UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Do Kindergarteners Develop Awareness of the Statistical Regularities They Acquire?

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

DOI
10.1111/lang.12445

Publication date
2021

Document Version
Final published version

Published in
Language Learning

License
CC BY

Link to publication

Citation for published version (APA):
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

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:28 Oct 2021

LANGUAGE LEARNING
A Journal of Research in Language Studies

Language Learning ISSN 0023-8333

EMPIRICAL STUDY

Do Kindergarteners Develop Awareness of the Statistical Regularities They Acquire?

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

Amsterdam Center for Language and Communication, University of Amsterdam

Abstract: Many studies suggest that detecting statistical regularities in linguistic input plays a key role in language acquisition. Although statistical learning is not necessarily implicit in nature, it is often defined as learning that happens without awareness. This article investigates whether statistical learning in young children is indeed implicit, as often assumed. We trained 63 kindergarteners on a miniature language and assessed learning using a picture-matching task. We used an opt-out task to measure whether the kindergarteners possessed awareness of an acquired meaningful grammatical marker. In the opt-out paradigm, participants demonstrate awareness by expressing uncertainty through a nonverbal response: opting out. Our results are compatible with an earlier study of which the present study is a partial replication, suggesting that kindergarteners can acquire this marker from distributional properties in the input. Furthermore, although none of the children could verbalize knowledge of the structure during exit interviews, their behavior during the opt-out task indicated that they developed awareness of it.

This study is part of the PhD project "Meta-linguistic Awareness in Early (Second) Language Acquisition," which is funded by the University of Amsterdam. The authors declare that they have no conflict of interest. The authors would like to thank Dirk-Jan Vet, Tiffany Boersma, Channa van Dijk, Afra Klarenbeek, Klaas Seinhorst, Joris Wolterbeek, and Joanna den Blijker for their help with constructing the experiment, three anonymous reviewers for their helpful comments, and the participating schools for their hospitality. The long time between acceptance and publication is due to issues with making the earlier study that informed the current study available.

Correspondence concerning this article should be addressed to Sybren Spit, University of Amsterdam, Spuistraat 134, 1012 VB Amsterdam, the Netherlands. E-mail: S.B.Spit@uva.nl

The handling editor for this article was Emma Marsden.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Language Learning* published by Wiley Periodicals LLC on behalf of Language Learning Research Club, University of Michigan
DOI: 10.1111/lang.12445

Keywords artificial language learning; kindergarteners; metalinguistic awareness; replication; statistical learning; young learners

Introduction

When learners acquire a language, they have to accomplish several seemingly distinct tasks. Learners have to detect words within a continuous stream of speech; they need to map meanings onto these words; and they have to group these words into abstract categories and determine the grammatical relations between these categories. Accumulating evidence suggests that these tasks might be accomplished through the learners' ability to detect statistical regularities within linguistic input (e.g., Erickson & Thiessen, 2015; Frost & Monaghan, 2016; Romberg & Saffran, 2010; Walker, Monaghan, Schoetensack, & Rebuschat, 2020). Evidence that such statistical learning plays a role in different domains of language comes from several types of task results. Artificial grammar learning experiments have shown that tracking statistical regularities allows learners to infer word boundaries and establish dependencies between linguistic elements (e.g., Aslin, Saffran, & Newport, 1998; Gómez & Gerken, 1999; Saffran, Johnson, & Aslin, 1996, for infants; Gómez, 2002, for infants and adults; Endress & Bonatti, 2007; Perruchet & Vinter, 1998; Thiessen, Kronstein, & Hufnagle, 2013, for adults). Studies using cross-situational word-referent learning tasks have indicated that statistical learning plays a role in mapping form onto meaning (e.g., Frank, Goodman, & Tenenbaum, 2009; Frank, Tenenbaum, & Fernald, 2013; Smith & Yu, 2008; Smith, Suanda, & Yu, 2014; Vlach & Johnson, 2013, for infants; Kachergis, Yu, & Shiffrin, 2014; Vouloumanos, 2008; Yu & Smith, 2007, for adults). Several studies have also shown that distributional properties in the input can help learners to determine the semantic category to which words belong (e.g., Lany, 2014; Lany & Saffran, 2013, for infants; Chen, Gershkoff-Stowe, Wu, Cheung, & Yu, 2017; Monaghan, Mattock, Davies, & Smith, 2015, for adults).

Although statistical learning is not necessarily implicit in nature (Batterink, Reber, Neville, & Paller, 2015; Bertels, Boursain, Destrebecqz, & Gaillard, 2015), many scholars have defined it as learning that happens without awareness (e.g., Arciuli & Simpson, 2012; Aslin et al., 1998; Baker, Olson, & Behrmann, 2004; Kidd, 2011; Reber, 1967; Rebuschat, 2013). Rebuschat and Williams (2012) note that, from time to time, statistical learning is explicitly equated with implicit learning (Conway & Christiansen, 2006; Perruchet & Pacton, 2006; Turk-Browne, Jungé, & Scholl, 2005), and that the two are sometimes even concatenated, in speaking of "implicit statistical learning" (e.g., Conway, Bauernschmidt, Huang, & Pisoni, 2010; Kidd, 2011). Because statistical learning appears to play an important role in child language acquisition (e.g., Erickson & Thiessen, 2015; Frost & Monaghan, 2016; Romberg & Saffran, 2010), it is often assumed that children acquire a target language without any awareness of the regularities in it. This might be a reasonable assumption, because children are probably unable to verbalize or explicitly reflect on their knowledge. Indeed, regardless of the theoretical framework they adopt, many scholars assume children acquire linguistic patterns without any awareness of them (Aslin et al., 1998; Chomsky, 1986; Radford, 2004; Saffran et al., 1996; Tomasello, 2003; Ullman, 2016; Wijnen, 2013). A study by Spit, Andringa, Rispens, and Aboh (2020) seemingly confirmed this idea in a statistical learning experiment, in which kindergarteners had to acquire a meaningful grammatical element without explicit instruction. In the exit interview after the test phase of this miniature language learning experiment, no child could verbalize knowledge of the grammatical element they had learned, which could indicate that they had acquired it implicitly. It remains to be seen, however, whether this lack of verbalization necessarily reflects an absence of awareness of the acquired knowledge in these children. Instead, we suggest that they might be aware of the knowledge they acquire in a way they cannot verbalize, and that research so far has not tapped into awareness at this level.

Background Literature

Awareness and Language Acquisition

Awareness is studied within many disciplines of cognitive science, such as visual perception (e.g., Lamme, 2010), artificial intelligence (e.g., Dennett, 1993), and animal cognition (e.g., Proust, 2013). Informally, awareness can be described as the mental state in which living creatures know what they are experiencing. This conceptual representation of awareness stems from the classical Cartesian view on awareness (Descartes, 1637/1968), which assumed that humans are aware of experiences if they can remember them coherently and can verbalize them. This view has been labeled *the Cartesian theater* (Dennett, 1993). In this fictive theater, living beings are spectators who are only aware of their experiences once they appear on the stage. If an experience does not appear on stage, the spectator is not aware of it at all. Hence, awareness in this classical view is an all-or-nothing phenomenon; someone is either aware of something or not.

However, over the years, scholars have proposed alternative models that question the importance of verbalization, coherence, and memory as necessary requirements for awareness (Allport, 1988; Dehaene, Lau, & Kouider, 2017; Dennett, 1993; James, 1890). This work suggests that awareness probably exists in different gradations and is not the dichotomy that is represented in the Cartesian theater (Cleeremans, 2008, 2011, 2014). Cleeremans (2011), for example, distinguished between awareness at a phenomenal level (P-awareness) and awareness at the level of access (A-awareness). People possess P-awareness when they have a subjective experience of the object of awareness. Applied to linguistics, this means that they are phenomenally aware when they are able to explicitly describe certain linguistic patterns: This is metalinguistic P-awareness, which bears resemblance to what has often been called *explicit knowledge* in the field of second language acquisition (SLA; Hulstijn, 2005; Rebuschat, 2013). People possess A-awareness when they have an experience that is accessible to their cognitive system but cannot be expressed subjectively. People have metalinguistic awareness at access level if they perceive that a particular linguistic pattern is present and can (verbally) detect that "something" is going on, without being able to determine which pattern this is exactly. This A-awareness resembles what has been called *noticing* in the second language literature (Schmidt, 1990) and could be described as a mental state that floats somewhere between explicit and implicit knowledge.

Within linguistics, scholars have studied awareness most extensively within SLA, often focusing on the distinction between implicit and explicit knowledge (e.g., Andringa & Rebuschat, 2015; Hulstijn, 2015; Suzuki & DeKeyser, 2017). Explicit knowledge can be described as knowledge that is available to awareness, whereas implicit knowledge is not available to our awareness (Rebuschat, 2013). The research and findings on this topic are diverse (e.g., DeKeyser, 2003; Ellis, 2015; Han & Finneran, 2014; Paradis, 2009) and have led to the formulation of several theories about the role that awareness could play when people acquire the grammatical structures of a new language. Some authors claim that awareness of such structures in a second language is necessary for their acquisition (DeKeyser, 2003; Schmidt, 1990), whereas others suggest that awareness merely facilitates the acquisition process (Ellis, 2003). Yet other researchers have argued that awareness is only circumstantial and a consequence of how the language is learned, and that it has no causal influence on the implicit linguistic knowledge that is gained (Krashen, 1981).

