TOPICS
TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science 11 (2019) 504–519

ISSN: 1756-8765 online
DOI: 10.1111/tops.12428

This article is part of the topic "Two Approaches, One Phenomenon: Aligning Implicit Learning and Statistical Learning," Padraic Monaghan and Patrick Rebuschat (Topic Editors). For a full listing of topic papers, see http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1756-8765/earlyview

Statistical Learning, Implicit Learning, and First Language Acquisition: A Critical Evaluation of Two Developmental Predictions

Inbal Arnon

Department of Psychology, Hebrew University of Jerusalem

Received 28 March 2018; received in revised form 22 March 2019; accepted 4 April 2019

Abstract

The role of distributional information in language learning, and learning more generally, has been studied extensively in both the statistical learning and the implicit learning literatures. Despite the similarity in research questions, the two literatures have remained largely separate. Here, we draw on findings from the two traditions to critically evaluate two developmental predictions that are central to both. The first is the question of age invariance: Does learning improve during development or is it fully developed in infancy? The combined findings suggest that both implicit and statistical learning improve during childhood, contra the age invariance prediction. This raises questions about the role of implicit statistical learning (ISL) in explaining the age-related deterioration in language learning: Children's better language learning abilities cannot be attributed to their improved distributional learning skills. The second issue we examine is the predictive relation to language outcomes: Does variation in learning predict variation in language outcomes? While there is evidence for such links, there is concern in both research traditions about the reliability of the tasks used with children. We present data suggesting that commonly used statistical learning measures may not capture stable individual differences in children, undermining their utility for assessing the link to language outcomes in developmental samples. The evaluation of both predictions highlights the empirical parallels between the implicit and statistical learning literatures, and the need to better integrate their developmental investigation. We go on to discuss several of the open challenges facing the study of ISL during development.

Correspondence should be sent to Inbal Arnon, Department of Psychology, Hebrew University of Jerusalem, Mount Scopus, Jerusalem, Israel 9190501. E-mails: inbal.arnon@gmail.com; inbal.arnon@mail.huji.ac.il

Keywords: Statistical learning; Implicit learning; Age invariance; Individual differences; Reliability; Children

1. Introduction

How children learn about their environment is one of the most fundamental and hotly debated questions in cognitive science, and one that has been at the center of the nature/nurture debate: Can children learn structure simply from the information available to them in their environment or does such learning necessitate innate knowledge or constraints? The answer, in part, depends on whether there is sufficient information in the environment, and whether humans can utilize it to learn higher order structure. If they can, such a distributional learning could provide an alternative to innate knowledge. In accordance with its theoretical significance, humans' ability to detect and learn recurring patterns has been studied for over a century (Esper, 1925; see Christiansen, 2019, and Perruchet, 2019, for reviews). The interest in such distributional learning has led to the emergence of two distinct literatures: that studying *implicit learning* and that studying *statistical learning*. While there is not one agreed-upon definition for either term, implicit learning can be broadly defined as *"adaptation to the regularities of the world that evolves without intention to learn, and without clear awareness of what we know"* (Perruchet & Pacton, 2006, pp. 233), while statistical learning is often construed more broadly as *"the discovery of patterns in the input,"* with less focus on the lack of intention or awareness (Romberg & Saffran, 2010, p. 906).

Despite the fundamental similarity in the domain of inquiry, the two literatures have remained largely separate, with little interaction between them. The findings are typically published in different journals, presented at different conferences, and favor different explanations for learners' behavior: For instance, while the statistical learning literature emphasizes statistical computations, the implicit learning one highlights chunk formation (Christiansen, this topic; Perruchet, this topic; Perruchet & Pacton, 2006). Is this separation theoretically justified? Is it empirically useful? What can be gained by integrating the findings?

The goal of this paper is to examine developmental issues through the combined lenses of implicit and statistical learning as a way to gain insight on specific questions, as well as on the overall similarity (or difference) between the two. I evaluate two developmental predictions that are of theoretical importance in both research traditions: age invariance (the idea that learning is fully developed in infancy and does not improve with age), and the predictive relation to language outcomes (I discuss both in more detail below). The examination of both issues is important for assessing the role of distributional learning in language acquisition, which is a central research topic in both traditions. This evaluation aims to provide novel evidence on both issues, while highlighting the empirical parallels in the developmental findings on implicit and statistical learning. These parallels, in turn, suggest that the two literatures are studying a fundamentally similar phenomenon, and that their developmental research should be better integrated. I end with a discussion of future directions and open challenges that face the study of implicit and statistical learning during development.

1.1. Choice of terminology and some caveats

Throughout this paper, I use the term "implicit statistical learning (ISL)" to refer to the research conducted in both traditions. The use of this term does not imply they are the same thing, but it is a matter of convenience. I return to the question of whether or not they investigate the same ability in the discussion. I have not included findings from the procedural learning literature, even though it is a closely related literature, with related developmental findings (Nicolson & Fawcett, 2007). This omission reflects the limited scope of the current paper, and the less central place of language acquisition in it. I focus on studies with typically developing children (not infants and not atypically developing children), because it is the most relevant population for examining the two predictions.

