The Effect of Explicit Instruction on Implicit and Explicit Linguistic Knowledge in Kindergartners

Spit, S.; Andringa, S.; Rispens, J.; Aboh, E.O.

Routledge
Taylor & Francis Group

The Effect of Explicit Instruction on Implicit and Explicit Linguistic Knowledge in Kindergartners

Sybren Spit [ID], Sible Andringa [ID], Judith Rispens, and Enoch O. Aboh

Faculty of Humanities, Amsterdam Center for Language and Communication, University of Amsterdam, The Netherlands

ABSTRACT

Research consistently shows that adults engaged in tutored acquisition benefit from explicit instruction in several linguistic domains. For preschool children, it is often assumed that such explicit instruction does not make a difference. In the present study, we investigated whether explicit instruction affected young learners in acquiring a morpho-syntactic element. A total of 103 Dutch-speaking kindergartners ($M = 5;7$) received training in a miniature language to learn a meaningful agreement marker. Results from a picture matching task, during which eye movements were recorded, provided no evidence that explicit instruction led to higher accuracy rates, but suggest that it did lead to earlier predictive eye movements. These data seem incompatible with the idea that explicit instruction does not make a difference when kindergartners learn a grammatical element, and tentatively suggest that explicit instruction has a different effect on explicit knowledge than on implicit knowledge in this age group.

Introduction

It is often noted there is a fundamental difference between young children who learn the grammatical regularities of a new language and adults who go through the same process (Bley-Vroman, 1990; Hartshorne et al., 2018; Krashen et al., 1982; Singleton & Ryan, 2004). A central question in linguistics is how we can characterize this difference between adult and child language learners. Scholars have proposed that differences between child and adult language learning might be due to different linguistic inputs, contextual factors, or general cognitive aptitude, or might arise from a critical period for language learning specifically (Andringa & Dabrowska, 2019; Birdsong, 2005, 2006; Birdsong & Vanhove, 2016; Brooks & Kempe, 2019; Doughty, 2019; Granena & Yilmaz, 2019; Hakuta et al., 2003; J.S. Johnson & Newport, 1989; Newport, 2020; Singleton & Munoz, 2011). One take on this issue is that there is a difference in the available learning mechanisms for adult and child language learners (Lichtman, 2016; Pakuklak & Neville, 2011; Zwart et al., 2017). In this scenario, children use more implicit or procedural learning mechanisms when they acquire grammatical structures, whereas adults rely on explicit or declarative learning mechanisms, and possibly in conjunction with implicit learning mechanisms (DeKeyser, 2000, 2003; R. Ellis, 2005, 2009; Paradis, 2004, 2009; Ullman, 2001).

Lichtman (2016) summarizes the idea that children and adults make use of different learning mechanisms in the form of the 'maturational hypothesis': both children and adults are maturationally constrained to using particular learning mechanisms. She also observes that some scholars take a fairly radical position with regard to this hypothesis by claiming children learn grammatical rules using implicit learning mechanisms only (Bialystok, 1994; DeKeyser & Larson-Hall, 2005; Paradis, 2004, 2009). This is perhaps best illustrated by Bialystok (1994) who posits "Rules make sense to adults; they

make little difference to young children" (p. 565), suggesting children do not rely on explicit learning when acquiring grammatical structures. However, hardly any studies have investigated specifically whether young children actually can learn grammar explicitly. It could well be the case that children utilize implicit learning only, but to our knowledge such claims have hardly been tested experimentally. In this study we try to fill a this void partly by investigating if young children could benefit from awareness of form through explicit instruction of a grammatical regularity. We conducted an artificial language learning experiment in which we investigated whether kindergartners are influenced by explicit instruction when they have to learn a meaningful grammatical element.

Explicit and implicit instruction, learning, and knowledge

The difference between implicit and explicit language learning may be captured as follows. Explicit learning occurs when learners are aware that they are learning and take effort to discover the regularities in the target language, whereas implicit learning happens when learners do not know they are learning and do not have such intentions (e.g., Dörnyei, 2009; N. Ellis, 2015; Hulstijn, 2015; Ullman, 2001, 2016). Often, these types of learning are defined by the type of knowledge they result in (e.g., Hulstijn, 2015; Reber, 1967), with explicit learning seen as learning that results in explicit knowledge and implicit learning as learning that results in implicit knowledge. Explicit knowledge is available to awareness, whereas implicit knowledge is not (Rebuschat, 2013). Knowledge is regarded explicit when learners can verbally reflect on it or when they are confident of their behavior on an explicit knowledge-based task. In contrast, learners might not be confident about their performance during a linguistic task, or they may be unable to verbally reflect on the knowledge they possess. In these cases, they are supposedly unaware of their linguistic knowledge, indicating it is implicit. Although explicit and implicit learning often result in explicit and implicit knowledge, respectively (e.g., Andringa & Rebuschat, 2015; N. Ellis, 2005, 2011; Z. Han & Finneran, 2014; Hulstijn, 2005, 2015), this need not be the case. Explicit learning might result in implicit knowledge, and implicit learning could lead to explicit knowledge (Batstone, 2002; DeKeyser, 2009; Williams, 2009). This occurs for example, when learners are aware that they are learning a grammatical structure in a second language, but nevertheless end up with knowledge they are not aware of (R. Ellis, 2009).

Furthermore, we can make a distinction between explicit and implicit instruction (DeKeyser, 1995; R. Ellis, 2005, 2009; Housen & Pierrard, 2006). When receiving explicit instruction, learners receive instruction on what the target structure is, while implicit instruction refers to a situation in which learners do not receive such instruction (R. Ellis, 2009). Hypothetically, explicit instruction could lead to implicit learning and result in implicit knowledge, for example, when learners are instructed on what they have to learn in the second language they are learning, but cannot understand this instruction and end up with knowledge they are not aware of. Yet, from empirical studies we know that explicit instruction often leads to a higher degree of explicit learning, and results in more explicit knowledge (e.g., Hamrick & Rebuschat, 2012), and the same holds for implicit instruction, learning and knowledge (e.g., Williams, 2005). Still, it is important to keep in mind that neither explicit learning, knowledge and instruction nor implicit learning, knowledge and instruction necessarily need to co-occur (Batstone, 2002; DeKeyser, 2009; Williams, 2009).

Studies on explicit and implicit learning

Many studies support the idea that explicit learning mechanisms play a role in adult language learning and that implicit learning mechanisms contribute to child language learning. Adult learners are often greatly helped by explicit instruction, as evidenced by studies that show that explicit instruction has a positive effect on learning outcomes (for meta-analyses see Goo et al., 2015; Norris & Ortega, 2000; Spada & Tomita, 2010). These meta-analyses show that, whether they are explicitly instructed on pragmatics (e.g., Alcón Soler, 2005; Fukuya & Martínez-Flor, 2009; Koike & Pearson, 2005; Martínez-Flor & Fukuya, 2005; Muranoi, 2000; Sheen, 2007), syntax (e.g., Gass et al., 2003; Robinson, 1997a, 1997b; Scott, 1989),

morpho-syntax (e.g., DeKeyser, 1995; Ellis et al., 2006; Gass et al., 2003; J. Williams & Evans, 1998) or phonetics (e.g., Kissling, 2013; Saito, 2013), adults who receive explicit instruction often show higher learning gains than uninstructed language learners in classroom or experimental settings.

Evidence from the statistical learning literature suggests that children make use of implicit learning mechanisms to acquire fundamental parts of language (e.g., Erickson & Thiessen, 2015; Frost & Monaghan, 2016; Romberg & Saffran, 2010). Statistical learning is defined as the ability to implicitly extract linguistic regularities from distributional properties in the input (e.g., Arciuli & Simpson, 2012; Aslin et al., 1998; Baker et al., 2004; Conway et al., 2010; Conway & Christiansen, 2006; Kidd, 2011; Perruchet & Pacton, 2006; Reber, 1967; Rebuschat, 2013; Rebuschat & Williams, 2012; Turk-Browne et al., 2005). This learning mechanism enables both children and adults to infer word boundaries (e.g., for children Aslin et al., 1998; Gómez, 2002; Gómez & Gerken, 1999; Saffran et al., 1996; for adults Endress & Bonatti, 2007; Perruchet & Vinter, 1998; Thiessen et al., 2013), acquire words and their meanings (e.g., for children Smith et al., 2014; Smith & Yu, 2008; Vlach & Johnson, 2013; for adults Kachergis et al., 2014; Vouloumanos, 2008; Yu & Smith, 2007), and form grammatical categories (e.g., for children Lany, 2014; Lany & Saffran, 2013; for adults Chen et al., 2017; Monaghan et al., 2015). Taken together, and under the assumption that statistical learning really is implicit, these statistical learning studies and the earlier described instruction studies seem to show that children learn implicitly, and that adults, in addition to learning implicitly, can learn explicitly as well. Intriguingly, we hardly know anything about the role of explicit learning processes when children learn grammatical rules, because this has hardly ever been researched.

There is some work on the role of explicit instruction in young children. Several studies have investigated the role of explicit instruction in the acquisition of vocabulary items in this age group (e.g., Coyne et al., 2009; Silverman, 2007; Vaahtoranta et al., 2018). A meta-analysis of such studies shows that vocabulary training in kindergartners is effective and that more explicit instruction conditions result in bigger effect sizes (Marulis & Neuman, 2010). Furthermore, the way in which explicit instruction in phonemic awareness relates to reading and spelling abilities has been widely studied in similar age groups as well (e.g., Degé & Schwarzer, 2011; Furnes & Samuelsson, 2011; Gillon, 2018; C. Johnson & Goswami, 2010; Saygin et al., 2013). A meta-analysis by Ehri et al. (2001) on this topic finds that instruction aimed at phonemic awareness significantly aids both reading and spelling in preschoolers, kindergartners, and first graders. The work on explicit instruction in children – although limited to the domains of vocabulary, and reading and spelling – indicates that young learners can benefit from receiving explicit instruction when acquiring particular linguistic knowledge. However, whether explicit instruction also influences learning grammatical structures in these age groups, is yet an unresolved question.

Lichtman (2016) addressed this question by investigating the effect of explicit instruction in a group of children from 5 to 7 years old, and compared this to adult learners. Participants learned the new language 'Sillyspeak' over six consecutive days, in which they had to acquire a particular word order and a determiner system. During these days, half of the participants received only exposure to this language, and the other half received an explicit training session instead of a part of the exposure. Participants who received explicit instruction were taught grammatical gender on the determiners using pictures, and word order by contrasting it to the English system. On the seventh day, they were tested on their knowledge of the language. The study did not show an effect of explicit instruction in children on accuracy when producing these grammatical structures, but also did not show such an effect in the adult group. Crucially, because explicitly instructed participants received less exposure than uninstructed peers, it is difficult to determine whether explicit instruction indeed does not make a difference for children. Perhaps explicit instruction helped these young learners to compensate for the extra exposure that uninstructed learners received. Alternatively, explicit instruction may not have added value to the input children received over the course of 6 days, even though this input was not equal for both groups. The imbalance of exposure seems to make it difficult to identify the actual effect of explicit instruction, which is not an uncommon problem in instruction studies (cf. Andringa et al., 2011). These results thus warrant further investigation. Do young children benefit from explicit instruction, or does it make no difference

whatsoever? To answer to this question, we will delve deeper into this issue and challenge claims as put forward by Bialystok (1994), which others followed up upon (DeKeyser & Larson-Hall, 2005; Paradis, 2004, 2009), that children learn grammatical regularities only implicitly, and do not make use of explicit learning mechanisms. Using a miniature language learning experiment, we investigated whether young learners who receive exposure and explicit instruction are better able to learn a morpho-syntactic element than young learners who receive the same amount of exposure but without explicit instruction. We were thus not interested in whether or not children use explicit instruction differently than adults, but merely whether they use it at all.

Current Study

To test whether children can make use of explicit learning mechanisms through explicit instruction, we investigated whether explicit instruction influences the acquisition of a morpho-syntactic agreement marker in kindergartners. We used the artificial language learning design from Spit et al. (2020); (2021), in which we tested whether kindergartners could learn a grammatical marker expressing a plurality feature from mere exposure. In the current experiment we compared kindergartners who had to learn this marker from exposure only with kindergartners who received not only exposure, but also explicit instruction about this marker. The previous studies probed into children's knowledge using a picture matching task only, which is considered a more explicit measure in second language acquisition. This study also followed their eye movements during the task to gauge knowledge using a more implicit measure as well. Apart from the fact that their eyes were followed, the present experiment was nearly identical to our previous experiments for children in the exposure only group. As such, the current study provides an excellent opportunity to partially replicate our previous findings. Replication is rare in language studies, yet highly necessary (Marsden et al., 2018), because a single study can never be conclusive and accidental findings might be published (Cumming, 2014).

