

ARTICLE

# Individual differences in non-adjacent statistical dependency learning in infants

Jill LANY\* and Amber SHOAB

Department of Psychology, University of Notre Dame, USA

\*Corresponding author: Department of Psychology, Haggard Hall, Notre Dame, IN 46556.

E-mail: [jlany@nd.edu](mailto:jlany@nd.edu)

(Received 22 August 2018; revised 14 January 2019; accepted 11 April 2019;  
first published online 13 June 2019)

## Abstract

There is considerable controversy over the factors that shape infants' developing knowledge of grammar. Work with artificial languages suggests that infants' ability to track statistical regularities within the speech they hear could, in principle, support grammatical development. However, little work has tested whether infants' performance on laboratory tasks reflects factors that are relevant in real-world language learning. Here we tested whether the language that infants hear at home, and their receptive language skills, predict their performance on tasks assessing the ability to learn non-adjacent statistical dependencies (NADs) at 15 months, and whether that in turn predicts sensitivity to native-language NADs at 18 months. We found evidence for some (though not all) of these relations, and primarily for females. The results suggest that performance on the artificial language-learning task reveals something about the mechanisms of grammatical development, and that females and males may be learning NADs differently.

**Keywords:** infant-directed speech; statistical learning; non-adjacent dependencies

Infants do not learn their language solely by rote, simply coming to recognize common sounds and words as a function of frequent exposure to them. Rather, identifying sounds and words reflects the operation of sophisticated learning mechanisms, whereby infants become attuned to statistical regularities by listening to language. For example, perceiving phonemes (or sounds like 'b' and 'p') as native listeners do is only partially the result of neuroanatomical properties of the auditory system – it is also shaped by months of tracking frequency distributions, or how often the sounds along an acoustic continuum occur (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Maye, Weker, & Gerken, 2002). Likewise, the beginnings and endings of words are not inherently marked in naturally produced speech, and the perception of words as coherent units results, in part, from infants' sensitivity to sequential statistics, such as the likelihood that one syllable will precede or follow the next (Saffran, Aslin, & Newport, 1996; Pelucchi, Hay, & Saffran, 2009). The acoustics of infant-directed

© The Author(s), 2019. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

speech (IDS) can serve to amplify statistical structure in speech, for example by enhancing distributional information relevant to phoneme perception (Liu, Kuhl, & Tsao, 2003), and promoting the use of sequential statistics for word segmentation (Thiessen, Hill, & Saffran, 2005).

There is now substantial evidence that infants' ability to track statistical regularities within the speech stream is important for real-world phoneme perception and word segmentation (Kuhl *et al.*, 1992; Ngon *et al.*, 2013; Shoaib, Wang, Hay, & Lany, 2018), and for encoding speech in real time (Lany, Shoaib, Thompson, & Estes, 2018). However, the extent to which statistical learning is relevant to other aspects of language development, such as learning grammar, is perhaps more controversial. Some theories suggest that statistical regularities in the speech stream are insufficient for discovering abstract grammatical structures, and that innate knowledge of some core grammatical principles is necessary to get learners off the ground (e.g., Pinker, 1989; Yang, 2004). However, close characterization of child-directed speech suggests that it contains statistical regularities that could aid the learning of some aspects of grammar (e.g., Cartwright & Brent, 1997; Mintz, Newport, & Bever, 2002; Monaghan, Christiansen, & Chater, 2007; Christiansen, Onnis, & Hockema, 2009). Furthermore, infants can use statistical regularities in artificial languages to learn simple patterns analogous to aspects of grammar in natural languages, such as sequential structure among words and word categories (e.g., Gómez & Gerken, 1999; Gómez, 2002; Saffran & Wilson, 2003; Gómez & Lakusta, 2004; Lany, 2014; Lany & Saffran, 2010).

In sum, there is evidence that statistical learning is, in principle, potentially relevant to learning grammar. However, the extent of the role it truly plays is still debated. For example, some researchers suggest that statistics play only a small, supporting role in learning grammar (Lidz & Gleitman, 2004, Yang, 2004). If statistical learning does substantively contribute to this aspect of language development, then we should be able to find evidence that infants' performance on statistical learning tasks is related to the language they hear, and to their native language development. However, there have been few attempts to assess these potential connections, and we aimed to address this critical gap in the current work.

Here we focused on infants' ability to learn 'remote' or non-adjacent co-occurrence relationships (also called 'non-adjacent dependencies', or NADs), because they are pervasive in natural languages but can nonetheless be very hard to learn (e.g., Gómez, 2002, Newport & Aslin, 2004; Wang & Mintz, 2018). An example of a such a relationship in English is the co-occurrence of the auxiliary *is* and the inflectional morpheme *-ing* that occurs on verbs, as in "She *is playing*". Note that the verb separates them, which is why this dependency is considered to be non-adjacent. Likewise, *can* predicts the occurrence of verb-pronoun combinations, and "can [verb] it" is one of the 'frequent frames' that bookend verbs in speech to infants (Mintz, 2003). We first describe what is known about when and how infants learn these dependencies, and then describe the current study, in which we tested key predictors of both NAD learning and sensitivity to native-language NADs using a longitudinal design.

### *When do infants learn non-adjacent dependencies?*

Over the past twenty years, researchers have made progress in describing the developmental trajectory of infants' sensitivity to NADs in their native language. For example, at 18 months, but not 15 months, English-learning infants distinguish between sentences containing grammatical NADs involving auxiliaries and

inflectional morphology (e.g., “Everybody *is* often baking bread”) and closely matched ungrammatical sentences (e.g., “Everybody *can* often baking bread”) (Santelmann & Jusczyk, 1998). By 19 months, infants learning German are sensitive to similar NADs involving tense marking (Höhle, Schmitz, Santelmann, & Weissenborn, 2006). Infants learning English are sensitive to NADs involving subject–verb agreement by 16 months, and infants learning French show evidence of learning similar dependencies by 14 to 18 months (van Heugten & Shi, 2010; Nazzi, Barrière, Goyet, Kresh, & Legendre, 2011; Culbertson, Kouluaguina, Gonzalez-Gómez, Legendre, & Nazzi, 2016). Dutch-learning infants appear to become sensitive to dependencies between determiner forms and plural and diminutive morphology on nouns sometime between 17 and 24 months of age. Thus, there is converging evidence that infants are beginning to track at least some NADs in their native language by sometime late in their second year.