1.2. The relation between implicit learning and statistical learning

The term "implicit learning" was first coined by Reber (1967) to describe participants' learning of abstract rules that governed the creation of letter strings. Importantly, Reber's (1967) participants were unable to verbally account for their performance, which was taken as evidence for the implicit (unconscious) nature of the acquired knowledge. The term was later adapted to refer more broadly to the learning of recurring patterns (abstract or not) that happens without instruction or awareness. In the 50 years since Reber's original formulation, implicit learning has been studied extensively, with a shifting focus from language to sequence learning more generally, and increased emphasis on the role of consciousness in such learning processes and on their neural realizations (see Cleeremans et al., 1998; Perruchet, 2008; Shanks, 2005 for reviews). There seems to be agreement that implicit learning is relatively inflexible (shows specificity of transfer), is associated with incidental and not intentional learning conditions, and is fairly robust (maintained under time pressure and preserved in certain disorders, such as amnesia, Knowlton, Ramus, & Squire, 1998). There is less agreement on the extent to which the acquired knowledge is abstract and unconscious, and the degree to which such learning is independent from other cognitive capacities like memory or attention.

The field of statistical learning is newer: Its more recent revival can be traced back to a seminal study showing that 8-month-old infants can use the transitional probabilities between syllables to detect word boundaries in an artificial language (Aslin, Saffran, & Newport, 1998; Saffran, Newport, & Aslin, 1996). This finding opened up novel theoretical and empirical horizons for investigating the centuries-old question of how children learn language. During the second half of the 20th century, following Chomsky's formulation of generative grammar (Chomsky, 1965) children's unique capacity for learning language was often attributed to innate language-specific learning mechanisms and constraints. A growing number of models and theories, however, advocate alternative explanations of how children learn language, emphasizing the importance of experience and learning in the acquisition and representation of linguistic knowledge (Tomasello, 2003). These approaches share the common assumption that it is feasible and possible to learn language structure from the linguistic input children hear. Saffran et al. (1996) provided an important demonstration of infants' ability to learn word boundaries by using distributional information. If substantiated, statistical learning could provide a plausible alternative to nativist accounts of first language acquisition.[1] Over the past 20 years, statistical learning has been investigated extensively. This body of research shows that statistical learning is present from infancy, including in newborns (Bulf, Johnson, & Valenza, 2011), found across modalities (Conway & Christiansen, 2005; Kirkham, Slemmer, & Johnson, 2002), and can facilitate learning of various linguistic relations and regularities, from the sound to the phrase level (see Romberg & Saffran, 2010; Saffran & Kirkham, 2018 for reviews).

These literatures have developed largely in parallel, in part because of differences in the specific research questions and methods. Research on statistical learning often examines *what humans can learn* (what kinds of relations, under what conditions) while the implicit learning research is also concerned with the question of how this knowledge is represented (e.g., abstract rules, exemplars, chunks) and the degree to which it is consciously available. The two literatures differ in their commitment to, and interest in, developmental research. While both make predictions about learning, the implicit learning research has largely focused on adult learners. Statistical learning research, in contrast, has been substantially influenced by studies on infants, and it has maintained a stronger developmental focus. The separation is also evident in the methods used: Studies of implicit learning often use artificial grammar learning (AGL), where participants have to learn the rules underlying letter strings and then distinguish between grammatical and ungrammatical strings, or sequence learning tasks (of which the serial reaction task, SRT, is a commonly used example) where response times are compared to structured and unstructured sequences. Both tasks examine learners' sensitivity to the regularities governing sequence generation. Studies on statistical learning, in contrast, often focus on the use of distributional information to detect recurring units in a continuous stream, using tasks modeled on Saffran et al. (1996). While there are studies that use this method to probe learners' sensitivity to more abstract rules (see Rabagliati, Ferguson, & Lew-Williams, 2019, for a recent review and meta-analysis), much of this work examines learners' sensitivity to transitional probabilities. That is, the two traditions differ in their specific research questions, the kinds of regularities they examine, and the tasks they use to do so, all of which contribute to their separation.

However, despite these differences, the literatures share fundamental commonalities. Both study humans' ability to implicitly detect patterns in their environment, and both are interested in the role of such learning in language acquisition. Children's acquisition of language is seen a prime example of distributional learning in the wild: It is a motivating force and a puzzle to explain in both traditions. While the statistical learning literature is more language-centric than the implicit learning literature, they nevertheless make (and share) some core developmental predictions. I focus here on two of these: age invariance and the predictive relation to language outcomes. If ISL plays an important role in language acquisition, it should be present from early on. Age (Reber, A. S. (1967); Esper, E. A. (1925); Cleermans, A., Destrebecqz, A. & Boyer, M. (1998)). invariance is also theoretically motivated by the impact of age on language learning: Unlike many other skills, language learning is at its prime during infancy and does not improve with age (Hartshorne et al., 2018; Johnson & Newport, 1989). If ISL plays an important part in this process, it may show a similar developmental trajectory. A second core prediction that is made in both literatures is that variation in ISL abilities should be predictive of variation in language outcomes: Individuals who are better at ISL should show better language outcomes. Such a link would substantiate the role of distributional learning for language: Group-level findings demonstrate that humans *can* learn regularities, but they do not show that such mechanisms modulate the actual process of language acquisition. In the next section, I evaluate the evidence for age invariance. In the following section I do the same for the predictive relation to language outcomes. I end with a discussion of what the combined findings can tell us about the relation between implicit and statistical learning.