When investigating the possible effects of explicit instruction on learning a grammatical element, it is important to include measures of both explicit and implicit knowledge (Andringa et al., 2011; Andringa & Rebuschat, 2015; R. Ellis, 2005; Y. Han & Ellis, 1998). Explicit instruction often leads to a higher degree of explicit learning and results in explicit knowledge (e.g., Hamrick & Rebuschat, 2012; Williams, 2005), but explicit instruction could also lead to implicit knowledge (Batstone, 2002; DeKeyser, 2009; Williams, 2009). To measure learning, researchers have used grammaticality judgment tasks (Bader & Häussler, 2010; Schütze & Sprouse, 2013), picture matching tasks (Pan, 2012; Schmidt & Miller, 2010), eye tracking paradigms (Sagarra & Hanson, 2011; Sedivy, 2010), track reaction times (Kaiser, 2013; Marinis, 2010), and other experimental methodologies (Blom & Unsworth, 2010; Hoff, 2012; Podesva & Sharma, 2013). None of these methods necessarily reveal anything about the implicit or explicit status of the knowledge, but each is more strongly associated with one type of knowledge or the other (Andringa & Rebuschat, 2015; Bowles, 2011; R. Ellis, 2005; Godfroid, Loewen, Jung, Park, Gass &, 2015; Y. Han & Ellis, 1998).

On the one hand, measures like timed judgment tasks, eye tracking paradigms, and reaction time studies have been claimed to measure implicit knowledge, whereas on the other hand, untimed judgment tasks, meta-linguistic knowledge tasks, and picture matching tasks are often more associated with explicit knowledge (Andringa & Rebuschat, 2015; Godfroid et al., 2015). The former types of measures are performed under time pressure or do not explicitly ask participants something about the grammaticality of test items. This makes it unlikely that participants consciously steer their behavior during the task and that explicit knowledge drives task performance (R. Ellis, 2005; Loewen, 2009). The latter types of measures often focus on the form and grammaticality of linguistic stimuli very explicitly. This makes it more likely that explicit knowledge drives task behavior of the participant, although implicit knowledge could still be involved during these tasks as well (R. Ellis, 2005; Sagarra & Hanson, 2011). Importantly, to get a better picture of the exact influence explicit instruction might have on learning, we should include measures that are associated with both implicit and explicit knowledge, as explicit instruction potentially influences both types of knowledge (Godfroid, 2015).

Method

Participants

113 Dutch speaking children (59 males, 54 females, M = 5;6 years, SD = 0;9) took part in this experiment. All children were in kindergarten and were recruited from four primary schools in the western and eastern regions of the Netherlands. They did not have any diagnosed language or communication disorders. No restrictions were imposed upon taking part in the experiment, and no further background information was collected. Ten participants had to be excluded, because of problems with calibrating (N = 2), other technical issues (N = 7) or because they did not want to participate anymore (N = 1). As a result, we present data from 103 Dutch speaking children (52 males, 51 females, M = 5;7 years, SD = 0;8). From this group, 50 children received explicit instruction during the training (24 males, 26 females, M = 5;7 years, SD = 0;8), the other 53 children did not receive such instruction and were only exposed to the miniature language (28 males, 25 females, M = 5;6 years, SD = 0;9). In the remainder we refer to the former group as the explicitly instructed group and to the latter as learners who received only exposure.

Materials

Miniature language and target structure

We used the artificial language from Spit et al. (2020); (2021), which consisted of four proper names, 3 verbs, 2 grammatical markers, s6 frequent nouns, 12 infrequent nouns and one conjunction.[1] The language had subject-verb-object word order. In this language, a noun phrase on its own does not encode number and could correspond to either singular and plural referents. In a sentence, however, a noun phrase must be introduced by a nominal marker (*pli* or *tra*). This type of nominal marker is quite common in languages with residual noun classes, such as the Kwa language Akan (Appah, 2003) or the Austronesian language Cebuano (Parnes, 2011). In our miniature language, a noun introduced by *tra* can refer to both singular and plural referents: the correct interpretation must be inferred from the visual context. The sole function of *tra* therefore is to turn a bare noun into an argument in the sentence (in this language: an object). The language includes another nominal grammatical marker *pli*, which encodes number and indicates that the noun necessarily refers to multiple referents. A similar type of marker appears in the Kwa language Gungbe (Aboh, 2004; Aboh & DeGraff, 2014). The marker *tra* can be seen as a default nominal marker with no number specification, whereas *pli* not only turns the noun into an argument but adds information about number. In short, the rule participants had to learn was that whenever *pli* preceded a noun, this noun always referred to multiple referents, and that when *tra* precede a noun, this noun could refer to a singular referent or multiple referents. See Spit et al. (2020) for a more information about certain particularities of the miniature language.

Structure of the experiment

Before describing the precise characteristics of the input to learn this grammatical regularity and the practical procedure in the next sections, we will briefly lay out the overall structure of the experiment, of which Figure 1 is a visual illustration. To learn the regularity, participants were told a story with four protagonists (Carlo, Julia, Marco, and Maria), who were going on a holiday to a country of which they did not speak the language. Participants were asked to help the protagonists learn this new language. They were told they would see pictures and hear things in the new language that matched the pictures they saw. The experiment consisted of three parts. It started out with a small vocabulary training. After the vocabulary training a rule training session followed. During this session, children received input to acquire the grammatical rule. Children who were in the explicit instruction group, received instruction after one third of this rule training session (see below for details about the instruction). After the rule training, participants performed a test phase to test whether they acquired the rule.

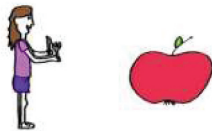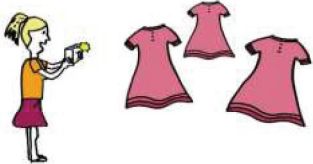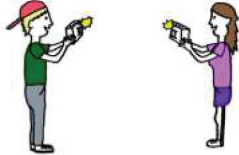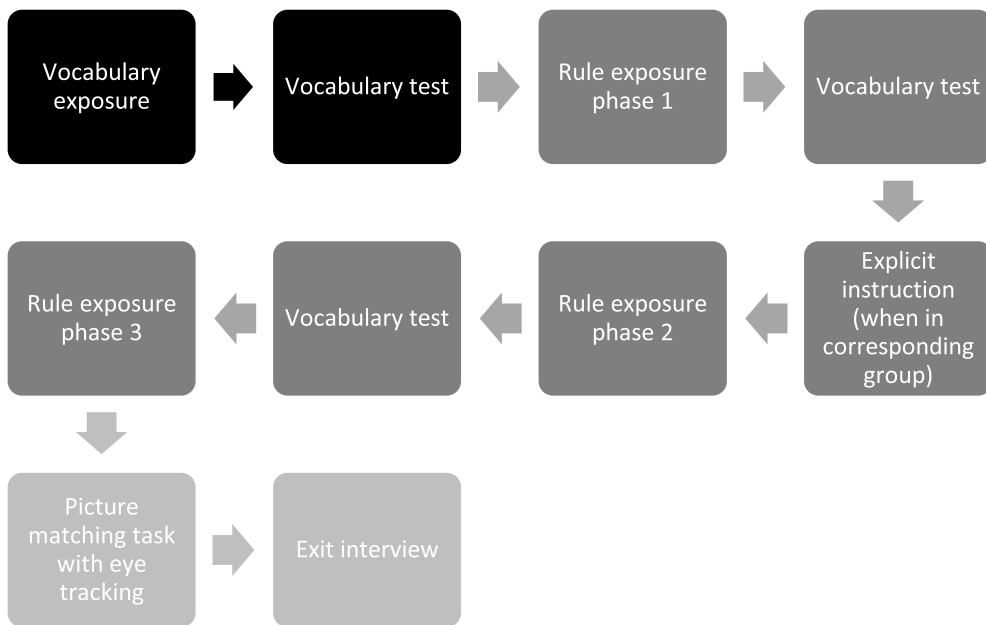[1] All words and their translations can be found in Appendix A.

Figure 1. A visual representation of the structure of the experiment. The top darkest boxes show the parts of the vocabulary training. The middle lighter boxes show the different parts of the rule training. The bottom lightest boxes shows the test phase.

As we were interested in the acquisition of the grammatical element, and not in that of lexical items, we bootstrapped participants slightly into learning the lexical items by exposing participants to the six frequent nouns of the language only. During this exposure, the nouns were presented auditorily, without the grammatical markers *pli* or *tra* accompanying them, and with a picture showing the meaning of the corresponding bare noun. Each noun was presented six times during this exposure, twice with one, twice with two and twice with three referents, to make sure a bare noun would not be associated with a particular number of referents. After this exposure, participants were given a picture matching task to test their vocabulary knowledge. They heard one of the frequent nouns and saw four pictures with referents of the other frequent nouns. This task was inserted to familiarize the participants with the picture matching task from early on during the experiment, but also to maintain participants' attention. Participants would progress to the rule training regardless of their performance on the vocabulary test.

After the vocabulary training, participants were exposed to three rule training phases in which they received input on the basis of which they could learn the grammatical rule. During each phase, 40 sentences were presented, adding up to a total of 120 sentences. Each sentence was accompanied by a picture showing the meaning of the sentence. A subset of 108 sentences consisted of a proper name, a verb, a marker and a noun. In these training sentences, the proper name functioned as subject, while the noun was the object. These sentences served as the input to learn the plural marking rule. The other 12 sentences were created using two proper names, the conjunction and a verb. These sentences did not contain a noun and thus did not provide evidence of the grammatical rule that had to be learned. These sentences were used to create the cover task (see below) and must be regarded as fillers. Examples of all sentence types and their accompanying pictures can be found in Table 1. Sentences in the language were always semantically plausible. After rule training phases one and two, participants were given a six-item vocabulary test (one item for each frequent noun), using the same procedure as described earlier. These vocabulary tests were inserted to maintain participants' attention. Participants received a sticker after each vocabulary test, regardless of their results.

All participants performed a cover task to maintain attention during learning. The children were made to believe that the four protagonists sometimes could not hear correctly what was said in the new

Table 1. Examples of artificial language training sentences and rough translations. For sentences with *tra* the noun could be translated both as a singular and a plural. The correct interpretation should be inferred from the visual context.

Structure	Example	Translation	Picture
Tra + **single referent**	*Maria rigarda tra zambo*	'Maria eats (an) apple(s)'	
Tra + **multiple referents**	*Carlo estima tra pano*	'Carlo looks at (a) cat(s)'	
Pli + **multiple referents**	*Julia pentura pli anso*	'Julia takes a picture of dresses'	
Filler	*Marco ut Maria pentura*	'Marco and Maria take a picture of each other'	

language. This was indicated by a questioning face of a protagonist after a certain stimulus. When participants saw this face, they had to repeat the previously heard stimulus. This cover task was inserted to keep participants focused; they did not have to repeat what was said correctly. Participants were introduced to this cover task during the vocabulary training during which they had to repeat each noun once. During the rule training phase, they had to repeat four filler sentences per training phase. Filler sentences did not contain a grammatical marker and a noun, and occurred at fixed moments during each phase.

Children who were in the explicit instruction group, received instruction immediately before the second rule learning phase, after they were exposed to 1/3 of the training sentences. During this instruction they were told the meaning of both grammatical markers *pli* and *tra* by one of the characters in the game. The Dutch instruction and its English translation can be seen in (1) and (2). The instruction was presented in smaller segments, which are indicated by dashes. Children would also see pictures accompanying parts of the instruction. When they would hear that a noun introduced by *pli* referred to two or three objects, they would see pictures with two or three objects respectively. Segments combined with an accompanying picture are in bold face.

(1)"Je ziet straks steeds weer een plaatje, en je hoort iets in de nieuwe taal. // Als je goed luistert, hoor je in de taal soms het woordje *tra*. Heb je dat al eens gehoord? Luister maar. // *Maria rigarda tra herbo* // Elke keer als je *tra* hoort zie je op het plaatje // **één** // **twee** // **of drie dingen** // **Kijk maar**. // Misschien heb je soms ook wel het woordje *pli* gehoord. Luister maar. //*Carlo rigarda pli zambo*. // Als je *pli* hoort staan er altijd // **twee** // **of drie dingen op het plaatje**. // **Zie je dat?** // Zullen we nu weer verder gaan?

