found in Appendix B.

Presentation

Sentences were read aloud so that all children could move through the tasks at the same pace. Sentences were presented as a binary grammaticality judgment, and the children’s task was to circle either a tick (correct) or a cross (incorrect) on a piece of paper in response to the question “Is the sentence correct or incorrect?” As with the WM test phase, no feedback was given as to the accuracy of the children’s answers. This tick–cross procedure was familiar to the children from their course textbooks, and the numbers of correct/incorrect responses to target sentences were later recorded by the experimenter. Half the target sentences agreed in subject (number) and verb and the other half did not. Of those that did not agree, half were singular subject/plural verb mismatches and half were plural subject/singular verb mismatches.
The target sentences were randomly interspersed with the same number of fillers to avoid boredom, rehearsal and shallow processing strategies, and they consisted of two types: (a) sentences that contained no long-distance dependencies (e.g., “The cat is under the bed in David’s room”) to offer respite from the WM-heavy sentences and keep the children engaged and (b) yes/no questions regarding the truth of the preceding statement (e.g., “Is the cat under the bed?”) to encourage children to process the content, including the embedded clause, of all the sentences (because the order of the yes/no questions could not be predicted).

Working memory

We used the well-known \( n \)-back paradigm administered on a laptop to assess and train WM capacity (Kirchner, 1958). Children were presented with a series of visual nonlinguistic shapes, and their task was to determine whether it matched a shape \( n \) trials before. For example, in a 2-back task, children needed to decide whether the current shape is the same as the shape in trial \( n - 2 \). If children thought they had detected a match, they pressed the letter “A” on the laptop. If they thought there was no match, they were not required to respond. All children took part in a practice trial (1-back) so that they understood the rules of the game. During this phase, they received feedback on the accuracy of their answers. For each subsequent test trial, 20–22 shapes were selected from a pool of 8 irregular polygons (Jaeggi et al., 2010) to create a randomly ordered sequence (Fig. 3).

Only one shape appeared on the screen at a time for 2.7 s, followed by a 0.3s gap (no shape), after which the next shape was displayed and so on until the sequence was finished. For each WM session (whether in pretest, posttest, or training), the children completed six trials in total, taking approximately 25 min to complete with no feedback given as to the accuracy of their answers. Target correct matches represented approximately 28% of the presented stimuli. Correct responses were recorded if participants pressed the “A” key when the image repeated itself \( n \) positions back (true positive) and if they not press “A” when it did not (true negative). Incorrect responses were recorded if participants responded to a distractor (false positive) and if they failed to respond to an \( n \)-back target (false negative). The overall score was given by the true responses (accuracy) as a proportion (%) of the total stimuli shown to participants (Fig. 4).

For those children allocated to the memory training condition, where they would be engaging with this task for many weeks, we wanted to ensure that there was scope for development as their perfor-

---

**Fig. 3.** Pool of eight irregular polygons used for working memory assessment and training.
mance improved. This would increase the chances of training a strong WM capacity over time. For this reason, when children’s performance reached a threshold (~80% correct) on the 1-back test, the children progressed to the 2-back test.
Results

The results are organized according to how the different training groups performed on the outcome measures of interest: L1, L2, and WM. The most revealing posttest contrasts are between training groups and control—for the reason that we might expect performance to improve with time, disregarding condition. Therefore, any significant differences between posttest training groups and control are also significant differences between posttest and pretest. This significantly reduces the amount of pretest and posttest comparisons needed to the minimum required to answer the research question while still establishing change/no change over time.

Age, gender, and second language exposure (see Method) were included as covariates, all of which could plausibly account for some variance in the outcome. Significant differences between groups are reported after controlling for covariate variance. Finally, as a reminder of our measures, our independent variable had four levels—L1, L2, and WM training and control group—and our outcome variable was accuracy on tests of syntax and WM.

**L1 measure**

**Pretest versus posttest**

A paired-samples $t$ test was conducted comparing the percentage correct L1 score between pretest and posttest control group. No significant difference was found between pretest ($M = 61.53, SD = 7.45$) and posttest ($M = 61.73, SD = 6.31$), $t(25) = -0.15, p = .88$.