2. Age invariance during development

Implicit learning is traditionally assumed to be early maturing and age invariant, and to be present and fully developed early on.[2] This claim was first made by Reber (1993) on the basis of unpublished data suggesting that children between the ages of 4 and 14 perform similarly on a modified Artificial Grammar Learning (AGL) task (they were similarly accurate in distinguishing between grammatical and ungrammatical strings). Age invariance was widely accepted as a basic property of implicit learning, despite the scant empirical evidence for it, and despite the fact that most other cognitive abilities improve with age (e.g., working memory, Gathercole, 1998; executive function, Anderson, 2002; and many more). One of the first published papers examining implicit learning during development actually showed improvement with age: Eleven year olds were better at learning co-varying visual cues than 6 year olds (Maybery, Taylor, & O'brien-Malone, 1995). A subsequent paper, using SRT, provided support for age invariance by showing similar rates of learning in younger children (6 year olds), older children (10 year olds), and adults (Meulemans, Van der Linden, & Perruchet, 1998; see also Asmo & Davidow, 2012). Two recent papers even found better learning rates in younger children (ages 4–12) compared to adolescents and adults (ages 12–60) on an Alternating Serial Reaction Time (ASRT) task where repeating sequences alternate with random ones (Janacsek, Fiser, & Nemeth, 2012; Nemeth, Janascek, & Fischer, 2013, but see Lukacs & Kemeny, 2015, for a critique of their use of raw RTs and not normalized ones).

At the same time, there is growing evidence for a positive effect of age on implicit learning between childhood and adulthood (Fletcher, Maybery, & Bennett, 2000; Thomas et al., 2004; Weiermann & Meier, 2012). The most comprehensive assessment to date shows clear age-related improvements during childhood (Lukács & Kemény, 2014). This large-scale study (N = 480) examined performance on three implicit learning tasks (SRT, AGL, and the WP task testing non-sequential probabilistic categorization) in participants between the ages of 7 and 85 years and found that learning improved until 18, and then remained relatively stable before showing a decline after 65. They suggest that the effect of age may have been masked in previous studies by testing children of mixed ages (e.g., treating 6–10 as one age group) and using small sample sizes. A similar conclusion is reached in a systematic review of studies using the SRT task across development (Zwart et al., 2017).[3] This paper reviews 50 studies examining implicit learning across the life span: The majority of them support an age-variant model, whereby implicit sequence learning improves from childhood to adolescence and then remains stable before declining again with aging. Interestingly, the choice of dependent variable impacts the observed effect of age: Studies using raw RTs (that are not corrected for developmental changes in motor responses) report no improvement during childhood followed by a decline from early adolescence, while studies using corrected RTs show an improvement during childhood followed by a decline with aging. The convergent findings provide weak support for age invariance and suggest instead that implicit learning improves during childhood.

Statistical learning is also described as an early-maturing capacity that is not expected to improve with age (Saffran, Newport, Aslin, Tunick, & Barrueco, 1997). Early evidence supported this prediction. Six year olds were as good as adults in a word segmentation task (Saffran et al., 1997), although they were worse than adults in learning phrase-structure rules (Saffran, 2001, 2002). Findings in the visual domain, however, indicated better performance in adults compared to children: In two studies, adults showed higher accuracy compared to children (11 year olds) and adolescents (Bertels, Boursain, Destrebecqz, & Gaillard, 2015; Schlichting, Guarino, Shapiro, Turk-Browne, & Preston, 2017). As in the implicit learning literature, the support for the prediction of age invariance comes from a handful of studies, none of which used large samples of cross-sectional data.

Several recent papers using cross-sectional data suggest that statistical learning does improve with age, at least for non-linguistic stimuli. Children's visual statistical learning improved with age between the ages of 5 and 12 years (Arciuli & Simpson, 2011). Raviv and Arnon (2018) expanded on these findings to ask if the effect of age on performance is modality-specific. They examined visual statistical learning (using alien shapes) and auditory statistical learning (using syllables) in children between the ages of 5 and 12 years (N = 230). Visual statistical learning improved with age, but auditory statistical learning did not. These findings were interpreted to reflect modality-based differences in statistical learning. The impact of modality could help explain the mixed effects of age in prior work: Studies using auditory stimuli found age invariance (Saffran et al., 1997), while ones using visual stimuli did not (Arciuli & Simpson, 2011). Recent work casts doubt on this interpretation: Since the auditory task used linguistic stimuli (syllables), the differential effect of age may have been driven by stimulus type (linguistic vs. non-linguistic) rather than modality. Using a similar design, Shufaniya and Arnon (2018) examined children's performance on visual and non-linguistic auditory statistical learning (using familiar sounds instead of syllables) across the same age range (5–12 years, N = 229) and found a strikingly different pattern. Both visual and non-linguistic auditory statistical learning improved with age, suggesting that it was the linguistic nature of the stimuli, rather than its auditory modality, that led to age invariance in the previous study. The distinct trajectory for auditory linguistic stimuli could reflect the unique relevance of such learning for language. However, it could also reflect the greater sensitivity of the ASL task to prior knowledge (Siegelman, Bogaerts, Elazar, Arciuli, & Frost, 2018), and in particular to syllable phonotactics, which are already mastered by age 5 (see Shufaniya & Arnon, 2018, for further discussion). In sum, when larger cross-sectional samples are used, there is growing evidence for a positive effect of age on statistical learning as well.