(2)"You will see some pictures again, and will listen to the new language. // If you listen carefully, you will sometimes hear the word *tra* in the language. Did you hear that? Listen. // *Maria rigarda tra herbo*. // Every time you hear *tra* you will see // **one** // **two** // **or three things on the picture** // **See**? // Sometimes you might have heard

the word *pli*. Listen. // *Carlo rigarda pli zambo*. // Every time you hear *pli* there are always // **two** // **or three things on the picture**. // **See?** // Shall we continue?"
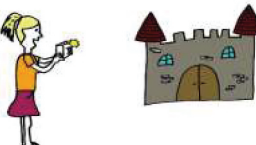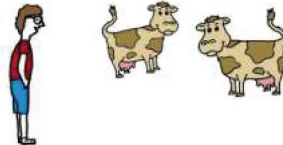
Characteristics of the input

From the 108 training sentences that participants could use to learn the rule during the rule training phases, 72 sentences contained the default marker *tra*, while 36 sentences contained the marker *pli*, which was indicative of number. In half of these 72 *tra* sentences, the noun had one referent. In the other half, the noun had multiple referents. In half of the sentences that contained a noun with multiple referents, two referents were shown, in the other half of the sentences that contained multiple referents, three referents were shown. This distribution was the same for sentences with *pli* that showed multiple referents and *tra* that showed multiple referents. As a result, 36 pictures showed one referent, 36 pictures showed two referents, and 36 pictures showed three referents.

Every frequent noun occurred twelve times in the input. It occurred four times with *pli* and eight times with *tra*. When a frequent noun occurred with *tra*, it referred to a single referent four times and multiple referents the other four times. Every infrequent noun occurred three times in the input, once with *pli* and twice with *tra*, of which it once referred to a single referent and once to multiple referents. Each noun referred to one, two or three referents equally often. Every noun occurred equally often over each of the three rule training phases, as did every grammatical marker.

Testing sensitivity to the rule

Immediately after the rule training phase, participants took part in a picture matching task to determine whether they became sensitive to the grammatical cue, and to measure the effect of explicit instruction on more explicit knowledge. In this task, participants heard 36 sentences based on the 12 infrequent nouns from the rule training phase, and had to choose which of two pictures matched each sentence. The infrequent nouns were used to test if participants could generalize the acquired knowledge to novel instances. Ideally, novel nouns are used for this. However, young children reportedly

Table 2. Examples of test items and their rough translations. For sentences with *tra* the noun could be translated both as a singular and a plural. The correct interpretation should be inferred from the visual context. In all three examples, the right picture was the target.

Marco rigarda pli bovo – 'Marco eats strawberries'

Pli

Julia pentura tra nutro – 'Julia takes a picture of (a) painting(s)

Tra

Carlo estima tra misto – 'Carlo looks at (a) dog(s)

Filler

have difficulties with tests in which they have to make decisions that are unrealistic to them (for a methodological review, see Pinto & Zuckerman, 2018), which would be the case when test items contained novel nouns and pictures. We tried to circumvent this by using infrequent rather than novel nouns in the picture matching task. This way, participants could base their response on what they had been exposed to; both pictures would be realistic options, as they would have seen them. Yet, by using infrequent nouns, we could still maximize the chance these children used rule knowledge when giving an answer, as chances are slim that they learned the meaning of these nouns from only three occurrences in the input.

24 out of 36 sentences used during the test phase were experimental items, 12 items were fillers. Examples of these different types of items can be seen in Table 2. For experimental items, participants had to choose between pictures with either one or multiple referents. The two pictures always referred to different referents. Twelve experimental items contained *pli*. For these items, the target picture always showed multiple referents. Another 12 experimental items contained *tra*. Any number of referents would be grammatical for these items. For half of the experimental items containing *tra*, the target picture showed multiple referents. For the other half of these items, the target picture showed a single referent. Every noun occurred once with *tra* and once with *pli* during the test. If participants learned the meaning of the infrequent noun, they should produce a target answer in both conditions. Possible learnability of particular nouns should thus have no effect on a difference between performance on *pli* and *tra* sentences. The performance on the sentences with *tra* can therefore be used as a baseline to which we can compare performance on the sentences with *pli*. We hypothesized that sensitivity to the statistical regularity in the input would lead to more target answers on trials with *pli* than on trials with *tra*, because for trials with *pli* participants could base their response on the number of referents the picture showed, whereas the number of referents was not indicative for trials with *tra*. Fillers included *pli* or *tra*, but showed the same number of referents on both pictures. These fillers were added to avoid that participants would link the grammatical markers to number during the test phase only, because they always had to choose between a picture with multiple and a picture with a single possible referent. Items of the different types were presented in a counterbalanced semi-randomized order.

To test whether explicit instruction also influenced more implicit knowledge, we measured participants' eye movements during the picture-matching task. Participants could see the two pictures from onset of the sentence, but we were interested at where they would look from the onset of the grammatical markers *pli* and *tra*, as this was the point at which we could expect eye movements to differ for the two conditions. Participants could look more often toward the target picture after the onset of *pli*, as this marker predicted the correct answer, whereas participants should not have a looking preference for any of the two pictures after the onset of *tra*, as this did not predict the correct answer. If participants acquired the regularity, this would lead toward more correct looks toward the target after the onset of the grammatical marker for sentences with *pli*, than for sentences with *tra*. If explicit instruction influences online processing in some way, a possible looking preference for the target picture for sentences with *pli* over sentences with *tra* might be observed at a different point in time in the explicitly instructed group than in the implicitly instructed group.

We presented the two pictures between which participants had to choose during the picture matching task in colored frames. Participants had to name the color of the frame of the correct picture. We used eight basic colors (black, blue, green, orange, pink, purple, red, yellow) to create these frames, which were semi-randomly distributed over all items. Color combinations were chosen such that the two colors contrasted with each other, to help participants distinguish between the colors they had to name. We used the color frames to minimize possible data loss during the task, as there are many threats to gathering eye tracking data in young participants (Léger et al., 2018). This way, participants did not have to point toward the correct picture, which might interfere with their eye movement measurements. They also did not have to press a button corresponding to the correct picture, which might have children looking at the buttons they need to press instead of the screen.

Because of possible difficulties with the distinction between left and right in this age group, we could also not let them call out the correct picture on the basis of the side of the screen.

After the picture matching task, an informal debriefing took place in which we examined whether participants were aware of their knowledge in a way they could verbalize. During this debriefing, participants were asked how they knew what the correct answer during the picture matching task was, and whether they knew the meaning of the words *pli* and *tra*.

Procedure

All stimuli were recorded by a female native speaker of Dutch. Sentences were recorded as a whole, leading to minor natural variation between similar words in the language. The test was administered in a quiet room at the participants' school. The task was presented on a laptop using E-prime (Psychology Software Tools, 2012). Eye tracking data was gathered using a Tobii Pro X3-120 Eye Tracker with an External Processing Unit. During vocabulary training, nouns and their accompanying pictures would be presented for 3 seconds, before automatically moving to the next noun. As the goals of this training were familiarization and keeping attention, we did not register scores during the vocabulary test systematically enough to report on. During rule training, sentences and their accompanying pictures would be presented for 4 seconds, before automatically moving to the next sentence. The instruction was prerecorded and segmented into smaller bits, such that the instructor could press a button on the keyboard to continue with the next piece of instruction. During the test phase, the experimenter pressed a button on the keyboard that corresponded to the answer the participant gave. Scores from this phase were registered automatically. We calibrated the eye tracker before the start of the test phase, and recorded eye movements at a rate of 120 Hz. After each test item, participants saw a fixation cross in the middle of the screen. The following test item would start automatically, if participants fixated for 100 ms (12 samples) on the cross. This ensured that participants were looking at the middle of the screen, between the pictures, at the onset of every test item. For every test item, the audio and pictures were presented simultaneously. Both the participant and the experimenter listened to the audio using headphones. The vocabulary training lasted 8 minutes on average, the rule training phase 15 minutes, and the picture matching task including calibration 10 minutes. The full experiment took approximately 30 minutes per participant. Ethical approval for this study was obtained from the University of Amsterdam and passive consent was obtained from children's parents or legal guardians before the start of the study.

Replication

Apart from the points listed below, the present and previous experiments were the same, and all statistical procedures to analyze the data from the picture matching task in the current experiment (see next section) were identical to the procedures we followed in our previous experiment.

(1) The vocabulary test after the vocabulary training at the start of the present experiment contained only two test items per noun. In Spit et al. (2020), all participants would continue to receive test items, until they identified all six nouns correctly four times. This proved frustrating, however, because most children had difficulties learning all nouns from the limited input they received. Therefore, it was problematic to keep the vocabulary training as it was, and it was altered from the test in Spit et al. (2021), as it is in this study. The vocabulary test was included to familiarize participants with the picture matching task, and participants did not receive any input to the grammatical regularity at this point. We think it is unlikely that this change would affect learning of the grammatical regularity.
(2) All participants in the present study performed the attention task, whereas only half of the participants performed this task in Spit et al. (2020). In our previous experiment, we observed a

tendency toward better performance by children in the attention task condition. Therefore, we decided to include the attention task for all children in Spit et al. (2021) and the current study.

(3) In Spit et al. (2021) the experiment ended with an opt out phase, in which we assessed whether children developed awareness of the grammatical regularity. It came after the picture matching task and could not have affected performance on this task. The opt out phase took place before the informal debriefing, and could have influenced the answers children gave during this debriefing in that study. This opt out phase was not part of Spit et al. (2020) and the current experiment.

(4) In the present study, we measured eye movements during the final picture matching task, which were not measured in the previous two studies. This means that the current picture matching task was slightly different from the previous two studies. In the current task, there was a small pause between items in which participants had to fixate on a cross in the middle of the screen. Also, in the previous studies, pictures were not presented in colored frames, which were necessary to enable children to give an answer without pointing to the screen. In the previous experiments, children could simply point toward the screen, which might have been cognitively less demanding.

Analysis

All analyses were carried out in R (R Core Team, 2015) using the lme4 package (Bates et al., 2015) and EyetrackingR package (Dink & Ferguson, 2015) where needed. To determine whether participants grasped the target regularity and whether explicit instruction affected this learning, two separate analyses were carried out. We used linear mixed effects modeling to analyze the results of the picture matching task and a cluster based permutation analysis to analyze the eye tracking data. We also carried out an additional cluster based permutation analysis on filler items, to test whether our groups are comparable in their referential processing. Scripts and data that were used for these analyses can be found on our Open Science Framework Page (Spit Andringa, Rispens, Aboh & Vet, 2019).

To investigate the effect of explicit instruction on learning as measured by the picture matching task, a generalized linear regression model with mixed effects was carried out to investigate the results from the picture matching task. This model took the responses from the task (1 for a correct answer, 0 for an incorrect answer) as a dependent variable, marker type as a within-participants fixed effect, type of instruction as a between-participants fixed effect, participant as a between-participants random effect and item as a within-participants random effect. Marker type was included as a random intercept for participant, because it was a within-participant fixed effect. Instruction type was included as a random intercept for item, because it was a between-participant fixed effect. Our fixed effects were included in this model, because we were *a priori* interested in their contribution to the outcome (Gelman & Hill, 2007). Additionally, as this model does not provide an outcome for the learning effect in just the exposure only group, we ran another generalized linear regression model in this group only to see whether we replicated the findings from our previous studies. This model took the responses from the task (1 for a correct answer, 0 for an incorrect answer) as a dependent variable, marker type as a within-participants fixed effect, participant as a between-participants random effect and item as a within-participants random effect. In both models, orthogonal sum-to-zero contrast coding was applied to our binary fixed effects (i.e., marker type and instruction type; Baguley, 2012, pp. 590-621). We aimed to keep both models as fully specified as possible by including random intercepts for participants and items, but the model would not fit including random slopes (Barr et al., 2013). We increased the number of possible iterations to 100.000 (Powell, 2009), to solve potential issues with non-converging models. This enabled us to report on a random effect structure justified by our data (Jaeger, 2009). We report simple rather than standardized effect sizes (Baguley, 2009) and Wald confidence intervals (Agresti & Coull, 1998).[2]

---

[2]See our scripts for other statistical output, such as unconverted estimates and standard errors.