The role of awareness in early first language acquisition has been studied too, but perhaps not as extensively as in the field of SLA. Many studies have investigated the role of awareness and explicit knowledge in the acquisition of vocabulary knowledge (e.g., Coyne, McCoach, Loftus, Zipoli, & Kapp, 2009; Silverman, 2007; Vaahtoranta, Suggate, Jachmann, Lenhart, & Lenhard, 2018), with a meta-analysis showing that vocabulary interventions in kindergarteners are effective and that more explicit instruction conditions result in bigger effect sizes (Marulis & Neuman, 2010). Furthermore, phonological awareness has been widely studied in young children (e.g., Degé & Schwarzer, 2011; Gillon, 2018), often with a focus on its important relation to reading (e.g., Furnes & Samuelsson, 2011; Johnson & Goswami, 2010; Saygin et al., 2013). Nevertheless, the general theoretical debate about the role of awareness during language learning that is present in the field of SLA seems to be absent in the field of child language acquisition in general, and in relation to statistical learning in particular.

We think there are good reasons to pursue the debate about the role of awareness in the fields of child language acquisition and statistical learning as well. Although the role of awareness is often not addressed explicitly in statistical learning studies, results from a study by Kerz, Wiechmann, and Riedel (2017) suggest that this topic requires more thorough investigation. In an artificial language experiment, they studied the acquisition of a grammatical marker and found that adults who did not possess any awareness of what they were learning scored slightly above chance (53%−56%) in a two-alternative forced-choice task measuring their acquired knowledge, but that they almost reached ceiling scores (88%−91%) when they did develop such awareness. These results suggest that awareness plays a pivotal role in the SLA process in adults. Whether awareness plays a similar role when young children acquire grammatical elements is an open question that merits investigation. Children might develop some degree of awareness of the regularities they acquire, and this awareness might facilitate their acquisition process. They could also develop awareness without really using it. Alternatively, awareness could be a prerequisite for them to acquire a language in the first place.

These questions are currently hardly asked, because the common assumption seems to be that when children learn from statistical regularities, this learning is implicit (e.g., Arciuli & Simpson, 2012; Aslin et al., 1998; Baker et al., 2004; Conway et al., 2010; Erickson & Thiessen, 2015; Frost & Monaghan, 2016; Kidd, 2011; Perruchet & Pacton, 2006; Reber, 1967; Rebuschat, 2013; Romberg & Saffran, 2010; Wijnen, 2013). The assumption that children do not develop awareness when learning from statistical regularities seems to be held mainly because they are unable to verbalize their awareness and seemingly do not possess metalinguistic P-awareness. Yet, in light of the graded nature of awareness (Cleeremans, 2008, 2011, 2014; Dehaene et al., 2017; Dennett, 1993), it would seem possible that children develop metalinguistic awareness at another level. To our knowledge, there is no research investigating whether this is the case. To gain insight into this issue, we would need to tap into metalinguistic A-awareness.

Measuring Awareness

Research on adults shows that measuring awareness is a difficult enterprise, and several measures of awareness can be distinguished (see Rebuschat, 2013, for a review). Awareness can be measured using retrospective verbal reports, in which researchers ask learners to verbally reflect on what they are learning (e.g., Curcic, Andringa, & Kuiken, 2019; Dienes, Broadbent, & Berry, 1991; Rebuschat & Williams, 2012; Williams, 2005). If learners can verbalize their linguistic knowledge in such reports, they are supposedly aware of it. Alternatively, researchers can use subjective measures, such as confidence judgment tasks, where learners have to indicate how confident they are of their task responses (e.g., Dienes, Altmann, Kwan, & Goode, 1995; Hamrick & Rebuschat, 2012) or are asked about the source of their knowledge (e.g., guess, intuition, memory, rule; see Rebuschat, Hamrick, Riestenberg, Sachs, & Ziegler, 2015). If learners are confident of their task behavior, this is taken as a sign that they are aware of the knowledge underlying this behavior. Another possible way to tap into awareness is the contrastive use of direct and indirect tests (e.g., Jiménez, Méndez, & Cleeremans, 1996; Reed & Johnson, 1994). Direct tests explicitly ask learners to rely on knowledge of which they are aware when executing a task, whereas indirect tests do not. (Indirect tests are described in more detail later.) If learners show they possess knowledge on either one of the two tasks (direct or indirect), but not on the other, this contrastive behavior can be used to make inferences about the status of their knowledge. Timmermans and Cleeremans (2015) label these types of measures *direct subjective methods*, because they directly assess the awareness that a person has, by relying on their subjective experience. Thus, these measures of awareness mainly tap into P-awareness: the type of awareness that people have when they possess a coherent experience of their object of awareness and can verbally reflect on it.

However, direct subjective methods have limitations (Rebuschat, 2013). In making use of retrospective verbal reports in which participants verbally reflect on whether or not they were aware of the gained knowledge, there is a tendency to overlook the fact that awareness can also exist in a way that cannot be verbalized (Dienes & Berry, 1997; Dienes & Fahey, 1995). Subjective measures are prone to a response bias, as participants set their own criteria for responding (Dienes, 2004, 2008; Reingold & Toth, 1996). Furthermore, participants' behavior on direct tests might be confounded by unconscious knowledge (Merikle, Smilek, & Eastwood, 2001; Reingold & Merikle, 1988, 1990). Because of these limitations and because we cannot expect children to verbalize rules or partake in confidence judgment tasks, given that they probably do not possess P-awareness, it seems unreasonable to employ these methods to test awareness in children. Other lines of research into awareness in young children, such as awareness of language itself at a macro-level (e.g., Atagi & Sandhofer, 2020) or metalinguistic awareness about written language (Ke, Miller, Zhang, & Koda, 2020) are also limited in that they do not give us methods for tapping into awareness of grammar during exposure to a new spoken language. Several tasks, such as grammaticality judgment tasks and wug tests, have been argued to tap into metalinguistic awareness and so might overcome these difficulties (Barac & Bialystok, 2012; Bialystok, 1986). Nevertheless, although such methods potentially require participants to possess metalinguistic A-awareness, we do not necessarily know whether participants are aware of *what* the regularities are that they have acquired, only that they are aware that there *are* regularities. Behavior on such tasks could equally well be the result of completely implicit knowledge of the rules. Therefore, if children possess awareness at the A-level, the question is what type of measurement we need to establish this, when we cannot use direct subjective methods, grammaticality judgment tasks, or wug tests.

One possibility for establishing whether awareness of the target regularities is involved when young learners acquire a language would be the use of indirect objective methods. These are methods that do not require subjective introspection by the participants (Timmermans & Cleeremans, 2015). An example of an indirect objective method that can be used to measure A-awareness is the opt-out paradigm. In this paradigm, participants can demonstrate awareness by exercising the option to express uncertainty through a nonverbal response: the so-called uncertainty response. Typically, participants in such a task perform a cognitive task with easy and difficult trials. If they possess awareness of their cognitive capacities, they should express their uncertainty in the more difficult trials by choosing the uncertainty response. In this paradigm, participants obtain a reward for a correct decision in a trial and are able to express their uncertainty by opting out of a trial. In the latter case, participants will always receive a reward, but this reward will be smaller than the one they get for a correct decision. If participants are aware of their cognitive capacities, they should take the risk to obtain a bigger reward in easy trials and opt out to obtain a smaller, but certain, reward in the more difficult trials.

Researchers have used this paradigm to investigate awareness in a variety of nonhuman animals (de Waal, 2016; Hampton, 2009; Smith, 2009). Experiments have shown that rhesus monkeys are aware of their memory of particular stimuli (e.g., Hampton, 2001), as are macaques (e.g., Fujita, 2009) and orangutans (e.g., Suda-King, 2008). Furthermore, rats are able to judge how well they distinguish long- from short-lasting noises (e.g., Foote & Crystal 2007); pigeons can indicate how good they are at classifying sparse and dense pictures (e.g., Sole, Shettleworth, & Bennett, 2003); and dolphins can assess their own performance when discriminating between high- and low-pitched sounds (e.g., Smith, 2010). Importantly, none of these animals had to engage in subjective introspection with regard to their cognitive capacity; they instead performed objective observable behavior that was an indirect effect of their awareness. Thus, the opt-out paradigm seems a useful indirect, objective measure to tap into A-awareness in children.

The Current Study

In this study, we exploited the opt-out paradigm to measure metalinguistic awareness in kindergarteners. This work extended two prior studies. The first study (Spit, Andringa, Rispens, & Aboh, 2019) investigated the balance in rewards for correct decisions and opt-out responses, to avoid a situation where the uncertainty response would be too appealing and would bias participants to pick this certain reward regardless of their awareness. Several studies have shown that children are much more likely to pursue a certain reward than an uncertain reward, because they cannot delay their gratification in an adultlike manner (Eigsti et al., 2006; Mischel, Shoda, & Rodriguez, 1989). Spit et al. (2019) showed that the rewards used in the present study strike the right balance: The children who participated in that study ($n = 48$, $M = 5;3$ years, $SD = 0;9$) did not show a bias toward the uncertainty response.

In the second study, an artificial language learning experiment was created in which the reward system could be incorporated (Spit et al., 2020). In this artificial language, kindergarteners ($n = 50$, $M = 5;5$ years, $SD = 0;10$) learned a grammatical marker, which could serve as a target structure that children could potentially develop awareness of. This grammatical marker expressed a plurality feature when it was combined with a noun, and its function could be learned from distributional properties in the input. Children seemed able to acquire this marker using statistical regularities, and posttest verbal reports suggested that the children were not aware at the phenomenal level (P-awareness) of their acquired knowledge.