The combined findings suggest that age variance is more prevalent than age invariance across the two literatures. They highlight the similarity between distributional learning and other cognitive abilities, which improve with age (memory, attention, executive function), as well as the similarity between implicit learning and statistical learning. Once large enough cross-sectional data were used, performance improved with age across tasks (SRT, AGL, ASL, VSL), learning measures (reaction times, explicit judgments), and regularities (sequence learning, transitional probabilities, grammar induction), without any clear difference between studies of implicit and statistical learning. At least from the perspective of age invariance, the two seem more similar than distinct. What do these findings tell us about the role of ISL in language acquisition and its relation to critical period accounts? There are several possible interpretations. The first is that ISL does deteriorate between infancy and early childhood, along with language learning abilities: None of the existing studies have examined the developmental trajectory in children younger than 5, because of the difficulty in finding tasks that can be used with that age range. A second interpretation, similar in spirit to the first, is that ISL is at its prime in infancy, and the detected improvement is driven by changes in other related abilities like memory or attention. While theoretically possible, it would imply a similar pace of change across modalities (e.g., for visual and auditory working memory). An additional, and more parsimonious, interpretation is that the deterioration in language learning abilities is related to the growing impact of prior knowledge and experience (MacWhinney, 2005). So that ISL improves with practice—experience with detecting patterns makes you better—while making it harder to learn relations that differ from the ones you know (Finn & Hudson Kam, 2008).

3. The predictive relation of implicit statistical learning to language outcomes

One of the ways to substantiate the role of ISL in language acquisition is by showing that variation in learning is predictive of variation in language outcomes: Such findings are crucial for evaluating the role of ISL in learning language. While the connection to language outcomes is predicted in both literatures, more of the research on individual differences and language is conducted within the statistical learning tradition, in line with its more language-centric and developmental focus. As a result, I focus in this section primarily on the developmental work conducted within the statistical learning literature and review the implicit learning findings in less depth. I return to the implicit learning findings at the end of this section to highlight the empirical parallels between the two bodies of research.

Over the past few years, a growing number of studies has examined variation in statistical learning performance on an individual level. These studies show that learners do vary in their statistical learning abilities, and that this variation is correlated with language outcomes in both children and adults. Statistical learning is positively correlated with second language reading, syntactic processing, and speech perception (Conway, Bauernschmidt, Huang, & Pisoni, 2010; Frost, Siegelman, Narkiss, & Afek, 2013; Misyak & Christiansen, 2012) in adults, and with literacy acquisition, syntactic processing, and vocabulary size in typically developing children learning their first language (Arciuli & Simpson, 2012; Kidd, 2012; Kidd & Arciuli, 2016; Spencer, Kaschak, Jones, & Lonigan, 2015; Qi et al., 2018). Correlations between statistical learning and vocabulary size were also found for children with SLI and autism (Haebig, Saffran, & Ellis Weismer, 2017; Mainela-Arnold & Evans, 2014). In infancy, visual statistical learning is correlated with later vocabulary scores (Shafto, Conway, Field, & Houston, 2012), although this was not found for all ages and language measures, and auditory statistical learning was predictive of real-time language processing in 15-month-old infants (Lany, Shoaib, Thompson, & Graf Estes, 2017).[4] Together, these findings strengthen the predicted link between statistical learning and language learning: They suggest that variation in statistical learning is associated with variation in language outcomes.

However, this interpretation is valid only if our statistical learning measures are reliable and capture stable individual variation in both children and adults. If this is not the case, then correlating them with other measures is not meaningful. Recent work with adults suggests there is reason for concern: A series of studies argue that commonly used statistical learning tasks may not be suitable for assessing individual differences (Siegelman, Bogaerts, Christiansen, & Frost, 2017; Siegelman, Bogaerts, & Frost, 2017). Many individual difference studies use Saffran's original word segmentation task (ASL) or a visual parallel using shapes instead of syllables (VSL, based on Turk-Browne, Junge, & Scholl, 2005, see table 1 in Siegelman, Bogaerts, Christiansen, et al., 2017). Because both tasks were designed to assess group-level performance, they suffer from several psychometric weaknesses when used as a measure of individual differences (Siegelman, Bogaerts, Christiansen, et al., 2017). In particular, they use relatively few testing trials, all at the same level of difficulty; repeat items and foils during testing; assess learning using forced choice trials; and have performance accuracy that is often close to chance, all of which impact task reliability and can lead to the detection of spurious correlations and to the underdetection of true ones (Siegelman, Bogaerts, & Frost, 2017). Despite their shortcomings, both the ASL and the VSL seem to show moderate reliability in adults (Potter, Wang, & Saffran, 2017; Siegelman & Frost, 2015; Siegelman, Bogaerts, & Frost, 2017). Reliability is increased when the tasks' shortcomings are modified by changing the nature and number of test trials (Siegelman, Bogaerts, & Frost, 2017) or assessing learning in more implicit ways, for example, by utilizing processing-based rather than reflection-based measures (Isbilen, Mccauley, & Christiansen, 2017; Siegelman, Bogaerts, Kronenfeld, & Frost, 2017).