To investigate whether explicit instruction influenced looking behavior during the picture matching task, we carried out a cluster based permutation analysis (see Curcic et al., 2019; Dink & Ferguson, 2015; Maris & Oostenveld, 2007) using the eyetrackingR package (Dink & Ferguson, 2015). This analysis enabled us to investigate the development of the proportion of looks toward the target picture over time. We analyzed trials from 500 ms before the onset of the grammatical marker until 4000 ms after the onset of the marker (see Figures 3–6 for a visual illustration of this time window). On average a grammatical marker lasted 350 ms, there was approximately 200 ms between the offset of the marker and the onset of the noun, and a noun lasted about 550 ms. A sentence thus took around 1100 ms from the onset of the marker. This means the endpoint of the frame we analyzed was approximately 3000 ms after the offset of the sentence. As children often took at least 3000 ms to give an answer during the task, we have eye track data for the majority of the test items in this timeframe. After this point in time, we have fewer data points available, as children might have given an answer, and the next test item would be played. When interpreting the eye tracking data, it is important to know that in this age group it takes approximately 400 ms before eye movements reflect the processing of a presented stimulus (Clackson et al., 2011; Fukushima et al., 2000).

For the cluster based permutation analysis, we analyzed all test items, regardless of whether participants gave a correct or incorrect answer during the picture matching task. For every item, we coded whether a single look was directed at the target picture (1) or not (0). The area of interest for both types of pictures was the whole area within the colored frames. We removed all looks that were irrelevant (14.11%), i.e., on screen but not on either of the two pictures, or missing (12.55%). After removing looks that were classified as either irrelevant or missing, we excluded trials in which more than 50% of the looks had been removed as a result of this procedure (9.1% of the total number of test items). This analysis, which involves two steps, allowed us to determine at which parts within the marked time frame, participants are likely to look more often toward the target picture for sentences with *pli* than for sentences with *tra*. In the first step, the data are split into 50 ms (6 samples) time bins, and by calculating a *t*-value for each time bin we determined whether participants looked significantly more often toward the target picture for *pli* sentences than for *tra* sentences. Adjacent time bins for which there is a significant effect are clustered together and the *t*-values are summed. In a second step, the data are reshuffled hundreds of times, to see how likely these clusters and accompanying summed statistics would be if the data were randomly generated. This way, *p*-values can be calculated which indicate how likely the observed clusters and summed *t*-values would be observed if the data were generated randomly.

Using this procedure, we can correct for false alarms, while not losing much of the sensitivity of the data. The expectation was that children who acquired the plurality marker would be more likely to look at the correct image in *pli* sentences than in *tra* sentences. Because we expected the effect to occur in this direction, we used a cutoff point of 0.05 for significance. Unfortunately, a cluster based permutation analysis only allows for one predictor and was thus unsuitable for an analysis in which we looked at the effect of marker and instruction type simultaneously.[3] Therefore, the data for our explicitly instructed group and the uninstructed group were analyzed separately. We had no clear expectations about the effect of instruction on eye movement behavior. Possible outcomes could be that instruction type would lead to predictive eye movements in one group, but not in the other, or that such prediction would occur at different moment in time.

As a check on whether referential processing was similar in both groups, we used a similar procedure to analyze filler items. For this cluster based permutation analysis, we analyzed all fillers, regardless of whether participants gave a correct or incorrect answer during the picture matching task. For every item, we coded whether a single look was directed at the target picture (1) or not (0). We removed all looks that were irrelevant (14.21%), i.e., on screen but not on either of the two pictures, or

---

[3]Chan et al. (2018), for example, compare a cluster based permutation analysis, similar to our analysis, to a regression model in which more than one predictor can be taken into account, but they argue the latter analysis is a weaker one for this type of data. This may be because scores need to be aggregated to make the data suitable for a regression analysis.

missing (11.23%). After removing looks that were classified as either irrelevant or missing, we excluded trials in which more than 50% of the looks had been removed as a result of this procedure (4.61% of the total number of fillers). However instead of comparing looks between *pli* and *tra* sentences, we here compared the proportion of looks toward the target for fillers sentences between instructed children and uninstructed children. If our groups are comparable in their general referential processing, we should not observe any differences between groups with regard to their eye movements during these filler sentences.

Results

An overview of the descriptive statistics from the picture matching task can be found in Table 3 and Figure 2. For this task we investigated whether children would give more correct answers for sentences with *pli* than sentences with *tra*. This would reflect learning of the grammatical marker. We also considered whether this behavior interacted with the type of instruction children received. The table shows that participants were slightly more likely to give target answers when sentences contained *pli* than when sentences contained *tra*, but this effect was not significant ($OR = 1.236$, 95% CI = [0.873, 1.750], $z = 1.194$, $p = .233$). The results also indicate that explicitly instructed learners gave slightly fewer target answers overall than learners who received exposure only, but this effect was not significant either ($OR = 0.938$, 95% CI = [0.774, 1.136], $z = -0.658$, $p = .511$). Additionally, we observed a non-significant interaction between type of instruction and grammatical marker

Table 3. Scores from the picture matching task indicating the number of target answers produced by children who received explicit instruction and exposure, and children who received only exposure. Sentences with grammatical marker *pli* were predictable, which should lead to more correct target answers in the case of learning. Scores could range from 0 to 12.

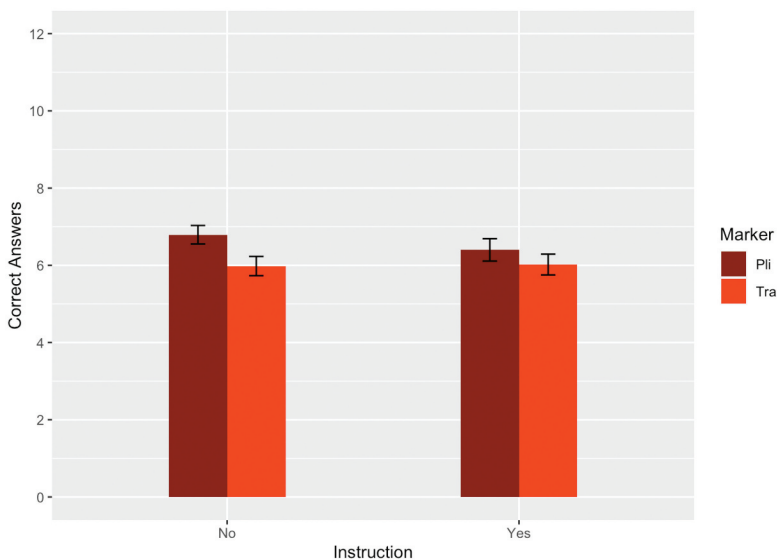|	Explicit Instruction (N = 50)			Exposure only (N = 53)		
	M	SD	Range	M	SD	Range
Pli	6.40	2.07	1–11	6.79	1.78	3–11
Tra	6.02	1.91	0–10	5.98	1.79	2–10

Figure 2. A graph depicting the results from the picture matching task for uninstructed (left) and explicitly instructed (right) learners. Scores from the picture matching task indicated the number of target answers. Scores could range from 0–12.

($OR = 0.855$, 95% CI = [0.618, 1.183], $z = -0.944$, $p = .345$). Although explicitly instructed learners showed less learning than their peers who received exposure only, this effect was very small and not significant. Furthermore, a mixed-effects model that was constructed to investigate learning in just the exposure only group, showed that these learners gave more correct answers for sentences with *pli* than for sentences containing *tra* ($OR = 1.334$, 95% CI = [0.913, 1.948], $z = 1.490$, $p = .136$). Again, this effect was not significant.

As a first step in the analysis of the eye tracking data, we looked at the explicitly instructed learners and learners who received only exposure together, to see whether their eye movements revealed sensitivity to the grammatical rule. To test this, we conducted a cluster based permutation analysis with the proportion of looks toward the target picture as a dependent variable, and grammatical marker as a within-participants factor. Figure 3 and Table 4 show the results of this analysis. In the first step of the analysis, we observed that participants behaved differently for sentences with *pli* than for sentences with *tra* during five clusters of time bins. In these cases, positive sum statistics were observed indicating that participants were more likely to look toward the target picture in *pli* sentences than in *tra* sentences. In the second step we found that participants looked significantly more often toward the target picture for sentences with *pli* than for sentences with *tra* in the cluster that lasted 1150 ms, and started at 700 ms after the onset of the marker till 1850 after this onset. This cluster starts 150 ms after
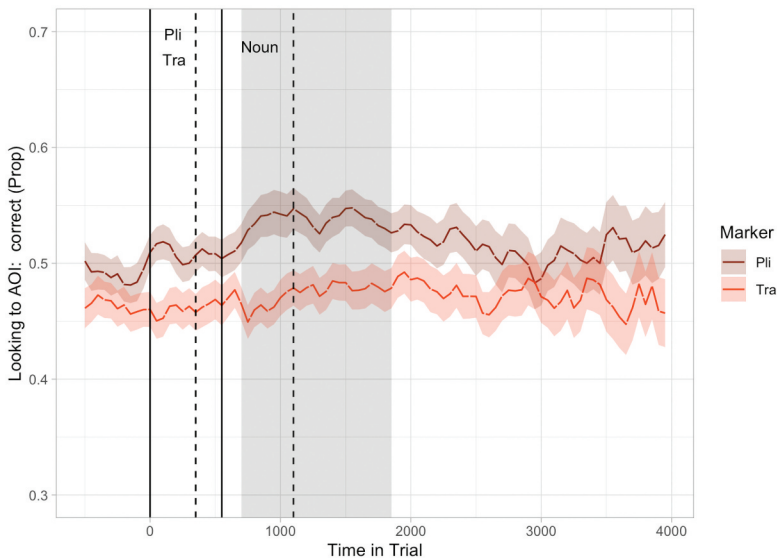
Figure 3. Proportion of looks toward the target picture over time for sentences with *pli* and sentences with *tra* for all learners together. Proportion of looks was calculated from all looks that were either on the target or on the non-target picture. Solid lines indicate the onset of the grammatical marker (*pli* or *tra*) and the noun respectively, dashed lines the offsets of these words. The shaded area indicates the cluster of time bins during which an effect of grammatical marker was observed.

Table 4. Results of the cluster based permutation analysis for explicitly instructed learners and learners who received only exposure together. The time range is measured from the onset of the grammatical marker. The cluster in bold turned out significant after the second step of the analysis.

Time cluster	Sum statistic	Time range in ms	Probability
1	11.06	0–200	.264
2	6.74	350–500	.417
3	**64.75**	**700–1850**	**.001**
4	2.09	2550–2600	.672
5	2.04	3950–4000	.704
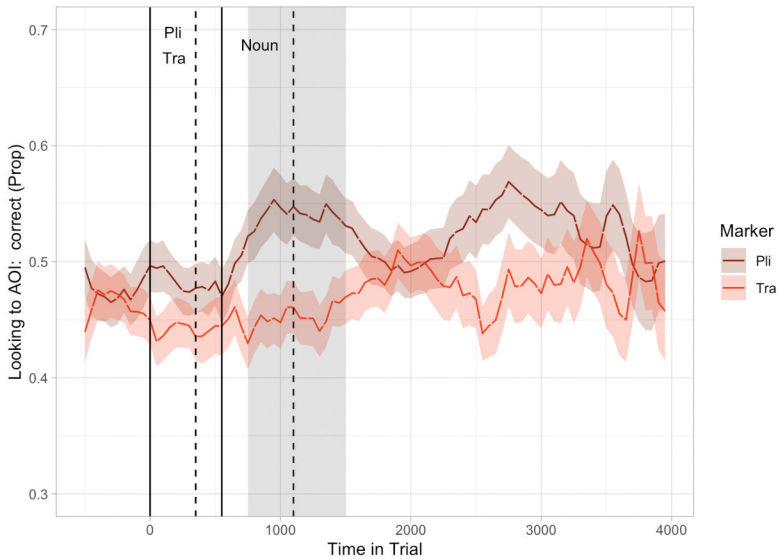
Figure 4. Proportion of looks toward the target picture over time for sentences with *pli* and sentences with *tra* for the explicitly instructed group. Proportion of looks was calculated from looks that were on the target or on the non-target picture. Solid lines indicate the onset of the grammatical marker (*pli* or *tra*) and the noun respectively, dashed lines the offsets of these words. The shaded area indicates the cluster of time bins during which an effect of grammatical marker was observed.

Table 5. Results of the cluster based permutation analysis for the explicitly instructed learners. The time range is measured from the onset of the grammatical marker. The cluster in bold turned out significant after the second step of the analysis.

Time cluster	Sum statistic	Time range in ms	Probability
1	4.48	50–150	.534
2	**41.62**	**750–1500**	**.009**
3	10.62	2550–2750	.266
4	2.22	2800–2850	.675

the onset of the noun and lasts till approximately 750 ms after the offset of the noun, and thus the end of the sentence.