The present study combined the artificial language learning task and the opt-out procedure. The main goal of this study was to see whether children develop A-awareness when they acquire a grammatical pattern on the basis of statistical regularities in the input. To test this, we used a novel measure in which children demonstrate such awareness by expressing uncertainty through a nonverbal response. If children develop awareness of their acquired knowledge, they should opt out more often on trials that are difficult and opt in on trials that are easy, when they have acquired this marker. In the spirit of recent calls for more replication (Marsden, Morgan-Short, Thompson, & Abugaber, 2018), because a single study can never be conclusive and chance or confounded findings might be published (Cumming, 2014), a secondary goal of our study was to replicate the results from Spit et al. (2020) that young learners are indeed able to acquire a meaningful grammatical marker from distributional properties in the input. Replicating our earlier findings would provide more evidence for the idea that children can use distributional cues to learn grammatical regularities (e.g., Lany, 2014; Lany & Saffran, 2013).

Method

Participants

Seventy native Dutch-speaking children (35 males, 35 females, $M = 5;5$ years, $SD = 0;6$, range: 4;3−6;10) took part in this experiment. All children were in kindergarten and were recruited from three primary schools in the central and western areas of the Netherlands. They did not have any diagnosed language or communication disorders. We did not collect any further information about the children's linguistic background. We had to exclude seven participants, because they consistently provided answers before they were exposed to the test items, probably due to a lack of concentration. For every test item, these children pointed toward the part of the screen they thought was showing the correct picture, before the pictures were on the screen. As a result, we present data from 63 children (32 males, 31 females, $M = 5;6$ years, $SD = 0;5$, range: 4;3−6;10).

Materials

Miniature Language and Target Structure

We used the artificial language from Spit et al. (2020), which consists of four proper names, three verbs, two grammatical markers, six frequent nouns, twelve infrequent nouns, and one conjunction. For all words and their translations, see Table 1. Apart from the proper names, which might occur in Dutch, all words were novel words. The language has subject–verb–object word order. In this language, a noun phrase on its own does not encode number and could correspond to both singular and plural referents. In a sentence, however, an argument noun phrase must be introduced by a nominal marker: *pli* or *tra*. This type of nominal marker included in our artificial language occurs in a number of languages that have residual noun classes, such as the Kwa language Akan (Appah, 2003) or the Austronesian language Cebuano (Parnes, 2011).

Table 1 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 nominal follows"
	Tra	"Any number nominal can follow"
Conjunction	*Ut*	"And"

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. The language includes another nominal grammatical marker, *pli*, that could be used instead of *tra*, and which encodes number and indicates that the noun necessarily refers to multiple referents. A similar type of marker appears in the Akan 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 that participants had to learn was that whenever *pli* preceded a noun, and not *tra*, this noun always referred to multiple referents.

Format of the Experiment

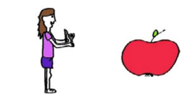
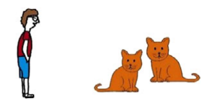
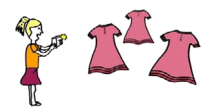
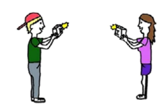
Before describing in the next section the precise characteristics of the input provided for learning this grammatical regularity, we will briefly lay out the overall structure of the experiment. So that participants had the opportunity to learn the regularity, they listened to an audio-recorded story with four protagonists (Carlo, Julia, Marco, and Maria), who were going on a holiday to a country whose language they did not speak. 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 four parts. It started out with a short vocabulary-training session. After the vocabulary training, a rule-training session followed. During this session, children received input that would allow them to acquire the grammatical rule. After the rule training, participants performed a picture-matching task designed to test whether they had acquired the rule. After the picture-matching task, participants continued with the opt-out phase, designed to test whether they were aware of the rule.

As in Spit et al. (2020), we were interested in the acquisition of the grammatical element, and not in that of lexical vocabulary items. That was why we included the initial short vocabulary-training session: We aimed to bootstrap participants slightly into learning the lexical items by means of training on the six frequent nouns of the language only. In this training, the nouns were presented auditorily in random order, 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 training: twice with one referent, twice with two referents, and twice with three referents, to make sure that a bare noun would not be associated with a particular number of referents. After this training, participants were given a short picture-matching task to test their vocabulary knowledge. They heard one of the frequent nouns and saw four pictures, one of which showed the matching referent and three of which showed some of the other frequent nouns. Our purpose in having participants perform this task was to familiarize them with the picture-matching task from an early point in the experiment and also to maintain their attention. Participants progressed to the rule-training phase regardless of their performance on the vocabulary test. In contrast, in our earlier study, participants continued this picture-matching task until they correctly identified each noun four times. This turned out to be problematic, because children were unable to learn the nouns from this short exposure. Therefore, we tested the knowledge of each noun in the current experiment only twice.

After the vocabulary training, participants were exposed to three rule-training phases in which they received input on the basis of which they could potentially 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, and the noun was the object. These sentences served as the input for learning 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 number marker and thus did not provide evidence of the grammatical rule to be learned. These sentences were used to create the attention task (see below) and can be regarded as fillers. Table 2 provides examples of all sentence types and their accompanying pictures. Sentences in the language were always semantically plausible. After rule-training phases one and two, participants completed a six-item vocabulary test (one item for each frequent noun), which was administered according to the same procedure as described earlier. The purpose of inserting these vocabulary tests was to maintain participants' attention. Participants would receive a sticker after each vocabulary test, regardless of their results.

Furthermore, because we wanted children to pay attention to both the visual and auditory input, all participants performed an attention task. They were made to believe that the four protagonists sometimes could not hear correctly what had been said in the new language. This was indicated by a questioning face of a protagonist after some of the stimuli. When participants saw this face, they had to repeat the previously heard stimulus. The purpose of inserting this attention task was to keep participants focused; we therefore did not analyze their answers for correctness. Participants were introduced to this attention task during the vocabulary training, where 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. By using sentences without a marker and noun as the fillers, we ensured that children who were better at repeating these sentences would not receive more input on the regularity through self-rehearsal (auto-input). Importantly, the current experiment deviates slightly from our previous experiment, in which only half of the

Table 2 Examples of artificial language training sentences and rough translations and how often they occurred per block and with each noun type

Structure	Example	Total occurrences	Occurrence per noun type	Translation	Picture
Tra + single referent	*Maria rigarda tra zambo*	36 times 12 times per block	4 times per frequent noun 1 time per infrequent noun	“Maria eats (an) apple(s)”	
Tra + multiple referents	*Carlo estima tra pano*	36 times 12 times per block	4 times per frequent noun 1 time per infrequent noun	“Carlo looks at (a) cat(s)”	
Pli + multiple referents	*Julia pentura pli anso*	36 times 12 times per block	4 times per frequent noun 1 time per infrequent noun	“Julia takes a picture of dresses”	
Filler	*Marco ut Maria pentura*	12 times 4 times per block	Did not contain a noun	“Marco and Maria take a picture of each other”	

Note. For sentences with *tra*, the noun could be translated as either a singular or a plural. The correct interpretation should be inferred from the visual context.

participants performed the attention task (Spit et al., 2020). That study showed a slight but not significant advantage on the picture-matching task for learners who had done the attention task. However, we nevertheless decided to include the attention task in the current study because this meant that there was less need for the experimenter to draw the attention of the children back to the task when they were not looking at the visual input.

Characteristics of the Input

Of the 108 training sentences that participants could use to learn the rule during the rule-training phases, 72 sentences contained the default marker *tra*, whereas 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 sentences with *tra* that showed multiple referents. As a result, 36 pictures showed one referent, 36 pictures showed two referents, and 36 pictures showed three referents. Furthermore, the probability that the noun following *pli* had multiple referents was 1. For a noun following *tra*, this probability was 0.5, and the probability that it referred to a single referent was 0.5 as well. Vice versa, when a single referent was shown, the probability of hearing *tra* before the noun was also 1. If a participant saw multiple referents, the probability of hearing *tra* before the noun was 0.5, as was the probability of hearing *pli* before the noun.

Every frequent noun occurred 12 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 to 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. For an overview of the characteristics of the input, see Table 2.

Testing Sensitivity to the Rule

Immediately after the rule-training phase, participants took part in a picture-matching task so that we could determine whether they became sensitive to the grammatical cue. 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. We used infrequent rather than novel nouns so participants would feel that a target response could be based on what they had been exposed to. Young children reportedly have difficulties with tests in which they have to make decisions that seem unrealistic to them (for a methodological review, see Pinto & Zuckerman, 2019), which would be the case if test items contained novel nouns and pictures (i.e., entirely absent from the training phase). We tried to circumvent this by using infrequent rather than novel nouns in the picture-matching task. This meant that 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 that these children used their knowledge of the grammatical marker when giving an answer, as chances are slim that they learned the meaning of these nouns from only three occurrences in the input.

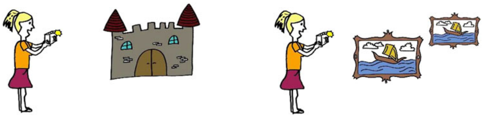
Of the 36 sentences used during the test phase, 24 were experimental items. For experimental items, participants had to choose between pictures with either one or multiple referents. The two pictures always referred to different referents (e.g., apple(s) in one picture and banana(s) in another). There were 12 experimental items containing *pli*. For these items, the target picture always showed multiple referents, and the alternative picture showed a single referent. Another 12 experimental items contained *tra*. Any number of referents would be grammatical for these items, so only the semantics of the noun could determine the correct response. For half of the experimental items containing *tra*, the target picture showed multiple referents and the alternative picture showed a single referent. For the other half of these items, the target picture showed a single referent and the alternative picture showed multiple referents. Every noun occurred once with *tra* and once with *pli* during the test. Although we did not expect children to have learned the meanings of the nouns in the training phase, if they did learn them then they could be expected to produce a target answer in both conditions. The performance on the sentences with *tra* can 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 (as well as the meaning of the noun, if they had learned it), whereas for trials with *tra* the number of referents was not indicative (and so only the meaning of the noun could serve as a cue).