Importantly, the reliability of statistical learning tasks has not been examined in children, even though there is reason to think it could be lower. Children may be more affected by the repetition of test items and foils and by the explicit nature of the test trials. Their lower overall accuracy means that there will be more participants who are at chance and whose scores do not reflect meaningful variation. Moreover, the need to keep the tasks child friendly may lead to further reducing the number of triplets and test items. Table 1 gives an overview of the tasks used and the correlations found in recent studies of individual differences in children. As can be seen, the majority of studies use either the VSL or the ASL with similar exposure and test properties. More worryingly, the correlations between SL and language outcomes during development are weaker than those found for adults (ranging between $r = .1–.46$), and somewhat inconsistent even when using the same measures. For instance, Kidd (2012) and Kidd and Arciuli (2016) found no correlation between statistical learning and vocabulary measures while Spencer et al. (2015) did. The low correlations and the fluctuations between studies may both stem from the tasks not having sufficient reliability.

In a recent study, we examined the internal consistency and test/retest reliability of three statistical learning measures (two auditory and one visual) in children and adults by looking at their performance on the same tasks twice, 2 months apart (Arnon, 2019). We used tasks that were closely modeled on ones previously used in the child individual difference literature in terms of their design properties: (a) a word segmentation task modeled on Saffran et al. (1996), (b) a non-linguistic version of this task that uses familiar sounds instead of syllables (e.g., door slamming, drum beating), and (c) a visual task modeled on Arciuli and Simpson (2011) that uses drawings instead of alien cartoons. The three tasks had identical exposure and test properties: Learners were exposed to a continuous stream made up of five recurring triplets. The TPs within triplets was 1 while the TPs between triplets was 0.25. Exposure lasted under 4 minutes, and learning was assessed using 25 2AFC trials. We tested children ($N = 42$, mean age 8;2) and adults ($N = 52$, mean age 23 years) with each participant completing all tasks twice.

Our results indicate that the tasks are not reliable in children. They showed moderate reliability in adults (see table 2 in Arnon, 2019). However, they did not capture stable individual variation in children, and they had internal consistency and reliability well below psychometric standards (see table 5 in Arnon, 2019).[5] The only task that showed any test/retest reliability was the ASL ($r = .33$, $p < .05$). This was, however, the only task where learners were exposed to the same stream (with the same triplets) in both sessions. The improved reliability could have been driven simply by memory of the specific triplets. To investigate this possibility, we ran an additional study looking only at the ASL (with a new sample of 44 children), but this time without repeating triplets between sessions. The task was no longer reliable (test/re-test of $r = -.15$), indicating that the previously found correlation reflected the repetition of triplets, and not an increased stability for the ASL. Taken together, the results suggest that these tasks cannot be used as a reliable measure of individual differences in children. Since they share important psychometric properties with previously used tasks, they raise a more general concern about the existing findings on the relation between statistical learning and language outcomes. If the tasks are not reliable, it is hard to interpret their correlation with other measures (see

Table 1
Overview of tasks used in recent statistical learning studies of child individual differences

Study	Task	Measure	Number of Triplets	Linguistic Measures	Obtained Correlations	Population
Arciuli and Simpson (2012)	VSL	64 2AFC	4	Reading	r = .33*	Children (ages 6–12)
Evans et al. (2009)	ASL	36 2AFC	6	Vocabulary	TD: r = .33** SLI: r = .02	Children with SLI and typically developing (ages 6–13)
Haebig et al. (2017)	ASL	32 2AFC	4	Vocabulary	TD: r = −.07 ASD: r = .46* SLI: .17	Children with SLI, children with autism, and typically developing (ages 8–12)
				Fast mapping	TD: r = .20 ASD: r = .26 SLI: r = .1	
Kidd (2012)	SRT	RTs	—	Vocabulary	r = .18	Children (ages 4–7)
Kidd and Arciuli (2016)	VSL	64 2AFC	4	Complex sentence comp	β = 2.33*	Children (ages 6–8)
				Vocabulary	r = .2	
Mainela-Arnold and Evans (2014)	ASL	36 2AFC	6	Gating task	TD: r = −.28 SLI: r = −.2	Children (ages 8–12)
				Word definition	TD: r = −.1 SLI: r = .2	Children with SLI (ages 8–12)
Spencer et al. (2015)	ASL visual AGL	4 2AFC	4	Ten literacy tasks	all r's < .2	Children (ages 4–10)
Qi et al. (2018)	ASL	32 2AFC	4	Word reading	r = .25	Children (ages 8–16)
				Non-word reading	r = .03	
	VSL	32 2AFC	4	Word reading	r = −.07	
				Non-word reading	r = −.07	
von Koss Torkildsen, Arciuli, and Wie (2019)	VSL	64 2AFC	4	Word reading	r = .3*	Children (ages 7–13)

Note. Only correlations for children are reported.
*Significance of $p < .05$.
**Significance of $p < .01$

Arnon, 2019 for a discussion of why these tasks may not be reliable during development).