**L1 pretest**

To confirm that counterbalancing (see Method) had resulted in groups that were matched for initial WM and Spanish skills, a one-way independent analysis of covariance (ANCOVA) with three levels was performed on L1, WM, and control, with age, gender, and L2 exposure as covariates and L1 pretest performance as our dependent variable. There was no main effect of group, $F(2, 72) = 0.104, p = .901, \eta^2 = .003$, confirming that our groups started from the same position.

**L1 posttest**

To find out whether 6 weeks of WM and syntax training resulted in cognitive transfer, a one-way independent ANCOVA with three levels was performed on L1, WM, and control, with age, gender, and L2 exposure as covariates and L1 posttest performance as the dependent variable (Fig. 5). There was a significant main effect of group, $F(2, 72) = 17.67, p < .001, \eta^2 = .329$. Pairwise comparisons with Sidak correction revealed a significant difference between L1 training ($M = 77.50, SD = 12.90$) and control ($M = 61.73, SD = 6.31$), $p < .001$, 95% confidence interval (CI) [7.87, 20.30], and between WM training ($M = 74.23, SD = 7.02$) and control, $p < .001$, 95% CI [5.35, 17.24]. No significant differences were found between L1 training and WM training on L1 scores. Covariate analysis revealed a significant effect of age, $F(2, 72) = 5.61, p = .021, \eta^2 = .072$, and no significant effect of gender or L2 exposure.

**Comparison with chance**

The pretest and posttest comparisons indicate where there is a significant change in performance over time; they do not demonstrate that this change led to performance that was significantly different from chance (50% correct). To demonstrate this, we conducted a one-sample $t$ test (test value = 0.5) comparing L1 training group posttest performance on L1 with chance, $t(25) = 10.86, M = .77, p < .001, d = .12$, and comparing WM training group posttest performance on L1 with chance, $t(25) = 17.58, M = .74, p < .001, d = .07$. This shows that the effect of training had led to performance that was significantly different from chance.
L2 measure

Pretest versus posttest

A paired-samples t test was conducted comparing the percentage correct L2 score between pretest and posttest control group. No significant difference was found between pretest ($M = 56.61, SD = 7.51$) and posttest ($M = 55.65, SD = 6.67$), $t(25) = 0.79$, $p = .43$.

L2 posttest

As was the case with L1, a one-way independent ANCOVA with three levels was conducted on L2, WM, and control, with age, gender, and L2 exposure as covariates and L2 pretest performance as the dependent variable. There was no main effect of group, $F(2, 72) = 0.42$, $p = .959$, $\eta^2 = .001$, confirming that our groups started from the same position.

L2 measure

Pretest versus posttest

To determine whether similar transfer effects to those found in L1 took place in L2, we carried out a three-way independent ANCOVA with three levels on L2, WM, and control, with age, gender, and L2 experience as covariates and L2 posttest performance as our dependent variable (Fig. 6). There was a significant main effect of group, $F(2, 72) = 40.11$, $p < .001$, $\eta^2 = .527$. Pairwise comparisons with Sidak correction revealed a significant difference between L2 training ($M = 68.63, SD = 5.76$) and control.
(M = 55.65, SD = 6.67), p < .001, 95% CI [8.96, 15.70]), between WM training (M = 62.86, SD = 3.78) and control, p < .001, 95% CI [3.27, 9.87], and between L2 training and WM training, p < .001, 95% CI [2.47, 9.04]. Covariate analysis revealed a significant effect of L2 exposure, F(2, 72) = 18.69, p < .001, η² = .272, and no significant effect of age or gender.

**Comparison with chance**

To demonstrate that training led to performance that was significantly different from chance, we conducted a one-sample t test (test value = 0.5) comparing L2 training group posttest performance on L2 with chance, t(25) = 16.49, M = .68, p < .001, d = .05, and comparing WM training group posttest performance on L2 with chance, t(25) = 17.33, M = .62, p < .001, d = .03. This shows that the effect of training had led to performance that was significantly different from chance.

**WM measure**

**Pretest versus posttest**

A paired-samples t test was conducted comparing the percentage correct WM score between pretest and posttest control group. No significant difference was found between pretest (M = 71.70, SD = 9.79) and posttest (M = 71.80, SD = 6.61), t(25) = –0.15, p = .88.