In addition to the 24 test items, we included 12 filler items. Fillers contained *pli* or *tra*, but showed the same number of referents on both pictures. We

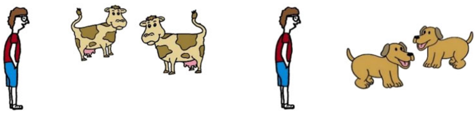
Table 3 Examples of test items and their rough translations

Example with *pli*: *Marco rigarda pli bovo* “Marco eats strawberries”

Example with *tra*: *Julia pentura tra nutro* “Julia takes a picture of (a) painting(s)”



Example of filler: *Carlo estima tra misto* “Carlo looks at (a) dog(s)”



Note. For sentences with *tra*, the noun could be translated as either a singular or a plural. The correct interpretation should be inferred from the visual context. In all three examples, the right-hand picture was the target.

included these fillers to avoid participants linking the grammatical markers to number during the test phase itself, because they always had to choose between a picture with multiple referents and a picture with a single possible referent. For examples of the different types of items, see Table 3. Items of the different types were presented in a counterbalanced semirandomized order.

Testing Awareness of the Rule

After the picture-matching task, there followed the opt-out phase, designed to determine whether participants were aware of the grammatical pattern they had acquired. Before this phase started, participants engaged in a short practice session. This practice session was necessary to prepare the children for the opt-out phase of the experiment and familiarize them with the rewards they could obtain in that phase. In this session, participants heard the six frequent nouns one by one from a recording. Each time they heard a noun, they saw two pictures and had to choose which of the two matched the noun they had heard. Importantly, for every correct decision a child would get a reward. Children helped the protagonists cross a river with 10 stepping-stones to return home after their holiday. The protagonist would move two stones for every correct answer. Children received a sticker for every protagonist that returned home in this practice session and in the opt-out phase.

In the opt-out phase, participants heard the same test items that were used in the picture-matching task, up to and including the grammatical marker, but without the last word, while seeing a blank screen. After hearing these incomplete sentences, participants were given two options. One option was that they could choose to hear the whole sentence including the last word. When hearing the full sentence, they would see the two pictures and they would have to make a choice between these pictures. As before, if they decided correctly they would earn two steps, and if they made a wrong decision, they would not obtain any reward. The other option was that participants could choose to opt out. In this case, they decided not to hear the full sentence, but simply to move on to the next test item. When the participant opted out, the protagonist would move on one stepping-stone. Thus, in order to bring home as many protagonists as possible, participants had to consider their chances of providing a correct answer. The assumption being that the more certain they were of their knowledge of the number of referents that would come after *pli* or after *tra*, the more likely they would be to not opt-out but, rather, to hear the whole sentence so as to have the opportunity to move two steps forward. If participants decided to hear the full sentence, the sentence would always end with one of the infrequent nouns that occurred in the training phase. This opt-out phase consisted of the same 24 test items from the picture-matching task, but with the nouns removed. That is, 12 items ended after *pli*, and 12 items ended after *tra*. Thus, if a participant chose to hear a full sentence every time they were given the option, that participant would hear all 24 test items from the picture-matching task.

After the opt-out phase, a 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 was during the picture-matching task, whether they knew the meaning of the words *pli* and *tra*, and why they decided to opt in or opt out. All participants were asked all questions.

Procedure

All stimuli were recorded by a female native speaker of Dutch. 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). During vocabulary training, nouns and their accompanying pictures would be presented for 3 seconds, before automatically moving to the next noun. Because the goals of this training were familiarization and maintaining 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. During both test phases, the experimenter pressed a button on the keyboard that corresponded to the answer the participant gave. Scores from this phase were registered by the software. 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, the picture-matching task 5 minutes, and the opt-out phase 5 minutes. The full experiment took approximately 30 minutes per participant. We obtained ethical approval for this study from the University of Amsterdam, and we obtained passive (opt-out) consent from children's parents or legal guardians before the start of the study. Full materials, data, and analyses codes can be found at osf.io/bp5qe (Spit, Andringa, Rispens, Aboh, & Vet, 2019) and on iris-database.org.

Analysis

We carried out all analyses in R (R Core Team, 2015), using the lme4 package (Bates, Maechler, Bolker, & Walker, 2015) where needed. All scripts and data are available on the OSF and on IRIS. To determine whether participants grasped the target regularity and whether they were aware of this regularity, we carried out a logistic regression model with mixed effects for the picture-matching task and opt-out phase separately. We expected that if children learned the regularity, they should score better on sentences with *pli* than on sentences with *tra* in the picture-matching task, and if they were aware of this regularity, they should opt out more often when sentences contained *tra* than when sentences contained *pli*. Both models took the responses from the tasks (in the picture-matching task, 1 for a correct answer, 0 for an incorrect answer; in the opt-out phase, 1 for opting out, 0 for not opting-out) as a dependent variable, marker type (*tra* or *pli*) as a within-participants fixed effect, participant as a between-participants random effect (as an intercept), and test item as a within-participants random effect (as an intercept). Our fixed effects were included in this first model, because we were a priori interested in their contribution to the outcome (Gelman & Hill, 2007). In both models, orthogonal sum-to-zero contrast coding was applied to our binary fixed effect (i.e., marker type; Baguley, 2012, pp. 590−621). Because we aimed to keep the models as fully specified as possible by including random intercepts for participants and test items (Barr, Levy, Scheepers, & Tily, 2013), we had to increase the number of possible iterations to 100,000 (Powell, 2009), to solve issues with nonconverging models. This enabled us to report on design-driven random effect structures that were still justified by our data (Jaeger, 2009).[1] We used an alpha level <0.05 for determining significance, and in line with Baguley's (2009) recommendation, we report simple (rather than standardized) effect sizes and their confidence intervals. We computed an $R^2_{marginal}$, which summarizes only the explanatory power of the fixed effects structure, and an $R^2_{conditional}$, which summarizes the combined explanatory power of the combined fixed and random effects structure, following Nakagawa, Johnson, and Schielzeth (2017). As an exploratory analysis, we conducted a Pearson correlation to investigate possible relationships between the scores on both parts of this experiment.

Replication

Before presenting our results, we state at which points the present study deviates from our previous study (Spit et al., 2020). Apart from the following three points, the present and previous experiment were exactly the same. For example, the same investigator executed both experiments, and all statistical procedures to analyze the data from the picture-matching task in the current experiment 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 our previous experiment, participants would continue to receive test items until they identified all six nouns correctly four times each. This proved frustrating, however, because most children had difficulties in learning all nouns from the limited input they received. Therefore, it was problematic to keep the vocabulary training as it was. The vocabulary test was executed 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 our previous study. In our previous experiment, we observed a tendency toward better performance on the picture-matching task by children in the attention task condition than by children who did not perform this attention task. Therefore, we decided to include the attention task for all children.
3. The opt-out phase included in this study, in which we assessed whether children developed awareness of the grammatical regularity, was added in this experiment. It came after the picture-matching task and therefore cannot have affected performance on that task. However, the opt-out phase took

place before the informal debriefing, and could have influenced the answers children gave during this debriefing.

Results

For the descriptive statistics from the picture-matching task and opt-out phase, see Table 4 and Figure 1. We first tested our expectation that children would be sensitive to the distributional properties in the input. If children were sensitive to these properties, they should have scored better on sentences with *pli* than on sentences with *tra* in the picture-matching task. Results from this task showed that participants gave more target answers when sentences contained *pli* than when sentences contained *tra*, $OR = 1.344$, 95% CI [0.980, 1.843], $z = 1.834$, $p = .067$, $R^2_{\text{marginal}} = .006$, $R^2_{\text{conditional}} = .054$, but this effect was not statistically significant. Based on this sample, we cannot claim that our kindergartener participants became sensitive to this regularity, although we see a tendency toward such an effect.

In the opt-out phase, we tested whether children developed A-awareness of the target structure. If they developed such awareness, they should have opted out more often when sentences contained *tra* than when sentences contained *pli*. Results from the opt-out phase showed that participants did indeed opt out more often when sentences contained *tra* than when sentences contained *pli*, $OR = 1.482$, 95% CI [1.089, 2.017], $z = 2.502$, $p = .012$, $R^2_{\text{marginal}} = .006$, $R^2_{\text{conditional}} = .0521$. This suggests that children were at some level aware of the regularity that *pli* always preceded plural referents. Note that the number of times participants opted out in general (i.e., regardless of the sentence type) varied greatly at the individual level (between children), which seems to indicate that, irrespective of whether children generally tended to favor rejecting the uncertainty response or not (i.e., whether they were risk-takers or not), a main effect was still observable that they were aware of the regularity. That is,

Table 4 Scores from the picture-matching task indicating the number of target answers produced and scores from the opt-out phase indicating the number of times participants opted out, for sentences in which the two different nominal markers occurred

	Picture-matching task				Opt-out phase			
Marker	*M*	*SD*	*SE*	Range	*M*	*SD*	*SE*	Range
Pli	7.05	1.90	0.24	3–12	6.95	3.68	0.46	0–12
Tra	6.21	1.82	0.23	2–10	7.71	3.37	0.42	0–12

Note. Scores could range from 0 to 12.