It is important to note that the term NAD is a general one, and can be used to describe a range of different underlying grammatical structures (e.g., subject–verb agreement vs. determiner–noun agreement). While these studies suggest that infants are sensitive to something about these NADs by about mid-way through the second year, they do not tell us what infants actually know about that relation. This is in part because most studies of NAD learning have used auditory discrimination paradigms, which assess whether infants listen differentially to strings containing grammatical instantiations of these dependencies vs. closely matched strings in which the dependencies are violated. To listen differentially, at minimum infants must have begun to track the co-occurrence of the surface forms involved in the NADs. In fact, it is likely that infants’ initial sensitivity is largely perceptually based, such that they first begin track the surface level co-occurrence of morphemes, and only later become sensitive to the more abstract nature of the dependency, including what the relation signals. Evidence for this suggestion comes from findings that French-learning infants discriminate between grammatical and ungrammatical strings involving subject–verb agreement by 14 to 15 months (Nazzi *et al.*, 2011; Culbertson *et al.*, 2016), but do not appear to understand what this dependency signals in comprehension tasks until closer to 30 months (Legendre, Barrière, Goyet, & Nazzi, 2010). Building on these findings, Culbertson *et al.* (2016) provided evidence that infants’ early sensitivity to NADs is likely to be perceptually driven in a study with French-learning infants between 14 and 24 months of age. Even the youngest infants showed evidence of discrimination between strings containing grammatical vs. ungrammatical dependencies involving subject–verb agreement morphology. Critically, the direction of preference for grammatical vs. ungrammatical strings shifted systematically with age. Culbertson *et al.* suggest that these shifts reflect changes in the strength and nature of infants’ knowledge about the dependencies, such that they initially represent them in terms of a co-occurrence relation between the surface forms, and between 18 and 24 months they begin to represent them in terms of their abstract morphosyntactic structure. Note that, in the current work, we focus on infants who are up to 18 months of age, at which time they are likely to be developing sensitivity to the surface-level statistical co-occurrence relations. Thus, in the current work, we use the term NAD to refer to these co-occurrence relationships rather than to the more abstract underlying relations.

In addition to broad developmental trends in the emergence of sensitivity to native-language NADs, there appear to be individual differences in when infants learn them. For example, Santelmann and Jusczyk (1998) found that 18-month-olds who were producing multi-word utterances were more likely to have learned

native-language NADs. Culbertson *et al.* (2016) reported that advances in sensitivity to native-language NADs were related to infants' age and to their productive vocabulary size. In both cases, these analyses were exploratory, but they suggest that NAD learning may share some underpinnings with word learning, and are consistent with evidence that lexical development in the second year strongly predicts later developments in learning grammar (Bates, Bretherton, & Snyder, 1991).

There is also evidence that females sometimes have an advantage over males in tasks assessing the ability to learn novel NADs. For example, in an evoked response potential (ERP) study Mueller, Friederici, and Männel (2012) found sex differences in NAD-learning at 3 months, with females displaying a more mature pattern of discrimination than males. At both 12 and 24 months, females have shown an advantage over males in generalizing sensitivity to newly learned NADs beyond trained instances (Lany & Gómez, 2008; Willits, Saffran, & Lany, 2017). It is not clear why females had an advantage in these NAD-learning tasks, but an explanation may be related to findings that sex hormone levels prenatally and early in infancy are linked both to neural structures important for language processing (Witte, Savli, Holik, Kasper, & Lanzenberger, 2010) and to phonological and syntactic processing ability (Friederici *et al.*, 2008).

### *How do infants learn non-adjacent dependencies?*

Beyond identifying *WHEN* infants learn NADs, researchers have also made progress in discovering the *MECHANISMS* by which infants learn them. Accumulating evidence suggests that statistical regularities play a role in forming initial, perceptually based representations of NADs. For example, although the co-occurrence relationship between *is* and *-ing* is highly frequent, it is probabilistic: Both *is* and *-ing* regularly occur outside of this dependency (i.e., “She is tired” and “Changing time!”). Infants often learn dependencies that are more frequent and reliable before learning those that are infrequent and more stochastic (van Heugten & Johnson, 2010), suggesting that these statistics do in fact matter for learning.

However, even when NADs are highly frequent and reliable, they can still be difficult to learn (Gómez, 2002). Critically, there is evidence that a different statistical property of the dependencies, namely a contrast between relatively variable and invariant structure, supports learning. For example, the highly frequent functional morphemes *is* and *-ing* bookend many different verbs. This frequency asymmetry between function words and open-class words (in this case, verbs) results in low reliability of adjacent relations relative to non-adjacent ones. Work with artificial languages suggests that this variability matters for learning NADs. Specifically, by 15 months of age, infants exposed to dependencies of the form *pel X rud* and *vot X jic* (where *X* denotes any one of a set of disyllabic words) learn that *pel* predicts *rud* but not *jic* when a large number of unique *Xs* can occur (i.e., *pel X<sub>1</sub> rud*, *pel X<sub>2</sub> rud*, *pel X<sub>3</sub> rud* ... *pel X<sub>24</sub> rud*), but not when 12 or fewer unique *Xs* appear across the corpus (Gómez & Maye, 2005). Such findings suggest that the presence of highly variable adjacent dependencies can highlight relatively invariant non-adjacent ones, and thereby facilitate NAD learning.<sup>1</sup>

<sup>1</sup>There is evidence from studies using ERP measures, rather than behavioral discrimination paradigms, that infants may even be able to track NADs with high variability in intervening elements by 3–4 months of age (Friederici, Mueller, & Oberecker, 2011; Mueller *et al.*, 2012).