**WM pretest**

A one-way independent ANCOVA with four levels was conducted on L1, L2, WM, and control as the main group variables, with age, gender, and L2 exposure as the covariates and WM pretest perfor-
mance as the dependent variable. There was no main effect of group, $F(3, 97) = 0.65, p = .978, \eta^2 = .002$, confirming that our groups started from the same position.

**WM posttest**

Finally, to assess whether language and WM training led to gains in WM, we carried out a one-way independent ANCOVA with four levels on L1, L2, WM, and control, with age, gender, and L2 experience as covariates and WM posttest score as the dependent variable (Fig. 7). There was a significant main effect of group, $F(3, 97) = 23.07, p < .001, \eta^2 = .416$. Pairwise comparisons with Sidak correction revealed a significant difference in WM scores between WM training ($M = 86.43, SD = 7.45$) and control ($M = 71.88, SD = 6.61$), $p < .001$, 95% CI [8.59, 19.75], between WM training and L1 ($M = 72.45, SD = 5.46$), $p < .001$, 95% CI [8.07, 19.40], and between WM training and L2 ($M = 72.06, SD = 9.54$), $p < .001$, 95% CI [8.70, 19.83], but no significant differences were found among the remaining groups. Covariate analysis revealed no significant effect of age, gender, or L2 exposure on WM performance.

**Comparison with chance**

To demonstrate that training led to performance that was significantly different from chance, we conducted a one-sample $t$ test (test value = 0.5) comparing WM training group posttest performance on WM with chance, $t(25) = 24.93, M = .86, p < .001, d = .07$. This shows that the effect of training had led to performance that was significantly different from chance.

---

Fig. 7. Working memory (WM) performance by training group. L1, first language; L2, second language; Ctrl, control. **$p < .01$.**
Individual differences

Thus far, we have established group differences such that those children assigned to the WM training group improved in their syntactic ability in L1 and L2 significantly more than those in the control group. What would be even stronger evidence of a causal relationship between WM and syntax is if the variance in WM accuracy predicted the variance in language gains made from pretest to posttest. If this was a significant relationship, it would suggest that those children who started from a lower WM baseline or engaged more with the WM training benefitted the most in their language performance. To examine this, we calculated individual variation in two ways. First, we correlated participants’ WM scores at pretest baseline with their language gains throughout training (Fig. 8).

Using multiple linear regression, we found that WM baseline accuracy was significantly associated with language gains, $R^2 = .30$, $F(2, 23) = 5.04$, $p = .015$, with L1 being a significant component of this model ($t = 2.76$, $p = .011$, CI [0.12, 0.89]) and L2 being a nonsignificant component ($t = 0.91$, $p = .369$, CI [−0.34, 0.88]). This shows that for children in the WM training group, their WM baseline score was a significant predictor of how much their syntactic performance improved overall with respect to L1.

Second, we calculated participants’ mean WM accuracy during training and the gains made in language performance from pretest to posttest for each participant and examined the relationship (Fig. 9).

Using multiple linear regression, we found that average WM training accuracy was significantly associated with language gains, $R^2 = .24$, $F(2, 23) = 3.74$, $p = .039$, with L1 being a significant component of the model ($t = 2.15$, $p = .042$, CI [0.14, 0.71]) and L2 being a nonsignificant component ($t = 1.17$, $p = .252$, CI [−0.24, 0.88]). This shows that for children in the WM training group, their WM training score was a significant predictor of how much their syntactic performance improved overall with respect to L1.

Fig. 8. Correlation between language gains (posttest accuracy − pretest accuracy [%]) and working memory (WM) baseline accuracy [%] for children who participated in the WM training, broken down by first language (L1) and second language (L2). Regression lines are with 90% confidence intervals.
Discussion

This study aimed to understand more about the way in which WM and syntactic abilities interact in development and draw inferences about the domain specificity of language. It explored four possible hypotheses about the way in which WM and syntactic training may affect performance: as a result of training, (1) WM improvements transfer to improved syntactic ability, (2) improved syntactic ability transfers to improved WM performance, (3) improvements in WM and syntax transfer in both directions, and (4) improvements in WM and syntax transfer in neither direction. After 6 weeks of training in language or WM, children showed evidence of cognitive transfer from WM to syntax, but not vice versa, in support of Hypothesis 1. This conclusion is supported by (a) a significant posttest difference between the WM training group and the control group on language measures (Figs. 5 and 6 and associated statistics in Results), (b) a failure to find a significant posttest difference between language training groups and the control on WM measures (Fig. 7), and (c) the individual differences in WM baseline and training performance predicting the variance in language gains (Figs. 8 and 9). The implications of these results are discussed with respect to the themes set out in the Introduction.