This analysis was repeated for the explicitly instructed and exposure only groups separately. Figure 4 and Table 5 show the results of the analysis for the explicitly instructed group. In the first step of this analysis, we observed four clusters of time bins during which participants looked more often toward the target picture for sentences with *pli* than for sentences with *tra*. In the second step we found that this effect was only significant in the cluster that lasted from 750 ms after the onset of the marker till 1500 after this onset. This cluster starts 150 ms after the onset of the noun and lasts till approximately 400 ms after the offset of the noun, and thus the end of the sentence. Figure 5 and Table 6 show the results of the analysis for the exposure only group. In the first step of this analysis, we observed four clusters of time bins during which participants looked more often toward the target picture for sentences with *pli* than for sentences with *tra*. In the second step, we found this effect reached significance for the cluster that lasted from 1500 ms after the onset of the marker till 2050 after this onset. This cluster starts 400 ms after the offset of the noun and lasts till approximately 950 ms after the offset of the noun, and thus the end of the sentence. In both groups, learners show looking behavior that might be associated with learning at some point in time. Although we were not able to make a direct comparison between the two groups, the analyses from the separate group suggest the moment at which this behavior occurs might be earlier in time for the explicitly instructed group than for the exposure only group.
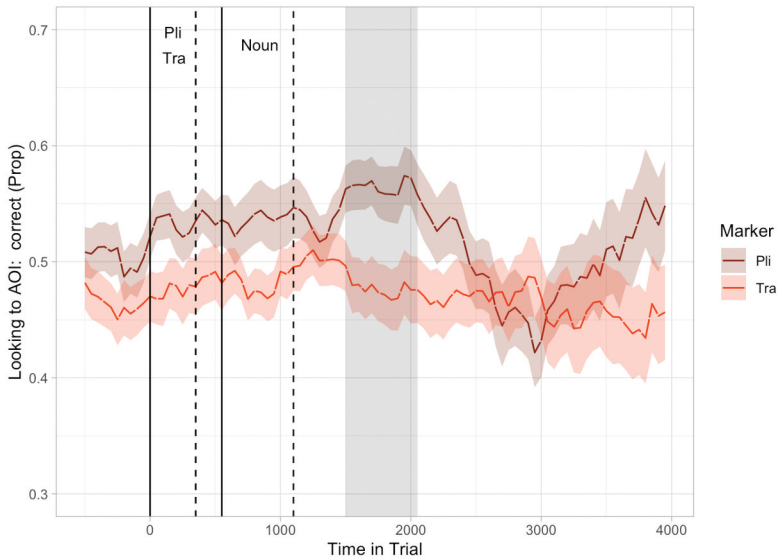
Figure 5. Proportion of looks toward the target picture over time for sentences with *pli* and sentences with *tra* for the exposure only group. Proportion of looks was calculated from looks that were on the target or on the non-target picture. Solid lines indicate the onset of the grammatical marker (*pli* or *tra*) and the noun respectively, dashed lines the offsets of these words. The shaded area indicates the cluster of time bins during which an effect of grammatical marker was observed.

Table 6. Results of the cluster based permutation analysis for the exposure only group. The time range is measured from the onset of the grammatical marker. The cluster in bold turned out nearly significant after the second step of the analysis.

Time cluster	Sum statistic	Time range in ms	Probability
1	2.18	−250 – −200	.605
2	4.33	50–150	.496
3	**27.83**	**1500–2050**	**.053**
4	2.41	3800–3850	.555

There is a possibility that the differences in behavior observed between our groups were caused by a general referential processing difference between groups. To investigate this, we ran an additional cluster based analysis, in which we compared looks toward the target picture in filler items between the instructed and uninstructed groups. A difference between groups on the filler items might indicate that one group consisted of faster processers. Figure 6 and Table 7 show the results of the analysis. In the first step of this analysis, we observed three small clusters of time bins during which participants from one group looked more often toward the target picture for fillers than the other. In second step, none of these three clusters turned out to be significant. Because these clusters were small and non-significant, we assume that the differences in processing of the grammatical markers between groups probably were not caused by a general referential processing difference between instructed and uninstructed children.[4]

---

[4]Ideally, this analysis would only have included filler items for which participants gave correct answers. However, participants gave a correct answer for approximately half of the filler items. As a result, we ended up with a rather low number of items that we could use for this additional analysis. We did run this analysis, but did not report in the article because of low statistical power. It can be found in the supplementary materials. The outcomes show a similar picture as presented in the analysis that used all (correctly and incorrectly answered) fillers.
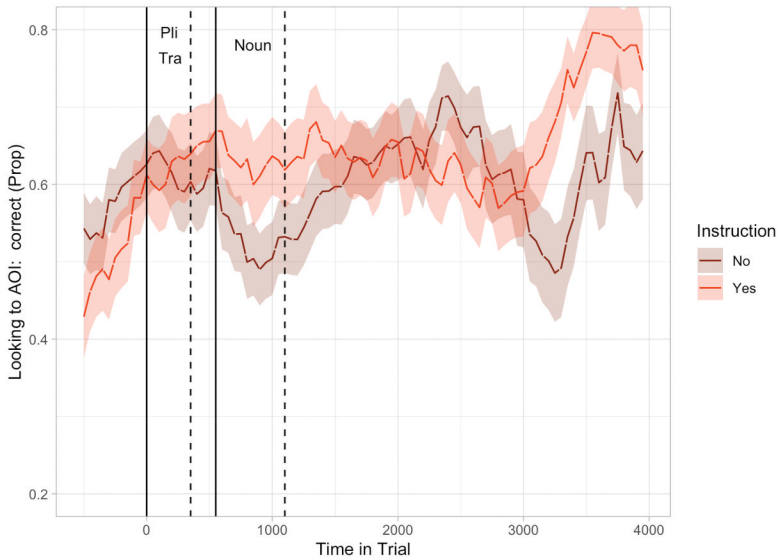
Figure 6. Proportion of looks toward the target picture over time for filler sentences for all learners together. Proportion of looks was calculated from all looks that were either on the target or on the non-target picture. Solid lines indicate the onset of the grammatical marker (*pli* or *tra*) and the noun respectively, dashed lines the offsets of these words.

Table 7. Results of the cluster based permutation analysis for all learners on the filler items. The time range is measured from the onset of the grammatical marker.

Time cluster	Sum statistic	Time range in ms	Probability
1	2.00	2450–2500	.726
2	4.10	2600–2700	.562
3	−6.75	3550–3700	.362

In the debriefing, when children were asked how they made a decision on the picture matching task, participants either reported they did not know how they made a decision (N = 83) or they claimed that they had heard the sentences before and remembered what they meant (N = 20) When asked for the meaning of *pli* and *tra*, most children either said they did not know or could not give a meaning (N = 81), gave the meaning of a noun they learned during the exposure (N = 13), or some other meaning (N = 7). However, when asked what *pli* meant, one child in the exposure only group answered 'to calculate'; when asked what *tra* meant, one child in the explicitly instructed group answered 'one, two or three,' which is more or less what the marker means.

Discussion

Many scholars have claimed there is a fundamental difference between children and adults learning the grammatical structures of a new language (Bley-Vroman, 1990; Hartshorne et al., 2018; Krashen et al., 1982; Singleton & Ryan, 2004). Adults would rely mostly on explicit learning mechanisms, but can combine these with implicit learning mechanisms (DeKeyser, 2000, 2003; R. Ellis, 2005, 2009; Paradis, 2004, 2009; Ullman, 2001), whereas children would only use implicit learning mechanisms. This implies that children do not learn grammar explicitly, and that explicit instruction about the grammatical regularities in the target language does not enhance learning (Bialystok, 1994; DeKeyser &

Larson-Hall, 2005; Paradis, 2004, 2009). The goal of this paper was to investigate whether children indeed do not learn grammatical structures explicitly, by testing the effect of explicit instruction on the acquisition of a morphosyntactic element in kindergartners. As explicit instruction potentially could influence both more explicit and more implicit knowledge, we measured learning using a picture matching task and eye tracking, because these measures are associated more with explicit and implicit knowledge respectively (Andringa & Rebuschat, 2015; Godfroid et al., 2015). While investigating the influence of explicit instruction on acquiring a morphosyntactic element in kindergartners, this study also enabled us to replicate our findings from earlier studies (Spit et al., 2020, 2021), in which we demonstrated that a meaningful agreement marker could be learned when children did not receive explicit instruction about this marker.

More evidence of learning the meaningful agreement marker would be provided if children in the exposure only group gave more correct answers during the picture matching task for sentences with the marker *pli*, as it was predictive of the target answer, than for sentences with the marker *tra*, which was not predictive of the target answer. When we isolated this group to allow for comparison with Spit et al. (2020); (2021), we saw they showed this behavior, but the effect was not significant. Although these results could be seen as null results, which suggest that we were unable to replicate our previous findings, they are compatible with the outcomes of the two previous studies, which are listed in Table 8. The effect size of the current study is similar to that of the previous two studies, and the effect sizes of the three studies fall within each other's confidence intervals. Taken together with the previous studies, we might think that these and our previous results show that kindergartners can acquire this meaningful grammatical marker from mere exposure to limited input. Perhaps, it was more difficult for children to show their knowledge of the grammatical marker in this experiment, because they had to answer more indirectly during the task. In our previous experiments children could point toward the screen to indicate their answer, whereas now they had to name the color of the frame, because pointing would interfere with their eye movements. Naming colors might be cognitively more demanding, and likely requires more explicit processing. As a result, the current task may tap more into the explicit knowledge that participants possibly develop during learning (Andringa & Rebuschat, 2015; Bowles, 2011; R. Ellis, 2005; Godfroid, Loewen, Jung, Park, Gass &, 2015; Y. Han & Ellis, 1998). Alternatively, it may have been difficult to establish the learning effect in the picture matching task with sufficient certainty in each individual study, because the learning effect observed in each of the studies is really small.

Replicating the learning effect in the exposure only group was not the main goal of this paper. We were primarily interested to see whether there could be a difference between the explicitly instructed group and the group that received exposure only. The interaction effect between learning and instruction type was only small and not significant, but the direction of the effect suggests that learning might be hindered by explicit instruction, as children who received rule explanation showed slightly less learning. Furthermore, the confidence interval of this interaction effect (95% CI = [0.618, 1.183]) suggests that an effect in the opposite direction could occur, but would likely not be very big either. Although we cannot interpret this null finding as indicating there is no effect of instruction type of learning, we might conclude it indicates that explicit instruction does not make a meaningful difference. Replication of this study would be desirable to make stronger claims about the possible effect of explicit instruction on learning, perhaps using a method that might generate larger learning

Table 8. Results from the generalized logistic effects model for the exposure only group only of our experiment, and two earlier studies that we were aiming to replicate with this part of the experiment.

Study	OR	95% CI	z	p
1	1.515	1.071–2.257	2.178	.029
2	1.344	0.980–1.843	1.834	.067
Current	1.334	0.913–1.948	1.490	.136

effects (i.e., by extending the exposure). Nevertheless, if the results of the present study are upheld, then that might mean that children learn more effective implicitly, and do not make use of explicit learning mechanisms (Bialystok, 1994; DeKeyser & Larson-Hall, 2005; Paradis, 2004, 2009).

The eye tracking data seem to present a slightly different picture. When measuring the eye movements of the children during the picture matching task, we expected that if children learned the meaningful grammatical element this would lead to more looks toward the target for sentences with *pli* than for sentences with *tra* at some point in the sentence after the onset of these markers. Importantly, we had no specific expectations for the effect of instruction on this looking behavior. We observed the expected looking behavior in the group as a whole, as children looked more often at the target from 700 to 1850 ms after the onset of the marker. When we interpret these data, it is important to take into account that in kindergartners it takes approximately 400 ms before eye movements reflect the processing of a presented stimulus (Clackson et al., 2011; Fukushima et al., 2000). When interpreting the present findings in the light of this knowledge, the start of the observed time cluster seems to have coincided with the offset of the grammatical marker, which suggests that it indeed triggered looks toward the target picture. The eye movements of all learners taken together seemingly reveal they develop some knowledge of the grammatical marker. Moreover, the discrepancy in results between the picture matching task (no effects) and the eye tracking task underlines to idea that both types of measures are ideally included in studies like these (Andringa & Rebuschat, 2015; Godfroid et al., 2015). These findings may also be taken to mean that eye tracking was a more sensitive measure of learning in this study.