Similar concerns about task reliability in children are found in the implicit learning literature. As mentioned earlier, fewer studies in this tradition examine the link between individuation variation in implicit learning and language outcomes during development.[6] Here also, the correlations with language outcomes seem relatively low and somewhat inconsistent. Gremp, Deocampo, Walk, and Conway (2019) found correlations between visual sequential processing and vocabulary size among deaf children ($r = .29$, $p = .014$), but not among hearing children (see also Conway et al., 2011). In contrast, Clark and Lum (2017) found a significant correlation between SRT performance and grammatical processing speed (measured using a picture matching task), for hearing children, but not deaf ones. However, SRT was not correlated with other language measures (word and non-word reading) in either group. As in the statistical learning literature, the reliability of these measures in children had not been assessed. A recent study examined the reliability of several implicit learning measures in children (among them the SRT task) and found that they displayed very poor reliability (West, Vadillo, Shanks, & Hulme, 2017). The authors conclude that such tasks cannot be used to investigate individual differences and their relation to other developmental measures.

Taken together, the findings raise serious concerns about our ability to evaluate the link between ISL and language outcomes. They suggest that existing findings about the predictive links to language cannot be taken as strong evidence for the role of ISL in language acquisition because of the low reliability of the tasks used. This is not to say that such learning mechanisms do not play an important role in learning language, but that our understanding of that role is constrained by the methodological limitations of the tasks used. Consequently, we should be careful about drawing strong conclusions about the impact of individual variation in implicit statistical learning. For instance, a number of recent studies report that the relation between ISL and language outcomes varies across populations of children (typically developing, SLI, ASD), and they use this to argue for underlying differences between the populations (Haebig et al., 2017). However, if the measures are not stable (no study examined their reliability in atypical populations), such differences are hard to interpret. Given the increased interest in the predictive power of ISL, there is pressing need for a systematic psychometric evaluation of the tasks used in both typical and atypical populations. The combined findings also point to striking parallels between the two literatures: Here also, the developmental patterns are very similar across the two literatures.

4. Conclusions

In this paper, we set out to evaluate two developmental predictions (age invariance and the predictive relation to language outcomes) by integrating findings from the implicit learning and statistical learning literatures. The combined findings show that ISL improves with age, across tasks, regularities and research traditions, contra the age invariance prediction. They also suggest that current findings on the predictive power of ISL should be interpreted with caution: Many of them seem to be based on tasks that do not capture stable individual variation in children. The discussion of both issues highlights some of the open challenges facing the study of distributional learning in development. The first is the development of reliable tasks that can be used from infancy to later childhood. Without such measures, we will not be able to assess the role of distributional learning in language acquisition, or its relation to other cognitive abilities. The second is the need to seriously ask whether there is a meaningful difference between implicit and statistical learning. The developmental data reviewed in this paper point to uncanny similarities in the findings from the two traditions: We would be hard pressed to find differences that cannot be traced back to the precise task used or the exact regularity tested. The continued separation of the two fields reduces our understanding of the power and limitations of distributional learning, with each tradition examining a corner of our ability to detect regularities. Future work should aim to integrate the two literatures by systematically examining the impact of task and regularity type on performance.

Acknowledgments

I thank Noam Siegelman, Louisa Bogaerts, and Ram Frost for helpful comments and discussions, as well as the editors of the volume and two anonymous reviewers for helpful and insightful comments. All remaining errors, are of course, my own. The research was funded by an Israeli Science Foundation grant to the author (grant number 584/16).

Notes

1. Acquisition and learning are often associated with different theoretical positions (nativist vs. usage-based). In this paper, I use the terms interchangeably to refer to the process of learning one's first language.
2. We focus here on the claim that performance does not improve during childhood, not on the possible deterioration with aging.
3. The paper looks at studies conducted with typically developing children, as well as children with specific language impairment (SLI) and autism spectrum disorder (ASD). I focus here only on the data from typically developing children.
4. Since the focus here is on SL tasks, we do not review the literature on the relation between speech perception measures in infancy and later language development but return to it in the discussion (see Cristia et al., 2014, for a review).
5. Qi et al. (2018) report higher internal consistency measures for a similar VSL task (they did not examine test re-test reliability). However, that study differs from ours in two important respects. First, that study looked at children spanning an older age range (8–16): The higher reliability may have been driven by the older participants. Second, each triplet was repeated eight times during testing, meaning that the

internal consistency measures may have reflected consistent performance on the same triplet, rather than consistency between items.

6. Many studies examine group-level differences in implicit learning, for instance between typically developing children and children with SLI or ASD, and relate those to differences in language outcomes at the group level. This work if of less relevance to the current discussion.