Domain specificity

As we stated in the Introduction, two different views of what language is and how it is acquired both agree that language interacts with domain-general cognitive processes at some level; the debate is over the extent of this interaction, the causal direction, the developmental periods in which it occurs, and the cognitive domains in which it happens. Our results contribute to this debate by showing a deep interaction between WM and syntax in children and that the causal direction is from WM to syntax. We were able to show greater “depth” of this interaction than previous literature by employing a purely nonverbal measure of WM. These results add further evidence to the view that the devel-
The developmental trajectory of language acquisition is contingent on the developmental trajectory of other cognitive capacities. The bigger question is whether this evidence is best thought of as developing cognitive abilities applying a performance filter through which the encapsulated (and largely nondynamic) language competence emerges or whether it is more straightforward to interpret this deep relationship as the result of general cognition being repurposed for linguistic ends. What counts as “straightforward” is of course subjective, but there is at least one thing on which everyone agrees. Capacities such as memory, categorization, and attention have a life independent of language; they emerge before language does so, they exist in species that have no language and probably existed in our species before language evolved. The fact that these capacities are needed, regardless of one’s view on language, is an independent motivation that we should fully exploit their explanatory power before constructing theories that require ad hoc assumptions unique to language. If future research continues in the trend toward greater integration of language and cognition, then the language-unique parts of language are in retreat to such an extent that we argue it becomes more parsimonious to use the framework of general cognition to understand linguistic processes rather than call on processes unique to language (Ibbotson, 2020).

The implications of this influence both developmental theory and applied practice. Theoretical accounts that emphasize deep integration between language and cognition see language acquisition as a process of repurposing general cognition for linguistic ends (Bybee, 2010; Dąbrowska & Divjak, 2015, 2019; Goldberg, 2005; Ibbotson, 2020; Tomasello, 2003). Over the first year of life, before children are fully productive with language or can fully comprehend the communicative intentions of others, there is a range of categorization, memory, inhibition, analogy, attention, and social cognition skills emerging during development, and these resources are available to the child at the time or before they begin to use language more productively. Whether there are any linguistic processes or cognitive structures that are unique to language or whether it is “cognition all the way down” is still an open question and one that we hope future research will continue to address given that the answers will inform our understanding of the nature of language and cognition.

For applied practice, deep integration between domains holds the promise that interventions targeting one domain will spill over into the untrained domain and ultimately improve developmental outcomes for children. To the recent evidence that has demonstrated the efficacy of this approach for children with developmental language disorder (Delage et al., 2022; Stanford et al., 2019) and children with ASD (Delage et al., 2021), our research adds evidence from a typically developing population and a strong causal role of WM in syntax (see Delage et al., 2021, for the lack of control group and the additional evidence our individual differences analysis adds). As others have argued for in populations with developmental language disorder and ASD, WM limitations are likely to influence the acquisition of complex syntax and improving WM capacities via a dedicated training program could free up cognitive resources to deal more effectively with syntax (Delage et al., 2021, 2022; Stanford et al., 2019). In this regard, children represent a particularly promising group to target transfer because of their greater neuroplasticity (e.g., Klingberg, 2010).

Cognitive transfer

Our results add a positive datum to a rather mixed overall picture of WM transfer in the literature. We suggest that the most parsimonious mechanism by which transfer occurred here is that there is a close functional mapping between the kinds of cognitive processes required to succeed in the WM task and the subject–verb agreement task. These abilities are connected in a functional hierarchy such that training one part of the hierarchy transfers to branches close by (Carroll, 1993; Lövdén et al., 2010; Noack et al., 2009, 2014; see also Fig. 1). We cannot say for sure why we saw transfer in only one direction, but one possibility is that subject–verb agreement WM is more task-specific than visual n-back WM, so that, broadly speaking, transfer is more likely from the general to the specific rather than vice versa. This is a speculation that warrants further follow-up investigations because it implies an asymmetry in the functional hierarchy of cognition, where training can transfer along one branch but not another (Fig. 1).