Interestingly, we observed the marker-induced looking behavior at different moments in time for the explicitly instructed group and the exposure only group. Explicitly instructed learners looked significantly more often toward the target picture in sentences with *pli* from 750 to 1500 ms after the onset of the grammatical marker. When applying the 400 ms correction to these data, explicitly instructed learners' preference for looking at the correct target coincided with the offset of the grammatical marker, which is about the first moment one can expect such looking behavior, until approximately the end of the sentence. This looking behavior emerged later in the learners who received exposure only, namely between 1500 and 2050 ms after the onset of the grammatical marker, and this effect approached significance ($p = .053$). Applying the 400 ms correction, we might say they started looking toward the correct picture in predictable sentences when the sentence ended. In short, both groups seemed to show eye movements that are indicative of learning, but explicitly instructed children showed this earlier than their peers who merely received exposure. Although we should remain cautious in interpreting these results, especially because they are borderline significant for the exposure only group and we did not directly compare the learning effect in the two experimental groups, it seems that the picture matching task accuracy data and eye tracking data stand in contrast to each other. While the accuracy data provided no evidence of an effect of explicit instruction on learning, the eye tracking data are compatible with the idea the children made use of explicit learning mechanisms (Bialystok, 1994; DeKeyser & Larson-Hall, 2005; Paradis, 2004, 2009). These results might show that explicit instruction lead to slightly better developed knowledge of the acquired grammatical element, which resulted in effects that appeared 750 ms earlier. Importantly, because we did not make a direct comparison between groups, we should also be cautious in interpreting the difference in timing between instructed and uninstructed children. Replicating this experiment is therefore highly necessary.

The question remains why we were unable to find a similar learning effect in the picture matching task. One potential explanation of our findings is that exposure to the miniature language over this short period of time (only 30 minutes) led to low level or unstable representations of the acquired linguistic knowledge. In addition, the instruction could have been too brief, as it lasted only 30 seconds. Because of its brevity, it might have merely raised attention for the form of the markers, rather than their function. Following this line of reasoning, we may have been unable to show an effect of explicit instruction on learning in the picture matching task because the children's linguistic knowledge was not developed strongly enough. It is possible that children need more exposure to develop stronger representations, or that effects of explicit instruction on acquisition become visible in accuracy over a

longer period of time. There are good reasons to assume learning not only happens during training, but also afterward when a learner is no longer involved in the learning task (Diekelmann & Born, 2010; Maquet, 2001; Nieuwenhuis et al., 2013; Stickgold, 2005). Knowledge that is initially stored in short term memory needs to be strengthened and consolidated in long term memory (Dumay & Gaskell, 2007; Frost & Monaghan, 2017; Gómez et al., 2006; Mirković & Gaskell, 2016; S.E. Williams & Horst, 2014). If these children were tested at a later moment in time or were given more exposure to the language, we might obtain stronger effects that become visible in the picture matching task as well. This could imply that the eye movements of the children in the current experiment are indicative of knowledge of the grammatical marker that is still under development. Future research should delve deeper into this issue, for example, by manipulating the amount of exposure that explicitly instructed children receive, or by testing children one or multiple days after a first training session.

Alternatively, our results could point toward a difference in learning between children and adults. Although children can make use of explicit instruction, this might not result in more explicit knowledge, as we typically see in adults (e.g., Hamrick & Rebuschat, 2012; Morgan-Short, Steinhauer, Sanz & Ullmann, 2012; Williams, 2005). Some authors claimed explicit instruction could also lead to implicit knowledge (Batstone, 2002; DeKeyser, 2009; Williams, 2009), and perhaps this applies to children in particular. The fundamental difference between children and adults might not lie in the learning strategies available to these learners, but in the type of knowledge they develop. This might explain why we see a difference in their eye movements, which are associated more with implicit knowledge, but hardly any in the picture matching task, which is associated more with explicit knowledge. This latter finding could have been reinforced by the more indirect nature of the task in comparison to our two previous studies. As children had to respond by naming the colors of the frames, explicit processing might have played a bigger role, and possible effects of explicit instruction on this task were more difficult to observe because we were less likely to tap into implicit knowledge. Furthermore, it explains why only a single child was able to report on the meaning of one of the two markers during the debriefing. In this view, our findings entail that children are able to make use of explicit instruction, but unlikely to develop explicit knowledge. Rather, children only develop implicit knowledge, and maybe even more when explicitly instructed, but this can only be observed when knowledge is measured appropriately. Possibly, children can use explicit learning mechanisms, but because they learn language in completely different contexts than adults, will not use them very often (Lichtman, 2016). However, to justify these speculative claims, a direct comparison between children and adults would be necessary, as we currently do not know whether adults would develop more explicit knowledge in this experiment.

The way we just characterized the fundamental difference between child and adult language learning relies on a dichotomous distinction between implicit and explicit knowledge. It assumes that when children cannot verbalize their knowledge, they are not aware of it and the knowledge is therefore implicit. However, awareness likely exists in gradations (Allport, 1988; Cleeremans, 2008, 2011, 2014; Dehaene et al., 2017; Dennett, 1993; James, 1890), and learners can also be aware of knowledge despite the fact that they cannot reflect on it verbally. Perhaps children develop knowledge that is explicit and available to awareness, but at a lower, non-verbalizable level. Possibly, we did not tap into implicit knowledge with the eye tracking measure, because we measured eye movements *during* the picture matching task, which they had to carry out quite consciously. Following this rationale, we might argue that the earlier predictive eye movements of the instructed children indicate increased awareness of the target regularity and development of some explicit knowledge. Yet, this knowledge is not available to children's awareness to be beneficial when they have to make a more conscious decision. As a result, explicit instruction did not lead to a meaningful difference in accuracy during the picture matching task. This would mean that a possible effect of explicit instruction is not fundamentally different for children and adults, but only gradually. In both groups explicit instruction may lead to an increase in explicit knowledge, but children lack the ability to verbalize this acquired knowledge. Future research could include more fine grained measures of awareness (see Spit Andringa, Rispens, Aboh et al., 2019 for an

example of such a measure) to investigate whether explicit instruction in young children indeed leads to an increase in knowledge that is explicit in nature, but on which they simply cannot reflect verbally.

Conclusion

Although many authors suggest children do not learn language explicitly, there is little experimental evidence for such claims. In this paper, we used a miniature language learning experiment to test whether children can learn explicitly, by investigating if explicit instruction influences kindergartners' learning of a grammatical marker. In a picture matching task, we found no clear evidence of an effect of explicit instruction on accuracy, but there was an effect on children's eye movements. These data seem incompatible with the idea that explicit instruction does not make a difference when kindergartners learn a grammatical element, and tentatively suggest that explicit instruction has a different effect on explicit knowledge than on implicit knowledge in this age group. However, learning effects were also relatively small, and we were unable to directly test the interaction effect between learning and instruction type in our eye tracking analyses. More research is needed to back up these findings. Perhaps the effect of explicit instruction in this age group becomes clearer after a longer period of exposure or consolidation, or perhaps children develop a different type of knowledge than adults when explicitly instructed. We hope that more studies on this topic will shed light on whether we are right in claiming that explicit instruction influences learning in children, and that, if this is the case, such research can give insight in the exact effect explicit instruction has on learning in this age group.

ORCID

Sybren Spit http://orcid.org/0000-0002-9724-8912
Sible Andringa http://orcid.org/0000-0002-0182-6977

References

Aboh, E. O. (2004). *The morphosyntax of complement-head sequences. Clause structures and word order patterns in Kwa.* Oxford University Press.

Aboh, E. O., & DeGraff, M. (2014). Some notes on bare noun phrases in Haitian Creole and Gungbe. In T. A. Åfarlí & B. Maehlum (Eds.), *The Sociolinguistics of Grammar,* 206-236. Amsterdam: John Benjamins Publishing Company.

Agresti, A., & Coull, B. A. (1998). Approximate is better than "exact" for interval estimation of binomial proportions. *The American Statistician, 52*(2), 119–126. https://doi.org/10.2307/2685469

Alcón Soler, E. (2005). Does instruction work for learning pragmatics in the EFL context? *System, 33*(3), 417–435. https://doi.org/10.1016/j.system.2005.06.005

Allport, A. (1988). What concept of consciousness? In A. J. Marcel & E. Bisiach (Eds.), *Consciousness in contemporary science* (pp. 159–182). Clarendon Press.

Andringa, S., & Dabrowska, E. (2019). Individual differences in first and second language ultimate attainment and their causes. *Language Learning*, 69(1), 5–12. https://doi.org/10.1111/lang.12328

Andringa, S., de Glopper, C. M., & Hacquebord, H. (2011). Effect of explicit and implicit instruction on free written reponses task performance. *Language Learning*, 61(3), 868–903. https://doi.org/10.1111/j.1467-9922.2010.00623.x

Andringa, S., & Rebuschat, P. (2015). New directions in the study of implicit and explicit learning. An introduction. *Studies in Second Language Acquisition*, 37(2), 185–196. https://doi.org/10.1017/S027226311500008X

Appah, C. K. I. 2003. *Nominal derivation in Akan. A descriptive analysis*. M.A. Thesis, Norwegian University of Science and Technology.

Arciuli, J., & Simpson, I. C. (2012). Statistical learning is related to reading ability in children and adults. *Cognitive Science*, 36(2), 286–304. https://doi.org/10.1111/j.1551-6709.2011.01200.x

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, 9(4), 321–324. https://doi.org/10.1111/1467-9280.00063

Bader, M., & Häussler, J. (2010). Toward a model of grammaticality judgments. *Journal of Linguistics*, 46(20), 273–330. https://doi.org/10.1017/S0022226709990260

Baguley, T. (2009). Standardized or simple effect size: What should be reported? *British Journal of Psychology*, 100(3), 603–617. https://doi.org/10.1348/000712608X377117

Baguley, T. (2012). *Serious Stats*. Palgrave Macmillan.

Baker, C. I., Olson, C. R., & Behrmann, M. (2004). Role of attention and perceptual grouping in visual statistical learning. *Psychological Science*, 15(7), 460–466. https://doi.org/10.1111/j.0956-7976.2004.00702.x

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed- effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Batstone, R. (2002). Making sense of new language: A discourse perspective. *Language Awareness*, 11(1), 14–29. https://doi.org/10.1080/09658410208667043

Bialystok, E. (1994). Representation and ways of knowing: Three issues in second language acquisition. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 549–569). Academic Press.

Birdsong, D. (2005). Nativelikeness and non-nativelikeness in L2A research. *IRAL - International Review of Applied Linguistics in Language Teaching*, 43(4), 319–328. https://doi.org/10.1515/iral.2005.43.4.319

Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning*, 56, 9–49. https://doi.org/10.1111/j.1467-9922.2006.00353.x

Birdsong, D., & Vanhove, J. (2016). Age of second language acquisition: Critical periods and social concerns. In S. Nicoladis & E. Montanari (Eds.), *Bilingualism across the lifespan: Factors moderating language proficiency* (pp. 163–181). American Psychological Association.

Bley-Vroman, R. (1990). The logical problem of foreign language learning. *Linguistic Analysis*, 20(1–2), 3–49. https://doi.org/10.1017/CBO9781139524544.005

Blom, E., & Unsworth, S. (2010). *Experimental methods in language acquisition research*. John Benjamins.

Bowles, M. A. (2011). Measuring implicit and explicit knowledge. *Studies in Second Language Acquisition*, 33(2), 247–271. https://doi.org/10.1017/S0272263110000756

Brooks, P. J., & Kempe, V. (2019). More is more in language learning: Reconsidering the less-is-more hypothesis. *Language Learning*, 69(1), 13–41. https://doi.org/10.1111/lang.12320

Chan, A., Yang, W., Chang, F., & Kidd, E. (2018). Four-year-old Cantonese-speaking children's online processing of relative clauses: A permutation analysis. *Journal of Child Language*, 45(1), 174–203. https://doi.org/10.1017/S0305000917000198

Chen, C. H., Gershkoff-Stowe, L., Wu, C. Y., Cheung, H., & Yu, C. (2017). Tracking multiple statistics: Simultaneous learning of object names and categories in English and Mandarin Speakers. *Cognitive Science*, 41(6), 1485–1509. https://doi.org/10.1111/cogs.12417

Clackson, K., Felser, C., & Clahsen, H. (2011). Chidlren's processing of reflexives and pronouns in English: Evidence from eye-movements during listening. *Journal of Memory and Language*, 65(2), 128–144. https://doi.org/10.1016/j.jml.2011.04.007

Cleeremans, A. (2008). Consciousness: The radical plasticity thesis. In R. Banerjee & B. K. Chakrabarti (Eds.), *Progress in brain research* (Vol. 168, pp. 19–33). Elsevier.