References

Amso, D., & Davidow, J. (2012). The development of implicit learning from infancy to adulthood: Item relations, salience, & cognitive flexibility. *Developmental Psychobiology*, *54*(6), 664–673.

Anderson, P. (2002). Assessment and development of executive function during childhood. *Child Neuropsychology*, *8*, 71–82.

Arciuli, J., & Simpson, I. C. (2011). Statistical learning in typically developing children: The role of age and speed of stimulus presentation. *Developmental Science*, *14*(3), 464–473.

Arciuli, J., & Simpson, I. C. (2012). Statistical learning is related to reading ability in children and adults. *Cognitive Science*, *36*(2), 286–304.

Arnon, I. (2019). Do current statistical learning tasks capture stable individual differences in children? An investigation of task reliability across modality. *Behavior Research Methods*. https://doi.org/10.3758/s13428-019-01205-5

Aslin, D. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *9*, 321–324.

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.

Bulf, H., Johnson, S. P., & Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition*, *121*(1), 127–132.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Cleermans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences*, *2*, 406–416.

Clark, G. M., & Lum, J. A. G. (2017). Procedural memory and speed of grammatical processing: Comparison between typically developing children and language impaired children. *Research in Developmental Disabilities*, *71*, 237–247.

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

Conway, C. M., Pisoni, D. B., Anaya, E. M., Karpicke, J., & Henning, S. C. (2011). Implicit learning in deaf children with cochlear implants. *Developmental Science*, *14*, 69–82.

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31*(1), 24–39.

Cristia, A., Seidl, A., Junge, C., Soderstrom, M., & Hagoort, P. (2014). Predicting individual variation in language from infant speech perception measures. *Child Development*, *85*, 1330–1345. https://doi.org/10.1111/cdev.12193.

Esper, E. A. (1925). *A technique for the experimental investigation of associative interference in artificial linguistic material*. Philadelphia: Linguistic Society of America.

Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *52*, 321–336.

Finn, A. S., & Hudson Kam, C. L. (2008). The curse of knowledge: First language knowledge impairs adult learners' use of novel statistics for word segmentation. *Cognition*, *108*, 477–499.

Fletcher, J. M., Maybery, M. T., & Bennett, S. (2000). Implicit learning differences: A question of developmental level? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(1), 246–252.

Frost, R., Siegelman, N., Narkiss, A., & Afek, L. (2013). What predicts successful literacy acquisition in a second language? *Psychological Science*, *24*(7), 1243–1252.

Gathercole, S. E. (1998). The development of memory. *Journal of Child Psychology and Psyciatry*, *39*, 3–27.

Gremp, M. A., Deocampo, J. A., Walk, A. M., & Conway, C. H. (2019). Visual sequential processing and language ability in children who are deaf or hard of hearing. *Journal of Child Language*, 1–15.

Haebig, E., Saffran, J. R., & Ellis Weismer, S. (2017). Statistical word learning in children with autism spectrum disorder and specific language impairment. *The Journal of Child Psychology and Psychiatry*, *58*, 1251–1263.

Hartshorne, J. K., Tenenbaum, J. B., & Pinker, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition*.

Isbilen, E. S., McCauley, S. M., Kidd, E., & Christiansen, M. H. (2017). Testing statistical learning implicitly: A novel chunk-based measure of statistical learning. In G. Gunzelmann, A. Howes, T. Tenbrink & E. Davelaar (Eds.), *Proceedings of the 39th Annual Conference of the Cognitive Science Society* (pp. 564–569). Austin TX: Cognitive Science Society.

Janacsek, K., Fiser, J., & Nemeth, D. (2012). The best time to acquire new skills: Age-related differences in implicit sequence learning across the human lifespan. *Developmental Science*, *15*(4), 496–505.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*, 60–99.

Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology*, *48*(1), 171–184.

Kidd, E., & Arciuli, J. (2016). Individual differences in statistical learning predict children's comprehension of syntax. *Child Development*, *87*(1), 184–193.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain-general learning mechanism. *Cognition*, *83*, 4–5.

Knowlton, B. J., Ramus, S. J., & Squire, L. R. (1998). Intact artificial grammar learning in amnesia: Dissociation of classification learning and explicit memory for specific instances. *Psychological Science*, *3*, 172–179.

Lany, J., Shoaib, A., Thompson, A., & Graf Estes, K. (2017). Infant statistical-learning ability is related to real-time language processing. *Journal of Child Language*, 1–24. https://doi.org/10.1017/S0305000917000253

Lukács, Á., & Kemény, F. (2015). Development of different forms of skill learning throughout the lifespan. *Cognitive Science*, *39*, 383–404.

MacWhinney, B. (2005). A unified model of language acquisition. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). New York: Oxford University Press.

Mainela-Arnold, E., & Evans, J. L. (2014). Do statistical segmentation abilities predict lexical-phonological and lexical-semantic abilities in children with and without SLI? *Journal of Child Language*, *41*(2), 327–351.