There is also evidence that the prosodic qualities of IDS may work in concert with statistical cues to facilitate learning NADs, just as they do in the case of phonetic tuning and word segmentation. For example, in artificial language materials NADs can be easier to learn when they are bracketed off from surrounding speech by pauses (Peña, Bonatti, Nespor, & Mehler, 2002; Wang & Mintz, 2018). In natural languages, NADs often bookend syntactic units, like noun phrases and verb phrases, and IDS contains relatively long pauses at the edges of these syntactic units (Fisher & Tokura, 1996). The exaggerated pitch contours characteristic of IDS also highlight the boundaries between such units (Fisher & Tokura, 1996). Moreover, infants can use the convergence of pauses and prosodic cues to segment speech into well-formed phrases by 9 months of age (Jusczyk *et al.*, 1992), and infants who are better able to detect pitch changes across a series of syllables are more likely to learn NADs (Mueller *et al.*, 2012). These findings suggest that NADs may be especially salient and easy to learn in IDS because of its exaggerated prosody.

The evidence from the studies that we have just reviewed suggests that the language infants hear, and their ability to track statistics within it, plays an important role in learning NADs, at least initially, as they build perceptual sensitivity to co-occurring morphemes. It also suggests that infants who are more advanced in their lexical development, and who are also female, should be better able to learn NADs. However, most existing studies have included either assessments of NAD learning or analyses of speech to infants (see van Heugten & Johnson, 2010, for an exception), and thus there is little evidence linking these factors to individual infants' NAD learning abilities. In the current work, we used a longitudinal design to address this gap. We tested three specific predictions: First, given that the exaggerated prosodic contours in IDS highlight the boundaries between grammatical units (Fisher & Tokura, 1996), we predicted that hearing more IDS at home would promote infants' ability to learn NADs in an artificial language and in their native language. Second, given evidence of concurrent relations between lexical development and sensitivity to native-language NADs, we predicted that infants with larger vocabularies would be better able to learn novel and native-language NADs. Third, we predicted that NAD-learning ability would be related to emerging sensitivity to native-language NADs. Finally, we also considered the possibility that females might be better at learning NADs, and thus that some of the effects might emerge more strongly in females than in males, given prior evidence of sex differences (Lany & Gómez, 2008; Mueller *et al.*, 2012; Willits *et al.*, 2017).

## Methods

### Participants

We tested 56 English-learning infants from a medium-sized Midwestern city, beginning at 12 months of age. Infants who were being raised bilingual, or who had major health or cognitive issues according to parental report, were not considered eligible, and therefore not enrolled. We also only enrolled infants who were primarily at home during the day, rather than in daycare, and instructed parents to record on typical days while avoiding noisy or potentially chaotic events such as birthday parties. We excluded data from six infants who were enrolled but developed more than five ear infections, had tubes put in, or began receiving speech or language therapy over the

course of the study. We also excluded four infants who did not have usable data for both of the NAD tasks. There were 46 remaining infants (21 females and 25 males). Of those infants, 45 contributed to a measure of IDS, 34 contributed to the 15-month NAD-learning measure, and 30 contributed to the 18-month native-language NAD measure, with the remaining infants failing to contribute because of fussiness and inattention. The attrition rate in the NAD tasks is comparable to other similar studies (e.g., Culbertson *et al.*, 2016). Infants were all reported to be Caucasian, and the average level of maternal education was a college degree. Infants were an average of 12.3 months (range: 11.8–13.2) when we recorded the home language samples, 15.4 months (range: 14.7–16.1) when we assessed their NAD learning ability, and 18.5 months (range: 17.8–19.5) when we assessed their sensitivity to native-language NADs.

### Materials and procedures

Figure 1 depicts the measures collected at each of the three time-points; 12, 15, and 18 months of age.

#### Language samples

An experimenter visited infants' homes when they were approximately 12 months of age to deliver a LENA recorder and provide instructions on its use. The LENA is a digital recorder held in the pocket of clothing worn by the infants that captures speech in the immediate environment. We instructed parents to record for 8 hours a day on two typical days, and all the recorders contained 16 hours of data upon their return. We obtained the language samples at 12 months because it is standard in studies examining connections between language input and language development to obtain language samples at a time prior to conducting key language assessments (see Weisleder & Fernald, 2013; Ramirez-Esparza, Garcia-Sierra, & Kuhl, 2014, for two recent examples).

The LENA system is designed to distinguish between adult and infant utterances, and to filter out electronic speech (Xu, Yapanel, & Gray, 2009). After filtering, the software generates an estimate of the number of adult words spoken in the infant's presence. We used this word count as a starting point for obtaining a measure of how much of the speech to infants occurred in the IDS register. We adapted the methods of Ramirez-Esparza *et al.* (2014) to identify and code segments of the LENA recordings that contained IDS. First, we used the LENA software to identify

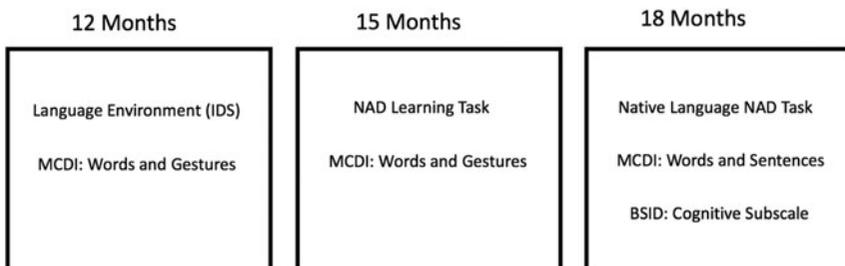


Figure 1. Measures at the three time-points.