In alignment with process-specific theories of transfer, we saw the greatest transfer where there is functional overlap in the processes that trained and untrained tasks engage, the content they use, or
the mnemonic strategies participants employ to succeed (e.g., Soveri et al., 2017). For example, using a strategy for chunking items in WM to link them to multi-chunk items in long-term memory for later retrieval is common to both visual–spatial processing and language (Christiansen & Arnon, 2017; Cowan et al., 2012; Miller, 1956). Also consistent with our findings is a more skill-based theory of transfer (e.g., Gathercole et al., 2019; Taatgen, 2013). Such theories argue that task overlap alone is not a sufficient condition to predict transfer, and a necessary further restriction is that training tasks require the establishment or refinement of a cognitive routine that is not already fully developed. This approach predicts transfer even if tasks differ on “low-level” elements of processing or content that might be quite specific to the task as long as they share task-general skills. Whatever the precise mechanism by which transfer occurred, the result was that by improving children’s nonlinguistic WM, we were also improving their ability to succeed on a linguistic task that also required WM. As well as establishing cognitive transfer at the level of group differences, we also established that individual variation in WM performance predicted the extent to which children’s syntax improved. We did this to implicate an even stronger causal role for WM in syntactic processing, but these results are of relevance to a broader debate about whether individuals with higher levels of task-relevant cognitive resources gain more from training or whether they gain less (e.g., Lövdén et al., 2012). The current data provide suggestive evidence for a magnification account of training, namely that the largest gains in syntax were shown by the children with the highest WM performance (at both baseline and training), rather than a compensation account, which would predict that high performers are already at ceiling levels of performance and so stand to gain less from any interventions. Note that we pitched the difficulty of the n-back test at baseline so it was sufficiently challenging for the majority of children and that there was enough “headroom” for them to get better as a result of training. Thus, it is difficult to say whether the most cognitively efficient children at pretest were already performing at ceiling without varying the n-back difficulty on an individual-level basis. Further research that was designed to specifically look at the magnification versus compensation issue should consider this possibility along with the fact that any negative correlations between the highest performers and least training gains could be a result of regression to the mean (Smolen´ et al., 2018).

Whether the children were high WM performers relative to other samples of the population is also worth considering. We know that socioeconomic status predicts health, well-being, and cognitive ability, including executive function (e.g., Last et al., 2018). The children in this study were selectively drawn from private schools, so only those parents who could afford the fees were able to send their children there. Although the history in Cuba has generally been one of state intervention informed by a socialist ideology to reduce inequalities in provision of, for example, free health care and higher education, inequalities do persist, and therefore it is plausible that higher executive function might be expected in our sample. Given that our data tentatively support a magnification account of training, that would mean that the children in our sample were especially well placed to gain from any intervention in comparison with children from less wealthy families. Some caution is warranted here, however, because the very notion of socioeconomic status, with the societal stratification that implies, does not transfer very meaningfully onto the everyday experience of most Cubans. Furthermore, it is unclear how socioeconomic status itself could be a mechanism driving the main transfer effect (rather than a moderator of it) or could explain why we saw transfer from WM to syntax but not vice versa.

**WM and language**

We focused on WM because of its importance in underpinning a wide range of complex cognitive tasks and developmental outcomes for children, including language acquisition. Although the n-back is ostensibly a test of WM, we cannot rule out that it also engaged (and thus trained) other executive functions as well. Indeed, if these abilities are organized according to a functionally related hierarchy, then that is what one would predict (Carroll, 1993). For example, in the n-back, irrelevant items must be inhibited and abandoned from WM, and a counting and matching process between the upcoming and stored stimuli in WM is necessary to make a decision of whether the stimuli are the same to initiate a correct response (Rac-Lubashevsky & Kessler, 2016). From other work, we know that training adults on a task that required high levels of cognitive control produced significant benefits on several
untrained transfer tasks, including memory and language (Hussey et al., 2017). In addition, children’s ability to interpret temporarily ambiguous sentences relates to individual differences in their executive functioning abilities (Woodward et al., 2016).