Maybery, M., Taylor, M., & O'brien-Malone, A. (1995). Implicit learning: Sensitive to age but not IQ. *Australian Journal of Psychology*, *47*(1), 8–17.

Meulemans, T., Van der Linden, M., & Perruchet, P. (1998). Implicit sequence learning in children. *Journal of Experimental Child Psychology*, *69*(3), 199–221.

Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: An individual differences study. *Language Learning*, *62*(1), 302–331.

Nemeth, D., Janascek, K., & Fiser, J. (2013). Age-dependent and coordinated shift in performance between implicit and explicit skill learning. *Frontiers in Computational Neuroscience*, *7*, 147.

Nicolson, R. I., & Fawcett, A. J. (2007). Procedural learning deficits: Reuniting the developmental disorders? *Trends in Neurosciences*, *30*, 135–141.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*(5), 233–238.

Perruchet, P. (2008). Imlpicit learning. In J. Bryne (Ed.), *Learning and memory: A comprehensive reference*. (Vol. 2, pp. 597–621). Oxford, UK: Elsevier.

Potter, C. E., Wang, T., & Saffran, J. R. (2017). Second language experience facilitates statistical learning of novel linguistic materials. *Cognitive Science*, *41*, 913–927.

Qi, Z., Araujo, S. Y., Georgran, W. C., Gabrieli, J. D. E., & Arciuli, J. (2018). Hearing matters more than seeing: A cross-modality study of statistical learning and reading ability. *Scientific Studies of Reading*, *23*, 101–115.

Rabagliati, H., Ferguson, B., & Lew-Williams, C. (2019). The profile of abstract rule learning in infancy. Meta-analytic and experimental evidence. *Developmental Science*, *22*, e12704.

Raviv, L., & Arnon, I. (2018). The developmental trajectory of children's auditory and visual statistical learning abilities: Modality-based differences in the effect of age. *Developmental Science*, *21*, e12593.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*, 855–863.

Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. Oxford, England: Oxford University Press.

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

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science (New York, N.Y.)*, *274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926.

Saffran, J. R. (2001). The use of predictive dependencies in language learning. *Journal of Memory and Language*, *44*, 493–515.

Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language*, *47*(1), 172–196.

Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology*, *69*(1). https://doi.org/10.1146/annurev-psych-122216-011805

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the comer of your ear. *Psychological Science*, *8*(2), 101–105.

Schlichting, M. L., Guarino, K. F., Shapiro, A. C., Turk-Browne, N. B., & Preston, A. R. (2017). Hippocampal structure predicts statistical learning and associative inference abilities during development. *Journal of Cognitive Neuroscience*, *29*(1), 37–51.

Shafto, C., Conway, C. M., Field, S. L., & Houston, D. M. (2012). Visual sequence learning in infancy: Domain-general and domain-specific associations with language. *Infancy*, *17*, 247–271.

Shufaniya, A., & Arnon, I. (2018). Statistical learning is not age-invariant during childhood: Performance improves with age across modality. *Cognitive Science*, *42*, 3100–3115.

Siegelman, N., Bogaerts, L., Christiansen, M. H., & Frost, R. (2017). Towards a theory of individual differences in statistical learning. *Transactions of the Royal Society B*, *372*(1711), 20160059.

Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J., & Frost, R. (2018). Linguistic entrenchment: Prior knowledge impacts statistical learning performance. *Cognition*, *177*, 198–213.

Siegelman, N., Bogaerts, L., & Frost, R. (2017). Measuring individual differences in statistical learning: Current pitfalls and possible solutions. *Behavior Research Methods*, *49*(2), 418–432.

Siegelman, N., Bogaerts, L., Kronenfeld, O., & Frost, R. (2017). Redefining "learning" in statistical learning: What does an online measure reveal about the assimilation of visual regularities? *Cognitive Science*, *42*, 692–727.

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, *81*, 105–120.

Spencer, M., Kaschak, M. P., Jones, J. L., & Lonigan, C. J. (2015). Statistical learning is related to early literacy-related skills. *Reading and Writing*, *28*, 467–490.

Thomas, K. M., Hunt, R. H., Vizueta, N., Sommer, T., Durston, S., Yang, Y., & Worden, M. S. (2004). Evidence of developmental differences in implicit sequence learning: An fMRI study of children and adults. *Journal of Cognitive Neuroscience*, *16*(8), 1339–1351.

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

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*, 552–564.

von Koss Torkildsen, J., Arciuli, J., & Wie, O. B. (2019). Individual differences in statistical learning predict children's reading ability in a semi-transparent orthography. *Learning and Individual Differences*, *69*, 60–68.

Weiermann, B., & Meier, B. (2012). Incidental sequence learning across the lifespan. *Cognition*, *123*(3), 380–391.

West, G., Vadillo, M. A., Shanks, D. R., & Hulme, C. (2017). The procedural learning deficit hypothesis of language learning disorders: We see some problems. *Developmental Science*, *21*, e12552.

Zwart, F. S., Vissers, C. T. W. M., Kessels, R. P. C., & Maes, J. H. R. (2017). Procedural learning across the life span: A systematic review with implications for atypical development. *Journal of Neuropsychology*, *12*, 1–34.