80 5-minute segments (40 per day) with word counts of at least 140.<sup>2</sup> Coders listened to each segment to select the first 30-second clip that contained speech to the infant for further coding. If there were not 80 clips with at least 140 words, we chose a different interval from the 5-minute segments with the highest adult word counts (in other words, double-sampling from some 5-minute clips). When there was not enough speech to code different 30-second segments with those clips, we selected clips from the segments with the next-highest word counts to obtain 80 segments.

In a separate pass, trained coders indicated whether the speech to the infant in each 30-second clip was in the infant-directed (ID) register. This speech had to be from an adult, but was not restricted to any specific caregiver (i.e., it could be from the infant's mother, father, grandparent, or another adult). We used these codes to compute the proportion of clips in which the speech to the infants was in the ID register. To measure the reliability of our coding, we re-coded 15% of participants' samples. Coders agreed on the classifications of 94% of the segments.

### *NAD measures*

*15-month NAD-learning task.* At 15 months we tested infants' ability to learn novel NADs in an artificial language using the Head-Turn Preference Procedure, or HTPP (Gómez, 2002; Gómez & Maye, 2005). Infants were seated on a parent's lap in a small sound-attenuated booth. A large LCD screen was mounted on the wall in front of them, and two smaller screens were mounted on the side walls, with a speaker below each side screen. An experimenter administered the tasks from outside the room via a PC computer running custom software. The experimenter was blind to the experimental stimuli and monitored infants' visual attention via the digital video feed.

Infants were first familiarized with strings of the form *pel X rud* and *vot X jic* (Grammar 1), or *pel X jic* and *vot X rud* (Grammar 2). There were 24 unique Xs yielding 48 strings in each grammar (see 'Appendix'). A female speaker recorded the materials from Gómez (2002), and the best tokens of each word were spliced together to form the strings, with silences of 0.15s between the words within the strings, and 0.75s between the strings. The same tokens were used in both G1 and G2 strings to prevent idiosyncratic differences from driving discrimination at test. Infants were familiarized to the full set of 48 strings while seated on a parent's lap. During this time the auditory stream played continuously, while the images on the front and side screens were presented contingent on the infants' looking (see the description of the test phase immediately below for further details).

After they were familiarized to the NADs, infants' learning was tested using the HTPP. On each test trial four strings from either G1 or G2 were presented (see 'Appendix'). The same four strings from each grammar were repeated in different orders to create two unique trials for Grammar 1 and for Grammar 2. The set of four trials was repeated in two additional blocks for a total of 12 test trials (see Gómez, 2002, for details). Thus, for all infants the test consisted of four G1 trials and four G2 trials. The order of G1 and G2 trials within blocks was randomized.

The experimenter used the custom software to control the presentation of auditory stimuli and record infants' listening times. Each test trial began with a video of a red flashing light displayed on a monitor located in front of the infant. Once the infant focused on the screen, a video of a yellow flashing light appeared on a

<sup>2</sup>Because speech in the adult-directed register is typically more rapid than IDS, we did not strictly select the highest word count segments to avoid creating samples biased to over-represent adult-directed speech.

monitor 90 degrees to the right or left, and the center video disappeared. When the infant turned his or her head at least 30 degrees toward the side monitor, the experimenter signaled the computer to present an auditory stimulus. The auditory stimulus continued to play until the infant looked away for at least 2 consecutive seconds, or until 35 seconds had elapsed.

*18-month NAD task.* At 18 months we tested infants' sensitivity to native-language NADs. The goal of this task was to assess what infants had learned about these NADs from their language experience outside of the lab, and thus we did not include a familiarization phase. Infants were simply tested on their ability to discriminate between English sentences that contained grammatical and ungrammatical NADs using the HTPP, as described above, with the exception that the visual stimulus presented on the side monitors was a GIF of a train engine chugging, which we anticipated would be more engaging for older infants. To train infants on the contingency between their looking behavior and the presentation of auditory stimuli, the test phase began with two trials that consisted of music rather than the NADs.

In their seminal work, Santelmann and Jusczyk (1998) used a single grammatical dependency and its ungrammatical counterpart (the grammatical *is/-ing* and the ungrammatical *can/-ing*). Likewise, in the current experiment, each infant also heard one kind of grammatical construction and one kind of ungrammatical construction throughout the test. The grammatical dependencies took the form "She is Xing" or "She can X it". This served as a between-participant manipulation, with some infants hearing only the former on all grammatical trials, and some hearing the latter. In both conditions, the grammatical constructions were pitted against ungrammatical sentences of the form "She can Xing". In pitting the ungrammatical *can/-ing* against *is/-ing* in one materials set, we matched the contrast used by Santelmann and Jusczyk. We chose to also pit the ungrammatical *can/-ing* strings against grammatical strings containing *can/it* for two reasons. First, we wanted to ensure that any observed preferences were not driven by a preference for *is* over *can*, or vice versa. Second, using a second grammatical dependency allows us to increase the generalizability of our results beyond the *is/-ing* dependency.

Note, however, that the co-occurrence of *can/it* does not arise from the same kind of dependency as *is/-ing*: *Is* and *-ing* co-occur because both morphemes are required to mark the present progressive tense. In contrast, *can/it* reflects the co-occurrence of the modal (*can*) and a pronoun (*it*) in transitive constructions, and it is not obligatory that transitive constructions use these particular forms. Thus, the co-occurrence reflects a grammatical relation, but one of a different nature than the one underlying the co-occurrence of *is/-ing*. *Can/it* is also less frequent than *is/-ing*, at least in adult-directed speech. Nonetheless, both *is/-ing* and *can/it* involve relatively frequent, prosodically reduced functional morphemes that bookend verbs, and thus there are important similarities in their perceptual instantiation. Because there is little evidence that 18-month-olds represent these dependencies in ways that go beyond the perceptual/surface-level, we suggest that it is reasonable to use *can/it* alongside *is/-ing* to help ensure that if infants listen differentially to the ungrammatical *can/-ing* construction, it is not simply because they respond differently to strings containing *is* vs. *can*.<sup>3</sup>

<sup>3</sup>Using *is/it* constructions in ungrammatical strings (e.g., "She is pliz it.") with novel verbs could be perceived as a grammatical utterance, (e.g., "She is plizit."), and thus *can/-ing* was the only ungrammatical construction we used.