Cleeremans, A. (2011). The radical plasticity thesis: How the brain learns to be conscious. *Frontiers in Psychology*, 2(86), 1–12. https://doi.org/10.3389/fpsyg.2011.00086

Cleeremans, A. (2014). Connection conscious and unconscious processing. *Cognitive Science*, 38(6), 1286–1315. https://doi.org/10.1111/cogs.12149

Conway, C. M., Bauernschmidt, A., Huang, S. S., & Pisoni, D. B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition*, 114(3), 356–371. https://doi.org/10.1016/j.cognition.2009.10.009

Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities: Pitting abstract against stimulus-specific representations. *Psychological Science*, *17*(10), 905–912. https://doi.org/10.1111/j.1467-9280.2006.01801.x

Coyne, M. D., McCoach, D. B., Loftus, S., Zipoli, R., Jr, & Kapp, S. (2009). Direct vocabulary instruction in Kindergarten: Teaching for breadth, versus depth. *The Elementary School Journal*, *110*(1), 1–48. https://doi.org/10.1086/598840

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, *25*(1), 7–29. https://doi.org/10.1177/0956797613504966

Curcic, M., Andringa, S., & Kuiken, F. (2019). The role of awareness and cognitive aptitudes in L2 predictive language processing. *Language Learning*, *69*(1), 42–71. https://doi.org/10.1111/lang.12321

Degé, F., & Schwarzer, G. (2011). The effect of a music program on phonological awareness in preschoolers. *Frontiers in Psychology*, *2*, 124. https://doi.org/10.3389/fpsyg.2011.00124

Dehaene, S., Lau, H., & Kouider, S. (2017). What is consciousness, and could machines have it? *Science*, *358*(6362), 486–492. https://doi.org/10.1126/science.aan8871

DeKeyser, R. (1995). Learning second language grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition*, *17*(3), 379–410. https://doi.org/10.1017/S027226310001425X

DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, *22*(4), 499–533. https://doi.org/10.1017/S0272263100004022

DeKeyser, R. (2003). Implicit and Explicit learning. In C. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 312–348). Blackwell.

DeKeyser, R., & Larson-Hall, J. (2005). What does the critical period really mean? In J. F. Kroll & A. M. D. de Groot (Eds.), *Handbook of bilingualism psycholinguistic approaches* (pp. 88–108). Oxford University Press.

DeKeyser, R. (2009). Cognitive-psychological processes in second language learning. In M. H. Long & C. J. Doughty (Eds.), *Handbook of second language teaching* (pp. 119–138). Wiley-Blackwell.

Dennett, D. (1993). *Concsiousness explained*. Penguin UK.

Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience*, *11*(2), 114–126. https://doi.org/10.1038/nrn2762

Dink, J. W., & Ferguson, B. 2015. *eyetrackingR; An R library for eye-tracking data analysis.* http://www.eyetrackingr.com.

Dörnyei, Z. (2009). *The psychology of second language acquisition*. Oxford University Press.

Doughty, C. J. (2019). Cognitive language aptitude. *Language Learning*, *69*(1), 101–126. https://doi.org/10.1111/lang.12322

Dumay, N., & Gaskell, M. G. (2007). Sleep-associated changes in the mental representations of spoken words. *Psychological Science*, *18*(1), 35–39. https://doi.org/10.1111/j.1467-9280.2007.01845.x

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, *36*(3), 250–287. https://doi.org/10.1598/RRQ.36.3.2

Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, *27*(2), 305–352. https://doi.org/10.1017/S027226310505014X

Ellis, N. C. (2011). Implicit and explicit SLA and their interface. In C. Sanz & R. Loew (Eds.), *Implicit and explicit language learning: Conditions, processes, and knowledge in SLA & bilingualism* (pp. 35–47). Georgetown University Press.

Ellis, N. C. (2015). Implicit and explicit learning: Their dynamic interface and complexity. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 3–23). John Benjamins.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, *27*(2), 141–172. https://doi.org/10.1017/S0272263105050096

Ellis, R. (2009). Implicit and explicit learning, knowledge, and instruction. In R. Ellis, S. Loewen, C. Elder, R. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing, and teaching* (pp. 3–25). Multilingual Matters.

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, *28*(2), 339–368. https://doi.org/10.1017/S0272263106060141

Endress, A. D., & Bonatti, L. L. (2007). Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition*, *105*(2), 247–299. https://doi.org/10.1016/j.cognition.2006.09.010

Erickson, L. C., & Thiessen, E. D. (2015). Statistical learning of language: Theory, validity and predictions of a statistical learning account of language acquisition. *Developmental Review, 37,* 66-108. https://doi.org/10.1016/j.dr.2015.05.002

Frost, R. L. A., & Monaghan, P. (2016). Simultaneous segmentation and generalization of non-adjacent dependencies from continuous speech. *Cognition*, *147*, 70–74. https://doi.org/10.1016/j.cognition.2015.11.010

Frost, R. L. A., & Monaghan, P. (2017). Sleep-driven computations in speech processing. *PLOS ONE*, *12*(1), e0169538. https://doi.org/10.1371/journal.pone.0169538

Fukushima, J., Hatta, T., & Fukushima, K. (2000). Development of voluntary control of saccadic eye movements: I. Age-related changes in normal children. *Brain and Development*, *22*(3), 173–180. https://doi.org/10.1016/S0387-7604(00)00101-7

Fukuya, Y. J., & Martínez-Flor, A. (2009). The interactive effects of pragmatic-eliciting tasks and pragmatic instruction. *Foreign Language Annals*, *41*(3), 478–500. https://doi.org/10.1111/j.1944-9720.2008.tb03308.x

Furnes, B., & Samuelsson, S. (2011). Phonological awareness and rapid automatized naming predicting early development in reading and spelling: Results from a cross-linguistic longitudinal study. *Learning and Individual Differences*, *21*(1), 85–95. https://doi.org/10.1016/j.lindif.2010.10.005

Gass, S., Svetics, I., & Lemelin, S. (2003). Differential effects of attention. *Language Learning*, *53*(3), 497–546. https://doi.org/10.1111/1467-9922.00233

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Gillon, G. T. (2018). *Phonological awareness: From research to practice*. The Guilford Press.

Godfroid, A. (2015). The effects of implicit instruction on implicit and explicit knowledge development. *Studies in Second Language Acquisition*, *38*(2), 177–215. https://doi.org/10.1017/S0272263115000388

Godfroid, A., Loewen, S., Jung, S., & Park, J.-H. (2015). Timed and untimed grammaticality judgments measure distinct types of knowledge: Evidence from eye-movement patterns. *Studies in Second Language Acquisition*, *37*(2), 269–297. https://doi.org/10.1017/S0272263114000850

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*(5), 431–436. https://doi.org/10.1111/1467-9280.00476

Gómez, R. L., Bootzin, R. B., & Nadel, L. (2006). Naps promote abstraction in language-learning infants. *Psychological Science*, *17*(8), 670–674. https://doi.org/10.1111/j.1467-9280.2006.01764.x

Gómez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *70*(2), 109–135. https://doi.org/10.1016/S0010-0277(99)00003-7

Goo, J., Granena, G., Yilmaz, Y., & Novella, M. (2015). Implicit and explicit instruction in L2 learning. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 443–482). John Benjamins.

Granena, G., & Yilmaz, Y. (2019). Corrective feedback and the role of implicit sequence-learning ability in L2 online performance. *Language Learning*, *69*(1), 127–156. https://doi.org/10.1111/lang.12319

Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: A test of the critical-period hypothesis for second-language acquisition. *Psychological Science*, *14*(1), 31–38. https://doi.org/10.1111/1467-9280.01415

Hamrick, P., & Rebuschat, P. (2012). How implicit is statistical learning? In P. Rebuschat & J. Williams (Eds.), *Statistical learning and language acquisition* (pp. 365–382). Mouton de Gruyter.

Han, Y., & Ellis, R. (1998). Implicit knowledge, explicit knowledge and general language proficiency. *Language Teaching Research*, *2*(1), 1–23. https://doi.org/10.1177/136216889800200102

Han, Z., & Finneran, R. (2014). Re-engaging the interface debate: Strong, weak, none, or all? *International Journal of Applied Linguistics*, *24*(3), 370–389. https://doi.org/10.1111/ijal.12034

Hartshorne, J. K., Tenenbaum, J. B., & Pinker, S. (2018). A critical period for language acquisition: Evidence from 2/3 million English speakers. *Cognition*, *177*, 263–277. https://doi.org/10.1016/j.cognition.2018.04.007

Hoff, E. (2012). *Research methods in child language: A practical guide*. Wiley-Blackwell.

Housen, A. M., & Pierrard, M. (2006). Investigating instructed second language acquisition. In A. Housen & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (pp. 235–269). Mouton de Gruyter.

Hulstijn, J. (2005). Theoretical and empirical issues in the study of implicit and explicit language learning. *Second Language Acquisition*, *27(2)*, 129–140.

Hulstijn, J. (2015). Explaining phenomena of first and second language acquisition with the constructs of implicit and explicit learning. The virtues and pitfalls of a two-system view. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 25–46). John Benjamins.

Jaeger, T. F. *Post to HLP/Jaeger lab blog*. May 14 2009, https://hlplab.wordpress.com/2009/05/14/random-effect-structure/

James, W. (1890). *The principles of psychology*. Henry Holt and Company.

Johnson, C., & Goswami, U. (2010). Phonological awareness, vocabulary, and reading in deaf children with cochlear implants. *Journal of Speech, Language, and Hearing Research*, *53*(2), 237–261. https://doi.org/10.1044/1092-4388 (2009/08-0139)

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*(1), 60–99. https://doi.org/10.1016/0010-0285(89)90003-0

Kachergis, G., Yu, C., & Shiffrin, R. M. (2014). Cross-situational word learning is both implicit and strategic. *Frontiers in Psychology*, *5*, 588. https://doi.org/10.3389/fpsyg.2014.00588

Kaiser, E. (2013). Experimental paradigms in psycholinguistics. In R. J. Podesva & D. Sharma (Eds.), *Research methods in linguistics* (pp. 135–169). Cambridge University Press.

Kidd, E. (2011). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology*, *48*(1), 171–184. https://doi.org/10.1037/a0025405

Kissling, E. M. (2013). Teaching pronunciation: Is explicit phonetics instruction beneficial for FL learners? *The Modern Language Journal*, *97*(3), 720–744. https://doi.org/10.1111/j.1540-4781.2013.12029.x

Koike, D. A., & Pearson, L. (2005). The effect of instruction and feedback in the development of pragmatic competence. *System*, 33(3), 481–501. https://doi.org/10.1016/j.system.2005.06.008

Krashen, S. D., Scarcella, R. C., & Long, M. A. (1982). *Child-adult differences in second language acquisition*. Newbury House.

Lany, J., & Saffran, J. R. (2013). Statistical learning mechanisms in infancy. In J. L. R. Rubenstein & P. Rakic (Eds.), *Comprehensive developmental neuroscience: Neurel circuit development and function in the brain* (Vol. 3, pp. 231–248). Elsevier.

Lany, J. (2014). Judging words by their covers and the company they keep: Probabilistic cues support word learning. *Child Development*, 85(4), 1727–1739. https://doi.org/10.1111/cdev.12199

Léger, P.-M., Sénécal, S., Karpova, E., Brieugne, D., Di Fabio, M.-L., & Georges, V. (2018). Setup guidelines for eye tracking in child and teenager research in the context of learning by interaction with a tablet. *Neuroeducation*, 5(1), 33–40. https://doi.org/10.24046/neuroed.20180501.33

Lichtman, K. (2016). Age and learning environment: Are children implicit second language learners? *Journal of Child Language*, 43(3), 707–730. https://doi.org/10.1017/S0305000915000598

Loewen, S. (2009). Grammaticality judgment tests and the measurement of implicit and explicit L2 knowledge. In R. Ellis, S. Loewen, C. Elder, H. Reinders, R. Erlam, & J. Philip (Eds.), *Implicit and Explicit knowledge in second language learning, testing and teaching* (pp. 94–112). Multilingual Matters.

Maquet, P. (2001). The role of sleep in learning and memory. *Science*, 294(5544), 1048–1052. https://doi.org/10.1126/science.1062856

Marinis, T. (2010). Using on-line processing methods in language acquisition research. In E. Blom & S. Unsworth (Eds.), *Experimental methods in language acquisition research* (pp. 139–162). John Benjamins.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 38(4), 843–847.