This raises the question of whether one should conceive of the transfer established here as one of WM or broader inhibitory control. The answer to that question very much depends on the theoretical position one takes of what WM is and how domain specific it is. For example, the theoretical framework of Cowan (1988) places greater emphasis on the possibility of domain interference within WM than does that of Baddeley (1986), and more generally there has been a continuing controversy about the extent to which modes (verbal and nonverbal) and components of the executive function (attention and WM) interfere with one another (e.g., Cowan & Morey, 2007; Fougnie & Marois, 2011; Morey & Bieler, 2013).

Finally, it is well known that center-embedded relative clauses of the type “The reporter that the senator attacked admitted the error” make a large demand on WM, causing even adults to misinterpret who did what to whom 15% of the time (Larkin & Burns, 1977). The difficulty arises not only because of the demand the sentence places on WM but also because “The reporter” is both the agent of an action in one clause (“admitted”) and the recipient of an action in another (“attacked”); this is a superposition of roles that seem to pose particular difficulties for linguistic processing (Bever, 1970). A similar conflict is in operation in our study, where it is not just length of the clause that may call on WM, but there is also an intervening noun that might cause interference with the head noun and require inhibitory control. Thus, we predict that any facilitatory boost of WM or more general executive function training would be less in sentences that did not contain this noun–noun superposition (e.g., “The book, tattered and unread, was white”), something that further work could clarify.

Second language learners

The transfer was most striking in the case of L1, where WM training was as effective as language training in boosting syntactic performance. L2 performance also received a boost from WM training (it was significantly different from the control, as was L1), but unlike L1 it was not as effective as L2 training itself. This is evidenced by a significant difference between L2 and WM training groups on measures of L2 (Fig. 6) and the failure to find a difference between L1 and WM training groups on measures of L1 (Fig. 5). The individual differences analysis showed that the strength of WM training was a significant predictor for L1 gains but not for L2 (although the regression line was also positive). From previous research, there were strong reasons to suspect that engaging with different L1/L2 stimuli might lead to different processing and parsing strategies and that this could plausibly result in different transfer effects. We briefly reviewed several models of L1/L2 parsing in the Introduction, and although not a direct test of any of them (our main hypotheses concerned cognitive transfer), our results are perhaps most interpretable under the framework offered by Cunnings (2017). Cunnings argued that L1/L2 differences in parsing are likely to be caused by L2 speakers’ greater susceptibility to similarity-based retrieval interference in part because of their greater reliance on semantic and pragmatic information when parsing. We took steps to control the semantic plausibility of subject–verb agreement such that local relationships (embedded noun–verb) were semantically plausible (e.g., a piano is the kind of thing capable of being white, as is a book) rather than implausible (e.g., a book can be written in German but the stairs cannot), which would have caused a confounding cue to the grammaticality of agreement between the subject and main verb. It is possible that this control could have disproportionately affected the L2 speakers, who rely more heavily on semantic and pragmatic information and thus had a harder time in general coordinating subject–verb agreement. The fact that L2 pretest started from 56% correct versus 65% correct in L1 is indicative evidence for this, and it may be that starting from a lower base requires more and longer training to achieve the same effects that we witnessed in our L1 group. Without corroborating online measures, it is difficult to be sure exactly how participants were parsing L1/L2 in a way that could have led to different transfer effects, something that requires further investigation. And of course it could have been differences in the morphosyntax of the languages used that drove the differences, an idea we explore further below in the “Limitations and future directions” section below.
One issue that could have affected both L1 and L2 is our use of ungrammatical sentences as part of the grammaticality judgment task. Post hoc investigations revealed a significant role of grammaticality in the effect of training. Language training improved the accuracy of identifying grammatical sentences ($M = 68\%$, $SD = 21$) more than it did ungrammatical ones ($M = 46\%$, $SD = 23$), $t(102) = 4.91$, $p < .001$. In addition, WM training improved the accuracy of identifying grammatical sentences ($M = 75\%$, $SD = 17$) more than it did ungrammatical ones ($M = 37\%$, $SD = 37$), $t(102) = 9.85$, $p < .001$. In sentences of the type used in our experiment, verbs are thought to trigger retrieval cues such as the grammatical number of the candidate noun phrase it is supposed to agree with as well as its case or syntactic position (Badecker & Kuminiak, 2007). The efficiency of this process is determined by whether there are encodings that match cues, how strongly they match the cues, and how uniquely they do so (Nairne, 2002). The greater cognitive cost of processing ungrammatical sentences caused by the mismatch means it is likely that participants had a harder time in general processing ungrammatical sentences.