To ensure that this task tapped sensitivity to the dependencies themselves, rather than to familiar vs. unfamiliar combinations of adjacent words, we created a set of 20 nonsense CVC verbs that served as the X-items (see 'Appendix'), similar to the approach taken by van Heugten and Johnson (2010). To make certain that infants could not discriminate between the grammatical and ungrammatical strings based on prosodic cues resulting from the disfluencies that native speakers can experience when intentionally producing ungrammatical utterances, we selected productions of "She can/She is" from grammatical utterances, and spliced them together with productions of "X it" and "Xing" also taken from grammatical productions. On each trial, a grammatical or ungrammatical dependency was presented with 10 of the nonsense verbs in random order. Infants were presented with four Grammatical and four Ungrammatical test trials. The test was administered using the HTPP, as described above. In this task, the maximum trial length was 18 seconds. To familiarize infants with the contingency between their looking behavior and the presentation of test trials, they were presented with two 'practice' trials in which they heard classical music, contingent on their looking behavior, before the test phase began.

In both the 15- and 18-month tasks, infants' sensitivity to non-adjacent structure was assessed using their listening behavior at test. For each task, we created a preference score by computing the average time spent listening to grammatical trials and dividing it by average listening summed across both grammatical and ungrammatical trials. In the HTPP, learning can be evidenced by preferences for familiar (i.e., grammatical) or novel (i.e., ungrammatical) strings. Here we were most interested in predicting individual differences in these preference scores. Given that, in previous work with a very similar NAD-learning task, 15-month-olds listened longer to grammatical over ungrammatical test trials, we predicted that better NAD learning would manifest in stronger preferences for grammatical strings. Because the stimuli we used to test sensitivity to native-language NADs had not been used previously, we did not predict a direction of preference that would be associated with better learning a priori.

#### *Standardized measures of language and cognitive development*

At all three ages, we assessed native language development using the MacArthur-Bates Communicative Development Inventory (MCDI) (the Words and Gestures version at 12 and 15 months, and the Words and Sentences version at 18 months). The MCDI is a commonly used parent-report assessment of multiple aspects of early language development, and here we focused on the measures of infants' receptive and expressive vocabulary size. For these measures, parents indicate on a checklist the specific words that they think their infant understands and/or says. Both raw counts of known words and percentile scores (which fall on a scale of 0–100) are available, with percentile scores computed by standardizing the raw counts based on infants' age and sex. We used the percentile scores because preliminary analyses revealed a small number of outliers on the raw vocabulary size scores but not their corresponding percentile scores, and because it can facilitate interpreting sex differences, as we discuss later.

At 18 months we also administered the cognitive subscale of the Bayley Scales of Infant Development (BSID; Bayley, 2006). The BSID involves a progression of calibrated tasks administered by a trained experimenter using a standard set of toys and objects. The cognitive subscale assesses infants' attention and memory, and their speed in completing tasks. Infants' performance is used to derive a normed score (out of 100). Both the MCDI and BSID have high reliability and validity, and are

widely used instruments for assessing early language and cognitive development, respectively. We included the standardized measure of cognitive development to test whether any of the relations we predicted between NAD learning and lexical development were selective, reflecting language-learning mechanisms in particular, vs. more general aspects of infants' development. If performance on the NAD-learning task taps native language development, it should be more strongly related to MCDI scores, which tap language development, than to BSID scores, which tap more general cognitive development.

## Results

**Table 1** contains the descriptive statistics for key measures reported separately for males and females, as well as the results of *t*-tests comparing males and females on these measures. With the exception of expressive vocabulary size at 12 months of age, which was not a key measure in these analyses, there were no significant differences between males and females on any of the measures, and the ranges of obtained values were also quite similar. Note that the MCDI scores were the normed scores, which adjust for sex (as females typically lead males), so we would not expect sex differences, even if raw vocabulary size does differ.

We first report group-level performance on the two NAD tasks, which are the key measures we wanted to predict. The measure of sensitivity to NADs was a preference score, in which we divided their mean listening time to grammatical trials by total mean listening time to grammatical and ungrammatical trials. A score of .5 represents equal listening to the two kind of trials. There was a substantial range in performance on both NAD tasks, with preference scores ranging from .27 to .68 on the NAD-learning task at 15 months, and from .30 to .74 on the native-language NAD task at 18 month and SDs of 0.10 in both tasks. However, one-sample *t*-tests revealed that average performance across the entire group did not differ from .5 (or chance) on the 15-month NAD-learning task ( $M = .51$ ,  $SE = .016$ ,  $t(33) = 0.66$ ,  $p = .517$ ),<sup>4</sup> or on the 18-months native-language NAD task ( $M = .50$ ,  $SE = .020$ ,  $t(29) = 0.20$ ,  $p = .845$ ). Females and males did not differ from each other on either task (independent samples  $t(32) = 1.43$ ,  $p = .162$  for the NAD-learning task at 15 months, and  $t(28) = 1.11$ ,  $p = .649$  for the native-language NAD task at 18 months; see **Table 1** for means).