Marsden, E., Morgan-Short, K., Thompson, S., & Abugaber, D. (2018). Replication in second language research: Narrative and systematic reviews and recommendations for the field. *Language Learning*, 1–71. https://doi.org/10.1111/lang.12286

Martínez-Flor, A., & Fukuya, Y. J. (2005). The effects of instruction on learners' production of appropriate and accurate suggestions. *System*, 33(3), 463–480. https://doi.org/10.1016/j.system.2005.06.007

Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research*, 80(3), 300–335. https://doi.org/10.3102/0034654310377087

Mirković, J., & Gaskell, M. G. (2016). Does sleep improve your grammar? Preferential consolidation of arbitrary components of new linguistic knowledge. *PLOS ONE*, 11(4), e0152489. https://doi.org/10.1371/journal.pone.0152489

Monaghan, P., Mattaock, K., Davies, R. A. I., & Smith, A. C. (2015). Gavagai is as Gavagai does: Learning nouns and verbs from cross-situational statistics. *Cognitive Science*, 39(5), 1099–1112. https://doi.org/10.1111/cogs.12186

Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M.T. (2012). Explicit and Implicit Second Language Training Differentially Affect the Achievement of Native-like Brain Activation Patterns. *Journal of Cognitive Neuroscience*, 24 (4): 933–947. https://doi.org/10.1162/jocn_a_00119

Muranoi, H. (2000). Focus on form through interaction enhancement: Integration formal instruction into a communicative task in EFL classrooms. *Language Learning*, 50(4), 617–673. https://doi.org/10.1111/0023-8333.00142

Newport, E. L. (2020). Children and adults as language learners: Rules, variation, and maturational change. *Topics in Cognitive Science*, 12 (1): 153–169.

Nieuwenhuis, I. L. C., Folia, V., Forkstam, C., Jensen, O., & Petersson, K. M. (2013). Sleep promotes the extraction of grammatical rules. *PLOS One*, 8(6), e65046. https://doi.org/10.1371/journal.pone.0065046

Norris, J.M., & Ortega, L. (2000). Effectiveness of L2 Instruction: A Research Synthesis and Quantitative Meta-analysis. *Language Learning*, 50(3), 417–528. https://doi.org/10.1111/0023-8333.00136

Pakuklak, E., & Neville, H. J. (2011). Maturational constraints on the recruitment of early processes for syntactic processing. *Journal of Cognitive Neuroscience*, 23(10), 2752–2765. https://doi.org/10.1162/jocn.2010.21586

Pan, B. A. (2012). Assessing vocabulary skills. In E. Hoff (Ed.), *Research methods in child language: A practical guide* (pp. 100–112). Wiley-Blackwell.

Paradis, M. (2004). *A neurolinguistics theory of bilingualism*. John Benjamins.

Paradis, M. (2009). *Declarative and procedural determinants of second languages*. John Benjamins.

Parnes, A. T.2011 *Ang marks the what? An analysis of noun phrase markers in Cebuano*. M.A. Thesis, Yale University.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Science*, 10(5), 233–238. https://doi.org/10.1016/j.tics.2006.03.006

Perruchet, P., & Vinter, A. (1998). PARSER: A model for Word Segmentation. *Journal of Memory and Language*, 39(2), 246–263. https://doi.org/10.1006/jmla.1998.2576

Pinto, M., & Zuckerman, S. (2018). Coloring book: A new method for testing language comprehension. *Behavior Research Methods,* 51, 2609-2628. https://doi.org/10.3758/s13428-018-1114-8

Podesva, R. J., & Sharma, D. (2013). *Research methods in linguistics*. Cambridge University Press.

Powell, M. J. D. 2009. *The BOBYQA algorithm for bound constrained optimization without derivatives*. Technical Report, Department of Applied Mathematics and Theoretical Physics, University of Cambridge.

Psychology Software Tools, Inc. (2012). *E-Prime 3.0.* http://www.pstnet.com/ .

R Core Team. (2015). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6(6), 317–327. https://doi.org/10.1016/S0022-5371(67)80149-X

Rebuschat, P., & Williams, J. N. (2012). Introduction: Statistical learning and language acquisition. In P. Rebuschat & J. N. Williams (Eds.), *Statistical learning and language acquisition* (pp. 1–12). Mouton de Gruyter.

Rebuschat, P. (2013). Measuring implicit and explicit knowledge in second language research. *Language Learning*, 63(3), 595–626. https://doi.org/10.1111/lang.12010

Robinson, P. (1997a). Generalizability and automaticity of second language learning under implicit, incidental, enhanced and instructed conditions. *Studies in Second Language Acquisition*, 19(2), 233–247. https://doi.org/10.1017/S0272263197002052

Robinson, P. (1997b). Individual differences and the fundamental similarity of implicit and explicit adult second language learning. *Language Learning*, 47(1), 45–99. https://doi.org/10.1111/0023-8333.21997002

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(6), 906–914. https://doi.org/10.1002/wcs.78

Saffran, J. R., Johnson, E. K., & Aslin, R. N. (1996). Word-segmentation: The role of distributional cues. *Journal of Memory and Language*, 35(4), 606–621. https://doi.org/10.1006/jmla.1996.0032

Sagarra, N., & Hanson, A. S. (2011). Eyetracking methodology: A user's guide for linguistic research. *Studies in Hispanic and Lusophone Linguistics*, 4(2), 1–13. https://doi.org/10.1515/shll-2011-1113

Saito, K. (2013). The acquisitional value of recasts in instructed second language speech learning: Teaching the perception and production of English /r/ to Adult Japanese learners. *Language Learning*, 63(3), 499–529. https://doi.org/10.1111/lang.12015

Saygin, Z. M., Norton, E. S., Osher, D. E., Beach, S. D., Cyr, A. B., Ozernov-Palchik, O., Yendiki, A., Fischl, B., Gaab, N., & Gabrieli, J. D. E. (2013). Tracking the roots of reading ability: White matter volume and integrity correlate with phonological awareness in prereading and early-reading Kindergarten children. *Journal of Neuroscience*, 33(33), 13251–13258. https://doi.org/10.1523/JNEUROSCI.4383-12.2013

Schmidt, C., & Miller, K. (2010). Using comprehension methods in language acquisition research. In E. Blom & S. Unsworth (Eds.), *Experimental methods in language acquisition research* (pp. 35–56). John Benjamins.

Schütze, C. T., & Sprouse, J. (2013). Judgment data. In R. J. Podesva & D. Sharma (Eds.), *Research methods in linguistics* (pp. 27–51). Cambridge University Press.

Scott, V. (1989). An empirical study of explicit and implicit teaching strategies in French. *The Modern Language Journal*, 72(1), 14–22. https://doi.org/10.1111/j.1540-4781.1989.tb05303.x

Sedivy, J. C. (2010). Using eyetracking in language acquisition research. In E. Blom & S. Unsworth (Eds.), *Experimental methods in language acquisition research* (pp. 115–138). John Benjamins.

Sheen, Y. (2007). The effects of corrective feedback, language aptitude, and learner attitudes on the acquisition of English articles. In Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 301–322). Oxford University Press.

Silverman, R. (2007). A comparison of three methods of vocabulary instruction during Read-Alouds in Kindergarten. *The Elementary School Journal*, 108(2), 97–113. https://doi.org/10.1086/525549

Singleton, D., & Munoz, C. (2011). Around and beyond the critical period hypothesis. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (Vol. 2, pp. 407–425). Routledge.

Singleton, D., & Ryan, L. (2004). *Language acquisition the age factor. Second edition.* Multilingual Matters Ltd.

Smith, L. B., Suanda, S. H., & Yu, C. (2014). The unrealized promise of infant statistical word-referent learning. *Trends in Cognitive Sciences*, 18(5), 251–258. https://doi.org/10.1016/j.tics.2014.02.007

Smith, L. B., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, 106(3), 1558–1568. https://doi.org/10.1016/j.cognition.2007.06.010

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning*, 60(2), 263–308. https://doi.org/10.1111/j.1467-9922.2010.00562.x

Spit, S., Andringa, S., Rispens, J., & Aboh, E. O. (2019). The opt out paradigm: First steps towards a new experimental method that measures meta-linguistic awareness. *Dutch Journal of Applied Linguistics*, 8(2), 206–227. https://doi.org/10.1075/dujal.17027.spi

Spit, S., Andringa, S., Rispens, J., & Aboh, E. O. (2020, November 18). Kindergarteners use cross-situational statistics to infer the meaning of grammatical elements. https://doi.org/10.31234/osf.io/kyvnz .

Spit, S., Andringa, S., Rispens, J., & Aboh, E. O. (2021). Do kindergarteners develop awareness of the statistical regularities they acquire? *Language Learning*, 71(2), 573–611. https://doi.org/10.1111/lang.12445

Spit, S., Andringa, S., Rispens, J., Aboh, E. O., & Vet, D. J. (2019, November 11). Meta-linguistic awareness in early (second) language acquisition. https://doi.org/10.17605/osf.io/bp5qe

Stickgold, R. (2005). Sleep-dependent memory consolidation. *Nature*, 437(7063), 1272–1278. https://doi.org/10.1038/nature04286
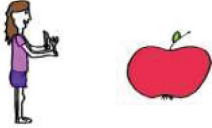
Thiessen, E. D., Kronstein, A. T., & Hufnagle, D. G. (2013). The extraction and integration framework: A two-process account of statistical learning. *Psychological Bulletin*, *139*(4), 792–814. https://doi.org/10.1037/a0030801

Turk-Browne, N. B., Jungé, J., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology*, *134*(4), 552–564. https://doi.org/10.1037/0096-3445.134.4.552

Ullman, M. (2001). A neurocognitive perspective on language: The declarative/procedural model. *Neuroscience*, *2*(10), 717–726. https://doi.org/10.1038/35094573

Ullman, M. (2016). The declarative/procedural model: A neurobiological model of language learning, knowledge and use. In G. Hickok & S. A. Small (Eds.), *The neurobiology of language* (pp. 953–968). Elsevier.

Vaahtoranta, E., Suggate, S., Jachmann, C., Lenhart, J., & Lenhard, W. (2018). Can explaining less be more? Enhancing vocabulary through explicit versus elaborative storytelling. *First Language*, *38*(2), 198–217. https://doi.org/10.1177/0142723717737452

Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition*, *127*(3), 375–382. https://doi.org/10.1016/j.cognition.2013.02.015

Vouloumanos, A. (2008). Fine-grained sensitivity to statistical information in adult word learning. *Cognition*, *107*(2), 274–729. https://doi.org/10.1016/j.cognition.2007.08.007

Williams, J., & Evans, J. (1998). What kind of focus and on which forms? In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 139–155). Cambridge University Press.

Williams, J. N. (2005). Learning without awareness. *Studies in Second Language Acquisition*, *27*(2), 269–304. https://doi.org/10.1017/S0272263105050138

Williams, J. N. (2009). Implicit learning in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *The new handbook of second language acquisition* (pp. 319–353). Emerald Press.

Williams, S. E., & Horst, J. S. (2014). Goodnight book: Sleep consolidation improves word learning via storybooks. *Frontiers in Psychology*, *5*, 184. https://doi.org/10.3389/fpsyg.2014.00184

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, *18*(5), 414–420. https://doi.org/10.1111/j.1467-9280.2007.01915.x

Zwart, F. S., Vissers, C. T. W. M., Kessels, R. P. C., & Maes, J. H. R. (2017). Procedural learning across the lifespan: A systematic review with implications for atypical development. *Journal of Neuropsychology 13*(2), 149-182. https://doi.org/10.1111/jnp.12139

Appendices

Appendix A: All words from the miniature language and their translations.

Word type	Word	Translation
Proper name	*Carlo*	'Carlo'
	Julia	'Julia'
	Marco	'Marco'
	Maria	'Maria'
Verbs	*Estima*	'Looking'
	Pentura	'Making a picture'
	Rigarda	'Eating'
Frequent nouns	*Domo*	'Tree'
	Herbo	'Banana'
	Kego	'Horse'
	Lito	'Flower'
	Pano	'Cat'
	Zambo	'Apple'
Infrequent nouns	*Ando*	'Carrot'
	Anso	'Dress'
	Arbo	'Castle'
	Bovo	'Strawberry'
	Halto	'Rabbit'
	Kobro	'Sheep'
	Misto	'Dog'
	Nego	'Egg'
	Nutro	'Painting'
	Teko	'Sandwich'
	Wiro	'Car'
	Wolgo	'Cow'
Grammatical marker	*Pli*	'Plural'
	Tra	'Any number'
Conjunction	*Ut*	'And'

Appendix B: Different types of sentence in the input and how often they occurred per block and with each noun type.

Structure	Total occurrences	Occurence per noun type	Picture
Tra + **single referent**	36 times 12 times per block	4 times per frequent noun 1 time per infrequent noun	
Tra + **multiple referents**	36 times 12 times per block	4 times per frequent noun 1 time per infrequent noun	
Pli + **multiple referents**	36 times 12 times per block	4 times per frequent noun 1 time per infrequent noun	
Filler	12 times 4 times per block	Did not contain a noun	