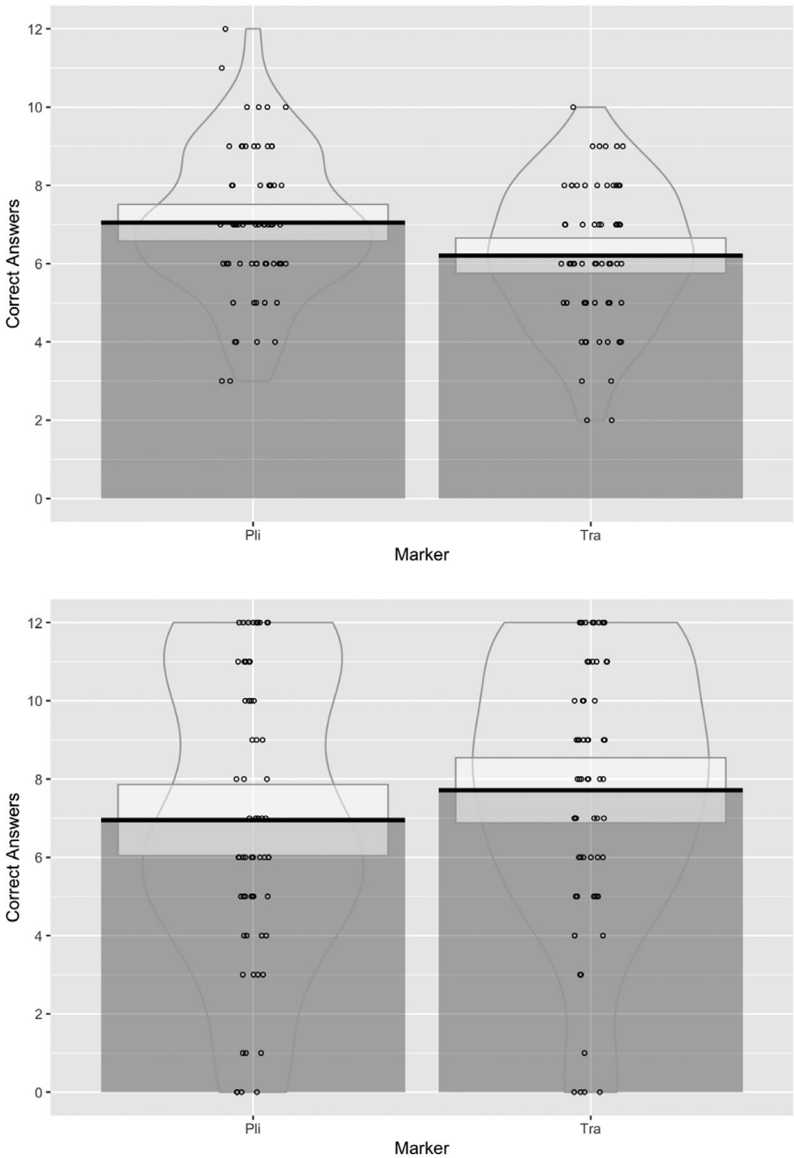


Figure 1 Graphs depicting the results from the picture-matching task (top panel) and the opt-out phase (bottom panel). Scores from the picture-matching task indicate the number of target answers produced, and scores from the opt-out phase indicate the number of times participants opted out. Scores could range from 0 to 12.

some children were likely to opt out in general whereas other children took a risk more often and hardly ever opted out so they had a chance at moving two steps, but, given that the type of marker was a within-participant factor, such personal preferences cannot have influenced the observed main effect.[2]

In the debriefing, when asked how they reached a decision on the picture-matching task, 75% of the participants ($n = 47$) reported that they did not know how they made a decision. The other 25% of them ($n = 16$) claimed that they had heard the sentences before and remembered what they meant. When asked for the meaning of *pli* and *tra*, 71% of the participants ($n = 45$) said that they did not know what these words meant; another 21% of them ($n = 13$) gave the meaning of a noun they had learned during the exposure; and, finally, 8% of the participants ($n = 5$) came up with some other meaning for these words. This suggests that children were not phenomenally aware of the regularity they acquired. When asked why they opted out or opted in, 22% of the children ($n = 14$) said that they had based their decision on the difficulty of the trials, but none of them linked this to plurality. The other 78% of the children ($n = 49$) did not report that they based their opt-out behavior on the difficulty of the trials, but, for example, because they wanted to hear the next sentence, or wanted to help the protagonists arrive back home.

To determine whether a correlation existed between the awareness-related behavior in the opt-out phase and the learning in the picture-matching task, we calculated a difference score for both phases. For the picture-matching task, this score was calculated by subtracting the number of target answers on sentences with *tra* from the number of target answers on sentences with *pli*. A larger difference in a positive direction indicates that the participant gave more target answers on sentences with *pli* than on sentences with *tra*. Difference scores per participant for the picture-matching task are shown in Figure 2. For the picture-matching task, 54% of the participants ($n = 34$) exhibited a positive difference score, 16% of the participants ($n = 10$) a neutral difference score, and 30% of the participants ($n = 19$) a negative difference score. For the opt-out phase, this score was calculated by subtracting the number of times a participant decided to continue when a sentence ended with *tra* from the number of times a participant decided to continue when a sentence ended with *pli*. If this difference score was higher in a positive direction, this indicates that the participant opted out more often when a sentence ended with *tra* than when a sentence ended with *pli*. Difference scores per participant for the opt-out phase are shown in Figure 2 as well. In this phase, 49% of the participants ($n = 31$) exhibited a positive difference score, 32% of the participants ($n = 20$) a neutral difference score, and 19% of the participants ($n = 12$) a

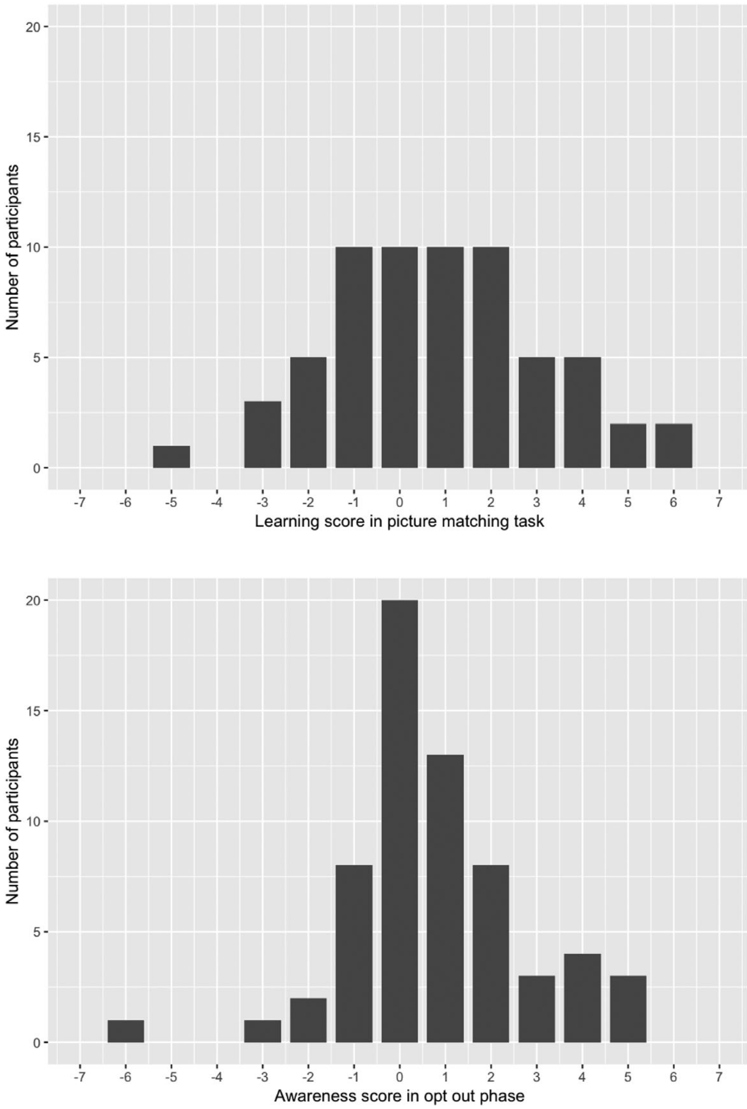


Figure 2 Graphs depicting the difference scores from the picture-matching task (top panel) and the opt-out phase (bottom panel) per participant. A positive score in the picture-matching task indicates that a participant gave more correct answers for sentences with *pli* than for sentences with *tra*, and thus indicates learning of the regularity. A positive score in the opt-out phase indicates that a participant opted out more often for sentences with *tra* than for sentences with *pli*, and thus indicates awareness of the regularity.

negative difference score. A Pearson's correlation test did not show a significant correlation between learning and awareness in this experiment ($r = -.133$, 95% CI [$-.369$, .118] $p = .23$).