We next tested the predicted relations between hearing IDS, lexical development, and performance on the NAD tasks with moderation analyses performed using the PROCESS macro for SPSS (Preacher & Hayes, 2004.) Rather than excluding infants with missing data from all analyses, we only excluded them for analyses that included that measure. As a result, the sample sizes across the moderation analyses differed. We chose to include sex as a moderator in these analyses based on our initial inspection of the data, which involved plotting the results of the predicted correlations separately for males and females. We did so even though the *t*-tests described above did not reveal difference in females' and males' average performance, because we reasoned that the relations might nonetheless emerge most strongly in females if males struggled in the NAD tasks. We first describe the key results of these analyses, which converged to suggest that sex does moderate at least some of

<sup>4</sup>We address the difference between our findings and those from prior studies in the 'Discussion'.

**Table 1.** Descriptive statistics

12 months	Females			Males			<i>t</i> -test
	<i>M</i>	<i>SD</i>	Range	<i>M</i>	<i>SD</i>	Range	<i>t</i> (df)
LENA words	18916.43	7298.41	7692–29614	19969.84	7827.08	7036–39881	−0.47 (44)
IDS	0.94	0.09	0.66–1.0	0.93	0.13	0.53–1	0.13 (40)
Receptive vocabulary	34.87	25.68	1–91	39.86	33.42	1–99	−0.56 (44)
Expressive vocabulary	45.87	17.83	20–87	56.60	18.37	25–88	−2.00 (44)*
<b>15 months</b>							
Receptive vocabulary	32.94	23.98	1–94	47.27	35.48	1–99	−1.54 (42)
Expressive vocabulary	39.12	19.87	10–80	45.37	22.83	5–83	−0.96 (42)
Novel NADs	0.53	0.10	0.27–0.68	0.49	0.08	0.35–0.63	1.43 (32)
<b>18 months</b>							
Expressive vocabulary	34.94	23.04	6–76	38.81	25.58	1–83	−0.52 (42)
BSID	44.68	29.4	1–99	45.18	20.14	16–95	−0.06 (33)
Native language NADs	0.49	0.10	0.35–0.62	0.51	0.12	0.30–0.75	−0.46 (28)

Notes. The table reports means and variances of the key variables separately for females and males, and also the results of *t*-tests comparing female and male scores on those variables. All values reported here were computed over all of the infants contributing usable data to that measure. The IDS score reflects the proportion of speech to infants in the IDS register; the receptive and expressive vocabulary measures are the normed MCDI scores. Preference scores for the NAD measures at 15 and 18 months reflect the average time spent listening to grammatical trials and dividing it by average listening summed across both grammatical and ungrammatical trials, such that larger numbers reflect longer listening to grammatical test trials. At 15 months the task assessed NAD learning in an artificial language, and at 18 months it assessed sensitivity to native language NADS. The BSID measure reflects the percentile score on the cognitive subscale of the Bayley. \* = *p* = .05.

the relations, and then describe the more specific findings for females and males separately.

In the first analysis (on a sample size of 31), IDS was entered as an independent variable, sex was entered as a potential moderator, and performance on the 15-month NAD-learning task served as the dependent variable. There was a significant interaction between sex and IDS in predicting NAD learning ( $B = -1.39$ ,  $t = -2.70$ ,  $p = .012$ ; 95% CI [-2.45, -0.33]). This result suggests that the relation between IDS and performance on the NAD-learning task differed for females and males. In the second analysis (sample size of 34), performance on the 15-month NAD task again served as the dependent variable, receptive vocabulary size at 12 months served as the independent variable, and sex as a potential moderator. Here we also found a significant interaction between sex and receptive vocabulary size ( $B = -0.0029$ ,  $t = -2.56$ ,  $p = .016$ ; 95% CI [-0.0052, -0.0006]), suggesting that the number of words understood, like IDS, was related to the NAD-learning task differently in females and males.

In the next three moderations, performance on the 18-month NAD task served as the dependent variable. When IDS was entered as the independent variable, and sex entered as a moderator, we found no evidence of an interaction ( $B = -0.40$ ,  $t = -0.59$ ,  $p = .560$ ; 95% CI [-1.80, 1.00]; sample size of 27). When receptive vocabulary size at 12 months was the independent variable, and sex the moderator, there was a non-significant interaction between them ( $B = 0.003$ ,  $t = 1.96$ ,  $p = .061$ ; 95% CI [-0.0001, 0.0060]; sample size 30). We note that the relation trended towards significance, and that it is consistent with the findings that receptive vocabulary size and sex interacted in predicting NAD learning at 15 months. However, when performance on the 15-month NAD task was the independent variable, and sex was entered as a potential moderator, we found a much weaker non-significant interaction between infant sex and performance on the 15-month NAD task in predicting performance on the 18-month NAD task ( $B = 0.846$ ,  $t = 1.35$ ,  $p = .20$ ; 95% CI [-0.50, 2.19]; sample size 18).

Altogether, the results from the moderation analyses suggest that language input and receptive language skills interact with sex to predict NAD learning at 15 months, and that some similar, but weaker, patterns may be present at 18 months. Given that the males and females did not differ on any of the predictor measures, including exposure to IDS and vocabulary size, these findings do not appear to reflect simple sex differences on any of these key measures, but rather suggest that there may be a different pattern of relations between the key variables for males and females. We thus consider these relations separately for females and males more carefully below.

### *NAD learning in females*

We first tested the factors that were associated with females' performance on the two NAD tasks. Table 2 contains pair-wise correlations between language input, NAD measures, and MCDI scores at both ages, as well as sample sizes. We focused on receptive vocabulary at 12 months in our predictions, but include the 15- and 18-months measures in the table. We report the results of the key correlations below, with 95% CIs of the  $r$  statistic.