Finally, we note that age was a significant covariate in L1 performance such that older children tended to be more accurate, as would be expected. In addition, despite being an imprecise measurement of L2 exposure (parental estimate), this was a significant predictor in the L2 performance such that children with more L2 exposure were likely to perform more accurately, again as would be expected. Our main levels of interest (WM, L1, L2, and control) remained significant after accounting for these covariates.

**Limitations and future directions**

Despite our best efforts to balance L1/L2 in terms of frequency, local agreement, semantic plausibility, dependency length, and phonetic distinctiveness, it was simply not possible to match L1 and L2 on all relevant aspects or indeed know what all the relevant measures were in advance. For example, the average intervening phrase was higher in L1 than in L2, so we could balance the error rate between different levels of proficiencies. One side effect of this was that the intervening phrase for Cuban Spanish contained more potential sources of conflict between the grammatical numbers of the embedded noun and the head noun, something that could make coordinating the longer distance agreement between subject and verb more or less difficult. L1 verb agreement was on the main verbs in the past tense, whereas in L2 it was carried by be in the present tense, so the singular and plural forms were distinct and these differences could have also affected the way in which WM interacted with L1 and L2. The morphological component of each language also differs, with Spanish being richer than English, and this has been shown to influence how comprehenders process subject–verb agreement dependencies (Acuña-Fariña, 2018). We would like to see future work address some of the limitations that between-language differences impose by employing groups who are more balanced in their bilingual dominance and by doing so create more closely matched linguistic training material than we were able to do here. Note however, differences between L1 and L2 do not invalidate the main effect of cognitive transfer, nor do they explain why transfer occurred in one direction rather than the other. Rather, the between-language condition differences point to the fact that this effect was demonstrated more clearly in L1 than in L2.

Like all cognitive tasks, the ones we employed here were an imperfect measure of the underlying cognitive ability they are designed to assess and, inversely, cognitive abilities never explain the total variance of indicator tasks (Noack et al., 2014). It would be of great value therefore to replicate this methodology with other related faculties and sub-faculties (inhibition, switch, control, and match) to differentiate which components of executive control and of wider cognition in general are carrying the transfer effect.

**Conclusions**

Cognitive transfer has proved to be a useful methodology both for inferring the interconnectivity of different components of the mind and for the promise it holds for intervention training. We found that, compared with the control, children in the training conditions showed cognitive transfer from WM to syntax but not from syntax to WM. The result was most striking in the case of their L1, where
WM training was as effective as language training in boosting syntactic performance. Following others, we argue that transfer occurred because there is functional overlap in the processes that trained and untrained tasks engage—that is, they call on similar cognitive resources to get the job done. As well as establishing cognitive transfer at the group level, we also found that individual differences in WM performance, both at baseline and in training, predicted the extent to which children’s syntax improved. The directionality of transfer and the group-level and individual-level results, established in the context of a randomized control design, all point to a strong causal role for domain-general cognition in the processes of language acquisition.

Acknowledgments

Thanks go to the parents and children who consented to take part in this study, the teachers and staff at both language schools who supported the research, and we acknowledge the helpful suggestions and feedback of the associate editor and three anonymous reviewers.

Data availability

The data are openly available at Open Research Data Online here https://doi.org/10.21954/ou.rd.21629879.v1. Based on the Figshare platform, data are kept for a minimum of 10 years after project completion.

Appendix A

Interested parents and children were briefed on-site in Cuba via oral presentation, with opportunities to ask questions of the lead researcher. Together with a written participant information sheet (in Cuban Spanish), this informed potential participants of the aims of the project, who was conducting the experiment and their affiliations, why participants were being invited to take part, what the study involved if they decided to take part, what we would be observing, and the main research methods. They were also informed how their data would be used and stored and that no individual would be identifiable from the published results. All communications with parents and children stressed their right to withdraw from the study at any point as well as the lead experimenter monitoring for signs of nonverbal assent or otherwise during the experiment. It was made clear that if one condition of the study performed significantly and robustly better than another because of an intervention, then that same intervention would be offered to all participants.