Discussion

The main goal of this opt-out experiment was to see whether kindergarteners develop awareness of the linguistic regularities they acquire using distributional properties of the input. We wanted to investigate this because statistical learning is not necessarily implicit in nature, although many scholars seemingly assume it is (e.g., Arciuli & Simpson, 2012; Aslin et al., 1998; Baker et al., 2004; Conway et al., 2010; Erickson & Thiessen, 2015; Frost & Monaghan, 2016; Kidd, 2011; Perruchet & Pacton, 2006; Reber, 1967; Rebuschat, 2013; Romberg & Saffran, 2010; Wijnen, 2013). To our knowledge, there have not been systematic investigations into the implicit or explicit nature of statistical learning in children. As well as investigating whether young children develop awareness during learning, this study also provided an opportunity to replicate our findings from an earlier study (Spit et al., 2020), in which we demonstrated that a meaningful grammatical marker could be learned on the basis of statistical information.

In the present study, we showed that participants gave more target answers when sentences contained *pli* than when sentences contained *tra*, but that this effect was not statistically significant. This could suggest that we were unable to replicate our earlier finding that kindergarteners use distributional properties from the input to acquire a meaningful grammatical marker. However, replication is not necessarily a matter of significance, and findings might be replicated in several ways. Marsden et al. (2018) explain that researchers may carry out narrative comparisons, compare descriptive statistics (including effect sizes and CIs), or interpret dichotomous findings from null hypothesis significance testing. Reasonable arguments might be given for each type of comparison, and it is unclear which comparison is most appropriate. If we look beyond p values and include effect sizes and confidence intervals in the comparison of this study and the previous one, then the outcomes are not so different. In Spit et al. (2020), we reported a main effect of grammatical marker in the same direction and found evidence that children could acquire this regularity, $OR = 1.515$, 95% CI [1.071, 2.257], $z = 2.178$, $p = .029$. Perhaps it was more difficult, for some reason, to establish the effect in the current study to one particular degree of certainty (the arguably arbitrary cut-off of $p < .05$), because it is slightly smaller than the effect in our previous study. One reason for this could be that the vocabulary training at the start of the experiment did not continue until participants demonstrated they knew all six frequent nouns, but stopped at a fixed point in time. Perhaps slightly less stable vocabulary knowledge at the start of the rule training phase made it a bit more difficult to grasp the meaning of the grammatical markers. Nevertheless, our current effect size (*OR*) falls within the 95% CI of our earlier study and vice versa: $OR = 1.344$, 95% CI [0.980, 1.843], $z = 1.834$, $p = .067$. Given each study, the results of the other can be expected because they fall within the range of most probable results. Thus, this study can be seen as adding support to the view that kindergarteners use distributional properties of the input to acquire a meaningful grammatical marker. If these findings were to hold in new studies, this would indicate that statistical learning not only plays a role in word segmentation, word referent learning, and the acquisition of agreement markers (e.g., Frost & Monaghan, 2016; Romberg & Saffran, 2010), but is also involved in acquiring other types of grammatical structures in young learners (e.g., as found by Lany, 2014; Lany & Saffran, 2013). However, caution is still needed in interpreting these results, because the p values from both studies are "dancing" around .05.

During the exit interviews, none of the children could reflect verbally on the target structure and none could explain what either *pli* or *tra* meant by, for example, explaining these markers had something to do with number or that they meant something like *the* or *these* or *a* or *some*. This seems to be in line with the idea that young children do not develop awareness of the regularities in the target language (Aslin et al., 1998; Chomsky, 1986; Radford, 2004; Saffran et al., 1996; Tomasello, 2003; Ullman, 2016; Wijnen, 2013). However, awareness probably exists at multiple levels, and not only in ways that can be verbalized (Allport, 1988; Cleeremans, 2008, 2011, 2014; Dehaene et al., 2017; Dennett, 1993; James, 1890). Perhaps children are aware of the regularity but cannot verbalize this during an interview. To tap into awareness at another level, we included the opt-out procedure and found that children opted out more often when they heard a sentence containing *tra* than when they heard a sentence containing *pli*. This indicates that children were more inclined to make a decision on *pli* trials, for which they could be more certain. It also showed that they were less willing to try to obtain a reward for the *tra* trials, possibly because they were less certain about their ability to give a correct answer in these trials. This pattern of results is what you would expect for children aware—at some level—of the target structure. When the structure is acquired, trials with *pli* are predictable, and when a participant is aware of this regularity, it is expected that this person does not opt out in these cases but keeps playing. Sentences with *tra* are unpredictable, even when the regularity is acquired, because *tra* can refer to both singular and plural referents. In these cases, children who have some level of awareness are expected to opt out more often. We observed precisely this behavior, but because the effect is only small and borderline statistically significant (at alpha = 0.05), we are cautious about making strong claims on the basis of this result. It is necessary to replicate this finding in order to provide further support.

If our findings indeed indicate that children develop awareness of the knowledge they acquired through statistical learning, this would have some important implications. In particular, it would mean that statistical learning is not necessarily as implicit as is often assumed. This result could be a first step in conducting more research on the role of awareness in early language acquisition, of the kind that is already taking place in the field of SLA (e.g., Andringa & Rebuschat, 2015; Curcic et al., 2019; DeKeyser, 2003; Ellis, 2003; Hulstijn, 2015; Krashen, 1981; Schmidt, 1990). Perhaps young learners also need some form of awareness—at some point—of such structures in a language in order to acquire them (e.g., DeKeyser, 2003; Schmidt, 1990, 2001), or perhaps it only facilitates their acquisition process (e.g., Ellis, 2003). Of course, it could also have no influence at all on the implicit linguistic knowledge that is gained, or it could hinder it (e.g., Krashen, 1981). Any of these scenarios may be possible for children as well, and a possible correlation between participants' learning and awareness might have shed light on this issue, but our findings were inconclusive in this regard—the correlation was in fact negative, but very small, its CI passed through zero, and it was not statistically significant. The absence of a statistically significant correlation between learning in the first phase and opt-out behavior in the second phase may seem counterintuitive, but this is not necessarily so. An absent correlation would be consistent with views that learning and awareness are uncorrelated (e.g., Krashen, 1981); a null effect like this, however, cannot be interpreted as proof of an "absence" of such a correlation, and there may be other explanations for the finding of a nonstatistically significant correlation.

One explanation for the absence of a significant correlation between learning and awareness could be that the group we tested is nonergodic (Lowie & Verspoor, 2019). This means that not every individual is representative of the group and vice versa. Perhaps awareness plays some role in acquiring language in young learners, but the extent to which it does so is subject to vast individual differences (see Curcic et al., 2019, for evidence of such variation between second language learners). This idea may be supported by the variation in difference scores from the picture-matching task and the opt-out phase. Not every individual gave more correct answers for sentences containing *pli* than for sentences containing *tra* during the picture-matching task, and not every individual opted out more often for sentences ending with *tra* than for sentences ending with *pli* during the opt-out phase. The answers that participants gave during the informal debriefing also point toward individual differences in learning the markers. Some children did not know the meaning of them at all, whereas others thought they had a nominal meaning.

Limitations and Future Research

The linguistic background of the children might play a role in explaining the individual differences in our study. Several authors have suggested, for example, that bilingualism might influence the development of metalinguistic awareness (Atagi & Sandhofer, 2020; Barac & Bialystok, 2012; Bialystok, 1986), which might also be the case in our sample. For this study, we opted not to collect any data that did not pertain to our research questions. We were investigating only whether children are able to develop awareness, and not necessarily whether all children develop such awareness or whether particular children develop more of this awareness than others. Thus our sample was representative of the population we were interested in and served our specific goals (see Kruskal & Mosteller, 1979; Kukull & Ganguli, 2012, for further discussion of different types of representativeness), but it does not enable us to draw firm conclusions about behavior in either specific individuals or specific populations. As a result, we were unable to explore whether the individual variation we observed in the present sample could be related to individual characteristics, such as the child's linguistic background. Hypotheses about possible relationships between awareness and participant features may be investigated in studies dedicated to investigating individual differences.

Another reason we were unable to uncover a possible correlation between learning and awareness is that the effect sizes for the picture-matching task and opt-out phase were small. Such small effect sizes in statistical learning studies are not in fact uncommon (e.g., Kerz et al., 2017), and we may wonder whether increased awareness in the age group in the current study, as in adults, leads to enhanced learning scores, which could have impact even in educational settings such as kindergarten. Future research, for example, could manipulate awareness in learners to see whether this leads to improved learning outcomes and distinct developmental patterns. Other avenues to consider are the length of exposure and complexity of the grammatical marker: Would children develop more awareness when they receive more input to learn the marker, or when there is a clear one-to-one correspondence between each marker and its meaning (unlike in the current study, where *tra* had both plural and singular functions and *pli* overlapped with one of those functions)?

This study was the first, to our knowledge, to apply direct objective measures to investigate whether kindergarteners develop awareness of target structures during language learning. Such measures have great potential for examining awareness because they do not require participants to verbalize their awareness, and they potentially tap into different levels of awareness (e.g., as proposed by Timmermans & Cleeremans, 2015). It would be worthwhile to develop these measures further in order to investigate this issue more fully. We hope that this study will lead toward more research on the topic, such that clearer hypotheses can be formulated about the relation between awareness and language acquisition in young learners.

Conclusion

The goal of this article was twofold. First, we wanted to replicate a previous study, in which we showed that kindergarteners might use statistical properties of the input to acquire a meaningful grammatical marker. Our current results partially replicate these findings, as the current effect size is broadly in line with our previous findings, but we cannot reject the null hypothesis based on our current sample. We argue that these two studies together are compatible with the idea that kindergarteners can acquire such a grammatical marker on the basis of statistical learning processes in a relatively short time span. Second, we wanted to see how implicit the acquisition of such a marker actually is. Many scholars assume statistical learning is an implicit process that happens without any awareness involved. Our opt-out results seemingly nuance this idea. Although children were unable to verbalize their awareness of the grammatical regularities they had acquired, they seemed to develop awareness of what these regularities were at some other level. The exact role of this awareness in the acquisition process is yet to be identified. We hope that the findings presented here are a first step in this direction.