Focusing first on the NAD-learning task at 15 months, we found that the amount of IDS that infants heard at 12 months was related to their preferences for grammatical strings ( $r = 0.65$ ,  $p = .012$ , 95% CI [0.18, 0.88]). This relation, in which infants who

Table 2. Correlations

	Females		Males	
	15 month AGL NAD	18 month native-language NAD	15 month AGL NAD	18 month native-language NAD
<b>12 months</b>				
LENA words	-0.10 (16)	-0.21 (15)	0.04 (18)	0.001 (15)
IDS	0.65* (14)	-0.39 (14)	0.13 (17)	-0.32 (13)
Receptive vocabulary	0.57* (16)	-0.65** (15)	-0.20 (18)	0.06 (15)
Expressive vocabulary	0.26 (16)	-0.44 (15)	-0.23 (18)	0.38 (15)
<b>15 months</b>				
Receptive vocabulary	0.45^ (15)	-0.43 (14)	-0.34 (17)	0.24 (14)
Expressive vocabulary	0.34 (15)	-0.24 (14)	-0.68** (17)	0.26 (14)
<b>18 months</b>				
Expressive vocabulary	0.38 (15)	-0.58* (15)	-0.49* (17)	0.20 (14)
BSID	0.24 (13)	-0.37 (15)	0.39 (10)	0.07 (14)
Native-language NAD	-0.64* (10)		0.18 (8)	

Notes. In this correlation table, which reports  $r$  values, the IDS score reflects the proportion of speech to infants in the IDS register, the receptive and expressive vocabulary measures are the normed MCDI scores. The BSID measure reflects the percentile score on the cognitive subscale of the Bayley. The sample size for each analysis appears in parenthesis next to the correlation coefficient. ^  $p < .1$ , \*  $p < .05$ , \*\*  $p < .01$ .

heard more IDS showed stronger preferences for the grammatical strings, is depicted in Figure 2. Note that females' performance on the NAD-learning task at 15 months was not related to the total amount of speech present in their environment (see the LENA measure in Table 2), suggesting that IDS, rather than speech in general, is relevant. Females with larger receptive vocabularies at 12 months also exhibited stronger preferences for grammatical strings at 15 months ( $r = 0.57$ ,  $p = .022$ , 95% CI [0.10, 0.83]) (see also Figure 3). Together, these findings suggest that stronger preferences for grammatical strings on this task reflect better NAD learning in females at this age.

Turning to performance on the native-language NAD task at 18 months, we found that females who displayed stronger preferences for grammatical strings on the NAD-learning task at 15 months showed stronger preferences for ungrammatical strings in the native-language NAD task at 18 months ( $r = -0.64$ ,  $p = .0045$ , 95% CI [-0.91, -0.02]) (Figure 4). Females with larger receptive vocabularies at 12 months also showed stronger preferences for ungrammatical strings in the 18-month NAD

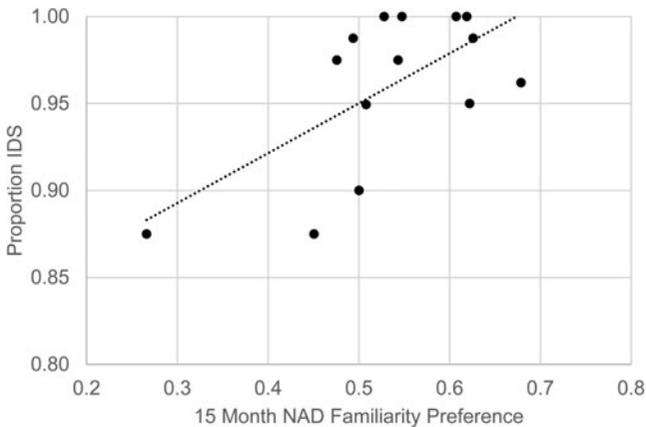


Figure 2. Relation between females' performance in the 15-month NAD task and their language input.

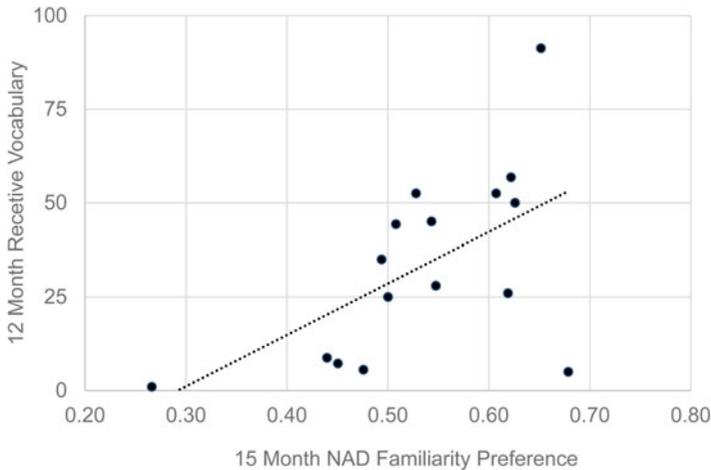


Figure 3. Relation between females' performance on the NAD task at 15 months and their receptive language at 12 months.

task ( $r = -0.65, p = .009, 95\% \text{ CI } [-0.87, -0.21]$ ) (see Figure 5), suggesting that a stronger preference for ungrammatical strings reflects more robust learning. This pattern of results suggests that females who had larger receptive vocabularies at 12 months, and who were better able to learn the novel NADs at 15 months (expressed as a preference for grammatical strings), were more sensitive to native-language NADs (at this age expressed in terms of longer listening to ungrammatical strings). Females' sensitivity to native-language NADs was not related to the amount of IDS they heard. The size and direction of the correlation coefficient ( $r = -0.39, p = .167; 95\% \text{ CI } [-0.76, 0.18]$ ) are suggestive, however: Infants hearing more IDS showed larger preferences for ungrammatical strings, just like those with larger vocabularies and those who exhibited stronger preferences for grammatical strings on the NAD-learning task.