Appendix B

Frequency

Because lexical frequency plays such an important role in language processing, comprehension and acquisition (Ambridge et al., 2015; Diessel, 2007; Tomasello, 2003), every effort was made to balance the frequency distribution of test and training items between first language (L1) and second language (L2). Using the Corpus del Español: Web/Dialects (Davies, 2016) and the Corpus of Contemporary American English (COCA) (Davies, 2008) as a baseline reference, items were selected for inclusion in the training and test phases such that the overall frequency of words in a larger corpus was comparable (Table B1).

Although we recognize the likelihood that the raw frequency of items will be different between corpora—with the dominant Cuban Spanish being experienced much more frequently than the sub-dominant English—we wanted to make sure that the frequencies of items within corpora were broadly equivalent such that items were comparable in their frequency distributions but not in their absolute levels.
Local agreement and semantic plausibility

Local agreement and semantic plausibility were also counterbalanced within the presentation of L1 and L2 sentences. Because the number of embedded noun can either aid or inhibit long-distance coordination of the head noun with the verb (Franck et al., 2002), we controlled for this by allowing the number of the noun proximal to verb to agree with the head noun 50% of the time (e.g., 2a) and to disagree 50% of the time (e.g., 2b). In this way, we could be sure that participants could not rely on a local coordination strategy to produce correct responses.

(2a) [The book [under the piano] is white]
(2b) *[The book [under the stairs] are white]

We also controlled the semantic plausibility of this relationship such that local relationships (embedded noun–verb) were semantically plausible (e.g., in 3a, a piano is the kind of thing capable of being white, as is a book) rather than implausible (e.g., in 3b, a book can be written in German but the stairs cannot), which would have caused an unwanted cue to the grammaticality of agreement between the subject and main verb.

(3a) *[The book [under the piano] are white]
(3b) [The book [under the stairs] is in German]

Dependency length

For L2 sentences, simple present third-person forms of the verb were used (is, are), taking into consideration students’ proficiency in English; sentences also had shorter antecedent gaps than in L1 like those in (4).

(4) The bus in the school is yellow

L1, L2, and working memory (WM) sessions (whether in pretest, posttest, or training) were matched in terms of time spent on intervention per child, resulting in a total of 20 sentences for L1 and 32 sentences for L2. In this way, we could ensure that both L1 and L2 were broadly comparable in terms of error rate before the training began; recall the pretest accuracy of 43.06% for L1 and 52.78% for L2.

Table B1
Summary statistics comparing the frequency of L1 and L2 items in a larger corpus

<table>
<thead>
<tr>
<th></th>
<th>Cuban Spanish (L1)</th>
<th>English (L2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corpus</td>
<td>Corpus del Español: Web/Dialects (Cuba)</td>
<td>Corpus of Contemporary American English</td>
</tr>
<tr>
<td>Corpus size</td>
<td>67,655,690</td>
<td>560,000,000</td>
</tr>
<tr>
<td>Word tokens</td>
<td>224</td>
<td>96</td>
</tr>
<tr>
<td>Word types</td>
<td>103</td>
<td>34</td>
</tr>
<tr>
<td>Type/token ratio</td>
<td>0.459</td>
<td>0.354</td>
</tr>
<tr>
<td>Average raw frequency</td>
<td>186 633</td>
<td>1 554 068</td>
</tr>
<tr>
<td>word type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ratio raw frequency/corpus</td>
<td>0.002758</td>
<td>0.002775</td>
</tr>
</tbody>
</table>

Note. L1, first language; L2, second language.
Phonetic distinctiveness

In English sentences, the forms of be (is, are) needed for successful agreement are relatively phonetically distinct from one another. Because more verbs were needed in Spanish to satisfy their greater dependency length, greater care was needed to offer L1 speakers the same key phonemic contrast and thus the same opportunity to succeed at the task. Because we fixed person and the agreement always relied on number, we were able to maximize the phonetic contrast by only using the past forms of the verb in third-person singular and plural [e.g., vio vs. vieron in (5)] and thus avoided using the less distinct option where the third-person singular and plural of the verb differ in only one sound [final /n/ in (6)].

(5) El maestro vió ... vs. *El maestro vieron ... (The teacher saw ...)
(6) El maestro ve ... vs. *El maestro ven ... (The teacher sees ...)

Appendix C. Supplementary material

Supplementary material to this article can be found online at https://doi.org/10.1016/j.jecp.2022.105593.

References


Further Reading