Open Research Badges

This article has earned Open Data and Open Materials badges for making publicly available the digitally-shareable data and the components of the research methods needed to reproduce the reported procedure and results. All data and materials that the authors have used and have the right to share are available at https://osf.io/bp5qe and https://www.iris-database.org/. All proprietary materials have been precisely identified in the manuscript.

Final revised version accepted 16 December 2019

Notes

1 The accepted version of this article reported a model with marker type as a random slope for participant. However, on rerunning the analysis with an updated version of the lme4 package, this model no longer converged. On removing that slope, the model converged and the results of that model are now reported here.

2 One reviewer suggested that children might have developed awareness during the opt-out task. We tested this by also running a model with item order as a fixed effect, instead of taking test item as a random effect, but did not find evidence for this hypothesis. In fact, descriptively, children gradually chose to opt out *more* toward the end of the test, perhaps due to fatigue or loss of attention.

References

Aboh, E. O. (2004). *The morphosyntax of complement-head sequences: Clause structures and word order patterns in Kwa*. Oxford: 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* (pp. 206–236). Amsterdam: John Benjamins.

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

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

Appah, C. K. I. (2003). *Nominal derivation in Akan: A descriptive analysis* (M.A. Thesis). Norway: 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*, 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*, 321–324. Retrieved from https://doi.org/10.1111/1467-9280.00063

Atagi, N., & Sandhofer, C. M. (2020). Early language environments predict aspects of explicit language awareness development. *Language Learning*, *70*, 464–505. https://doi.org/10.1111/lang.12381

Baguley, T. (2009). Standardized or simple effect size: What should be reported? *British Journal of Psychology*, *100*, 603–617. https://doi.org/10.1348/000712608.377117

Baguley, T. (2012). *Serious stats*. Basingstoke, UK: Palgrave Macmillan.

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

Barac, R., & Bialystok, E. (2012). Bilingual effects on cognitive and linguistic development: Role of language, cultural background and education. *Child Development*, *83*, 413–422. https://doi.org/10.1111/j.1467-8624.2011.01707.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*, 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*, 1–48. https://doi.org/10.18637/jss.v067.i01

Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language*, *83*, 62–78. https://doi.org/10.1016/j.jml.2015.04.004

Bertels, J., Boursain, E., Destrebecqz, A., & Gaillard, V. (2015). Visual statistical learning in children and young adults: How implicit? *Frontiers in Psychology*, *5*, 1–11. https://doi.org/10.3389/fpsyg.2014.01541

Bialystok, E. (1986). Factors in the growth of linguistic awareness. *Child Development*, *57*, 498–510. https://doi.org/10.2307/1130604

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*, 1485–1509. https://doi.org/10.1111/cogs.12417

Chomsky, N. (1986). *Knowledge of language: Its nature, origin and use*. New York, NY: Praeger.

Cleeremans, A. (2008). Consciousness: The radical plasticity thesis. In R. Banerjee & B. K. Chakrabarti (Eds.), *Models of brain and mind: Physical, computational and psychological approaches, progress in brain research* (Vol. 168, pp. 19–33). Amsterdam, Netherlands: Elsevier. https://doi.org/10.1016/S0079-6123(07)68003-0

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

Cleeremans, A. (2014). Connecting conscious and unconscious processing. *Cognitive Science*, *38*, 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*, 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*, 905–912. https://doi.org/10.1111/j.1467-9280.2006.01801.x

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

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, *25*, 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*, 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*, 486–492. https://doi.org/10.1126/science.aan8871

DeKeyser, R. (2003). Implicit and explicit learning. In C. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 312–348). Oxford, UK: Blackwell. https://doi.org/10.1002/9780470756492.ch11

Dennett, D. (1993). *Consciousness explained*. London, UK: Penguin.

Descartes, R. (1637/ 1968). *Discourse on method and the meditations*. Harmondsworth, UK: Penguin.

de Waal, F. (2016). *Are we smart enough to know how smart animals are?* New York, NY: Norton.

Dienes, Z. (2004). Assumptions of subjective measures of unconscious mental states: Higher order thoughts and bias. *Journal of Consciousness Studies*, *11*, 25–45.

Dienes, Z. (2008). Subjective measures of unconscious knowledge. *Progress in Brain Research*, *168*, 49–64. https://doi.org/10.1016/S0079-6123(07)68005-4

Dienes, Z., Altmann, G., Kwan, L., & Goode, A. (1995). Unconscious knowledge of artificial grammars is applied strategically. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1322–1338. Retrieved from http://doi.org/10.1037/0278-7393.21.5.1322

Dienes, Z., & Berry, D. C. (1997). Implicit learning: Below the subjective threshold. *Psychonomic Bulletin and Review*, *4*, 3–23. https://doi.org/10.3758/BF03210769

Dienes, Z., Broadbent, D. E., & Berry, D. C. (1991). Implicit and explicit knowledge bases in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 875–882. https://doi.org/10.1037/0278-7393.17.5.875

Dienes, Z., & Fahey, R. (1995). The role of specific instances in controlling a dynamic system. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 848–862. https://doi.org/10.1037/0278-7393.21.4.848

Eigsti, I-M., Zayas, V., Mischel, W., Shoda, Y., Ayduk, O., Dadlani, M. B., … Casey, B. J. (2006). Predicting cognitive control from preschool to late adolescence and young adulthood. *Psychological Science*, *17*, 478–484. https://doi.org/10.1111/j.1467-9280.2006.01732.x

Ellis, N. C. (2003). Constructions, chunking, and connectionism: The emergence of second language structure. In C. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 63–103). Oxford, UK: Blackwell. https://doi.org/10.1002/9780470756492.ch4

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). Amsterdam, Netherlands: John Benjamins. https://doi.org/10.1075/sibil.48.01ell

Endress, A. D., & Bonatti, L. L. (2007). Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition*, *105*, 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

Foote, A. L., & Crystal, J. D. (2007). Metacognition in the rat. *Current Biology*, *17*, 551–555. https://doi.org/10.1016/j.cub.2007.01.061

Frank, M. C., Goodman, N. D., & Tenenbaum, J. B. (2009). Using speakers' referential intentions to model early cross-situational word learning. *Psychological Science*, *20*, 578–585. https://doi.org/10.1111/j.1467-9280.2009.02335.x

Frank, M. C., Tenenbaum, J. B., & Fernald, A. (2013). Social and discourse contributions to the determination of reference in cross-situational word learning. *Language Learning and Development*, *9*, 1–24. https://doi.org/10.1080/15475441.2012.707101

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

Fujita, K. (2009). Metamemory in tufted capuchin monkeys *(Cebus apella)*. *Animal Cognition*, *12*, 575–585. https://doi.org/10.1007/s10071-009-0217-0

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*, 85–95. https://doi.org/10.1016/j.lindif.2010.10.005

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511790942

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

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

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

Hampton, R. R. (2001). Rhesus monkeys know when they remember. *PNAS*, *98*, 5359–5362. https://doi.org/10.1073/pnas.071600998
Hampton, R. R. (2009). Focusing the uncertainty about nonhuman metacognition. *Comparative Cognition & Behavior Reviews*, *4*, 56–57. https://doi.org/10.3819/ccbr.2009.40006
Hamrick, P., & Rebuschat, P. (2012). How implicit is statistical learning? In P. Rebuschat & J. N. Williams (Eds.), *Statistical learning and language acquisition* (pp. 365–382). Berlin, Germany: Mouton de Gruyter.
Han, Z., & Finneran, R. (2014). Re-engaging the interface debate: Strong, weak, none, or all? *International Journal of Applied Linguistics*, *24*, 370–389. https://doi.org/10.1111/ijal.12034
Hulstijn, J. (2005). Theoretical and empirical issues in the study of implicit and explicit language learning. *Second Language Acquisition*, *27*, 129–140. https://doi.org/10.1017/S0272263105050084
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). Amsterdam, Netherlands: John Benjamins. https://doi.org/10.1075/sibil.48.02hul
Jaeger, T. F. (2009, May 14). Random effect: Should I stay or should I go? [Blog post]. Retrieved from https://hlplab.wordpress.com/2009/05/14/random-effect-structure/
James, W. (1890). *The principles of psychology*. New York, NY: Henry Holt.
Jiménez, L., Méndez, C., & Cleeremans, A. (1996). Comparing direct and indirect measures of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 948. https://doi.org/10.1037/0278-7393.22.4.948
Johnson, C., & Goswami, U. (2010). Phonological awareness, vocabulary, and reading in deaf children with cochlear implants. *Journal of Speech, Language, and Hearing Research*, *53*, 237–261. https://doi.org/10.1044/1092-4388(2009/08-0139)
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
Ke, S., Miller, R. T., Zhang, D., & Koda, K. (2020). Crosslinguistic sharing of morphological awareness in biliteracy development: A systematic review and meta-analysis of correlation coefficients. *Language Learning*, https://doi.org/10.1111/lang.12429
Kerz, E., Wiechmann, D., & Riedel, F. B. (2017). Implicit learning in the crowd: Investigating the role of awareness in the acquisition of L2 knowledge. *Studies in Second Language Acquisition*, *39*, 711–734. https://doi.org/10.1017/S027226311700002X
Kidd, E. (2011). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology*, *48*, 171–184. https://doi.org/10.1037/a0025405

Krashen, S. D. (1981). *Second language acquisition and second language learning*. Oxford, UK: Pergamon Press.

Kruskal, W., & Mosteller, F. (1979). Representative sampling: 3: Current statistical literature. *International Statistical Review*, *47*, 245–265. https://doi.org/10.2307/1402647

Kukull, W., & Ganguli, M. (2012). Generalizability: The trees, the forest, and the low-hanging fruit. *Neurology*, *78*, 1886–1891. https://doi.org/10.1212/WNL.0b013e318258f812

Lamme, V. A. F. (2010). How neuroscience will change our view on consciousness. *Cognitive Neuroscience*, *1*, 204–220. https://doi.org/10.1080/17588921003731586

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

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

Lowie, W. M., & Verspoor, M. H. (2019). Individual differences and the ergodicity problem. *Language Learning*, *69*, 184–206. https://doi.org/10.1111/lang.12324

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*, *68*, 321–391. https://doi.org/10.1111/lang.12286

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*, 300–335. https://doi.org/10.3102/0034654310377087

Merikle, P. M., Smilek, D., & Eastwood, J. D. (2001). Perception without awareness: Perspectives from cognitive psychology. *Cognition*, *79*, 115–134. https://doi.org/10.1016/S0010-0277(00)00126-8

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science*, *244*, 933–938. Retrieved from http://doi.org/10.1126/science.2658056

Monaghan, P., Mattock, 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*, 1099–1112. https://doi.org/10.1111/cogs.12186

Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface*, *14*, 1–11. https://doi.org/10.1098/rsif.2017.0213

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

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

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*, 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*, 246–263. https://doi.org/10.1006/jmla.1998.2576

Pinto, M., & Zuckerman, S. (2019). Coloring Book: A new method for testing language comprehension. *Behavior Research Methods*, *51*, 2609–2628. https://doi.org/10.3758/s13428-018-1114-8

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.

Proust, J. (2013). *The philosophy of metacognition*. Oxford, UK: Oxford University Press.

Psychology Software Tools (2012). E-Prime 3.0 [Computer software]. Retrieved from http://www.pstnet.com/

Radford, A. (2004). *English syntax: An introduction*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511841675

R Core Team. (2015). R: A language and environment for statistical computing [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from R-project.org/

Reber, A. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*, 855–863. https://doi.org/10.1016/S0022-5371(67)80149-X

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

Rebuschat, P., Hamrick, P., Riestenberg, K., Sachs, R., & Ziegler, N. (2015). Triangulating measures of awareness: A contribution to the debate on learning without awareness. *Studies in Second Language Acquisition*, *37*, 299–334. https://doi.org/10.1017/S0272263115000145

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). Berlin, Germany: Mouton de Gruyter.

Reed, J., & Johnson, P. (1994). Assessing implicit learning with indirect tests: Determining what is learned about sequence structure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 585. https://doi.org/10.1037/0278-7393.20.3.585

Reingold, E. M., & Merikle, P. M. (1988). Using direct and indirect measures to study perception without awareness. *Perception & Psychophysics*, *44*, 563–575. https://doi.org/10.3758/BF03207490

Reingold, E. M., & Merikle, P. M. (1990). On the inter-relatedness of theory and measurement in the study of unconscious processes. *Mind & Language*, *5*, 10–28. https://doi.org/10.1111/j.1468-0017.1990.tb00150.x

Reingold, E. M., & Toth, J. P. (1996). Progress dissociations versus task dissociations: A controversy in progress. In G. Underwood (Ed.), *Implicit cognition* (pp. 159–202). Oxford, UK: Oxford University Press.

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*, 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*, 606–621. https://doi.org/10.1006/jmla.1996.0032

Saygin, Z. M., Norton, E. S., Osher, D. E., Beach, S. D., Cyr, A. B., Ozernov-Palchik, O., … 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*, 13251–13258. https://doi.org/10.1523/JNEUROSCI.4383-12.2013

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*, 129–158. https://doi.org/10.1093/applin/11.2.129

Schmidt, R. W. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge: Cambridge University Press.

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

Smith, J. D. (2009). The study of animal metacognition. *Trends in Cognitive Sciences*, *13*, 389–396. https://doi.org/10.1016/j.tics.2009.06.009

Smith, J. D. (2010). Inaugurating the study of animal metacognition. *Journal of Comparative Psychology*, *23*, 401–413.

Smith, L. B., Suanda, S. H., & Yu, C. (2014). The unrealized promise of infant statistical word-referent learning. *Trends in Cognitive Sciences*, *18*, 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*, 1558–1568. https://doi.org/10.1016/j.cognition.2007.06.010

Sole, L. M., Shettleworth, S. J., & Bennett, P. J. (2003). Uncertainty in pigeons. *Psychonomic Bulletin & Review*, *10*, 738–745. https://doi.org/10.3758/BF03196540

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*, 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., & Vet, D. J. (2019, November 11). Meta-linguistic awareness in early (second) language acquisition. https://doi.org/10.17605/osf.io/bp5qe

Suda-King, C. (2008). Do orangutans (*Pongo pygmaeus*) know when they do not remember? *Animal Cognition*, *11*, 21–42. https://doi.org/10.1007/s10071-007-0082-7

Suzuki, Y., & DeKeyser, R. (2017). The interface of explicit and implicit knowledge in a second language: Insights from individual differences in cognitive aptitude. *Language Learning*, *67*, 747–790. https://doi.org/10.1111/lang.12241

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*, 792–814. Retrieved from http://doi.org/10.1037/a0030801

Timmermans, B., & Cleeremans, A. (2015). How can we measure awareness? An overview of current methods. In M. Overgaard (Ed.), *Behavioural methods in consciousness research* (pp. 21–46). Oxford, UK: Oxford University Press. https://doi.org/10.1093/acprof:oSo/9780199688890.003.0003

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

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

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). Amsterdam, The Netherlands: 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*, 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*, 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*, 729–742. https://doi.org/10.1016/j.cognition.2007.08.007

Walker, N., Monaghan, P., Schoetensack, C., & Rebuschat, P. (2020). Distinctions in the acquisition of vocabulary and grammar: An individual differences approach. *Language Learning*, *70*(S2), 221–254. https://doi.org/10.1111/lang.12395

Wijnen, F. (2013). Acquisition of linguistic categories: Cross domain convergences. In J. Bolhuis & M. Everaert (Eds.), *Birdsong, speech, and language* (pp. 157–177). Cambridge, MA: MIT Press.

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

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

Appendix: Accessible Summary (also publicly available at https://oasis-database.org)

Do Kindergarteners Develop Awareness of the Grammatical Structures They Acquire?

What This Research Was About and Why It Is Important

Child and adult language acquisition may be different from each other in many ways. For example, adults are likely to become aware of grammatical patterns in languages they learn. Some theories even suggest that such awareness is necessary for learning. But children are thought to learn language implicitly, unaware of grammatical patterns. Indeed, awareness has played no role in theories of child language acquisition to date. But that could be because it is difficult to measure awareness in children. Unlike adults, they cannot put their knowledge of language into words. In this article, we present a new experimental method that assesses whether kindergarteners develop such awareness. We used a behavioral measure of awareness, and found that children were sufficiently aware of their language knowledge to make strategic decisions about the language. We argue that if researchers use such a method, young children might show awareness of grammatical structures.

What the Researchers Did

- 63 children in kindergarten (aged from 4 years and 3 months to 6 years and 10 months) were asked to learn a novel artificial language in approximately 20 minutes. They saw pictures and heard sentences like: *Maria rigarda tra zambo* (Maria eats [an] apple[s]) and *Julia rigarda pli bovo* (Julia eats strawberries). The word *pli* indicated the noun was plural; the word *tra* could be used before singular and plural nouns.
- After hearing the language, children's knowledge of these plural markers was tested: children heard a sentence from the artificial language and had to pick either a picture with a single or with a plural noun.
- In this test, the sentences with *pli* were easy if you had picked up that *pli* always indicated a plural noun; other sentences were difficult because they used *tra* and that did not help decide which picture was correct.

- In the final part of the experiment, children would obtain a reward for giving a correct answer. However, they were also allowed to "gamble": they could opt out of giving an answer to obtain a smaller reward.
- If children developed awareness of the plural marker, they should opt out when hearing a difficult sentence (with *tra*), because of uncertainty, but they should try to answer when hearing an easy sentence (with *pli*), because they could be certain that a picture with a plural noun would be the right one.

What the Researchers Found

- In our experiment, kindergarteners demonstrated knowledge of the plural marker; they were able to acquire it in the limited time they had, though this was not entirely statistically reliable.
- When asked directly about it, however, they were unable to describe the rule.
- Their gambling behavior showed evidence of strategic behavior, suggesting that at least some children were sufficiently aware of their language knowledge to use it to maximize their reward.

Things to Consider

- Using a task that does not require children to put their knowledge into words, this study suggests that awareness might also play a role in child language learning. Measuring such awareness is challenging and requires the use of methods that enable children to reveal their awareness behaviorally.
- New measures of awareness allow us to address new questions: Is some degree of awareness necessary for learning in children? Do all children develop awareness or do certain characteristics determine if they will?

Materials, data, open access article: Materials and data are publicly available at osf.io/bp5qe and iris-database.org.

How to cite this summary: Spit, S. (2021). Do kindergarteners develop awareness of the grammatical structures they acquire? *OASIS Summary* of Spit, Andringa, Rispens et al. (2021) in *Language Learning*. https://oasis-database.org

This summary has a CC BY-NC-SA license.