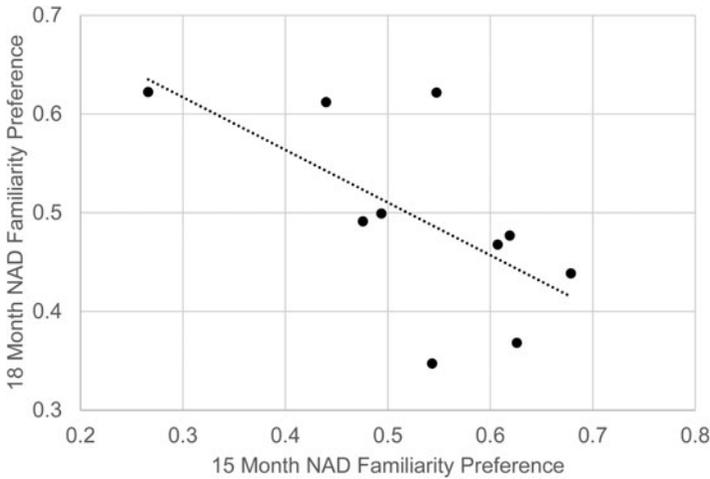


Figure 4. Relation between females' performance on the NAD tasks.

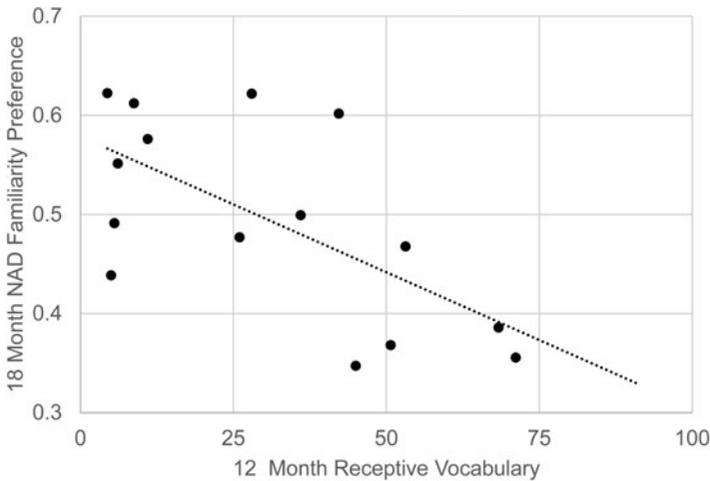
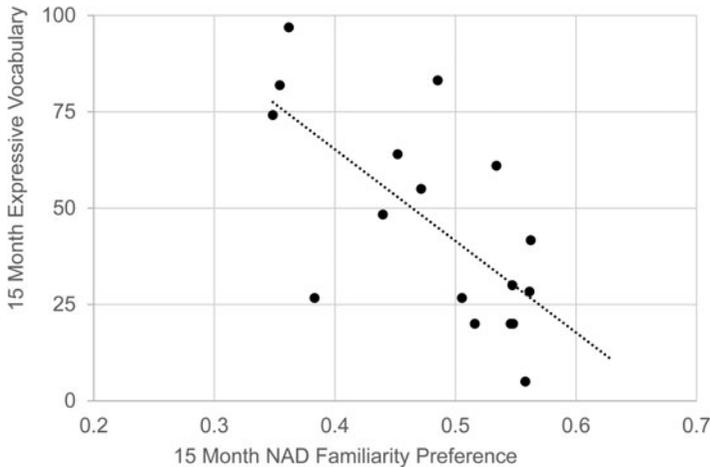


Figure 5. Relation between females' performance on the NAD task at 18 months and their receptive language at 12 months.

**NAD learning in males**

In contrast to the strong and relatively consistent relations that we observed in females, we found that males' performance on the two NAD tasks was not related to the amount of IDS they heard, or to their receptive vocabulary size at any age (Table 2). The association between NAD learning at 15 months and sensitivity to native-language NADs at 18 months was not significant. Likewise, the IDS measure was not related to sensitivity to native-language NADs in males ( $r = -0.32, p = .308, 95\% \text{ CI } [-0.84, 0.50]$ ).<sup>5</sup>

<sup>5</sup>Note that there was a non-significant trend towards a negative relation between IDS and performance on the 18-month NAD task for both males and females when these groups were considered separately. We thus tested whether there was a relation between IDS and performance on the 18-month NAD task when



**Figure 6.** Relation between males' performance on the NAD task at 15 months and their expressive language at 15 months.

While we did not find evidence for the predicted relations in males, there was some potential evidence that their performance on the NAD-learning task at 15 months was related to their lexical development: infants who exhibited stronger preferences for ungrammatical strings were saying more words at 15 months ( $r = -0.68$ ,  $p = .003$ , 95% CI  $[-0.88, -0.28]$ ) (see Figure 6). Their 15-month NAD performance was also related to expressive vocabulary at 18 months (see Table 2). Note that the associations involving males' 15-month NAD performance differed from those found in females, for whom stronger preferences for grammatical strings in the 15-month NAD-learning task were associated with earlier and concurrent measures of receptive vocabulary size measures.

### *NAD learning and the Bayley Scale*

The relations we observed between the NAD tasks and aspects of language input and lexical development suggest that the NAD tasks tap factors that are relevant to language learning. However, it may be that these relations, and particularly performance on the NAD tasks, also reflect more general learning mechanisms and cognitive processes. To address this possibility, we tested whether either of the NAD measures was related to a measure of general cognitive development, the scores on the cognitive subscale of the BSID. There were no significant relations between the BSID and either of the NAD tasks for males or for females (see Table 2). However, two relations approached significance. First, females who showed stronger novelty preferences on the native-language NAD task at 18 months were more likely to score highly on the BSID ( $r = -0.37$ ,  $p = .176$ , 95% CI  $[-0.74, -0.18]$ ). Stronger novelty preferences on this task were also predicted by larger receptive vocabularies at earlier ages, and thus this finding could suggest that sensitivity to native-language NADs is

---

collapsing across sex, but found no significant evidence of a relation ( $r(26) = -0.31$ ,  $p = .111$ , 95% CI  $[-0.62, 0.08]$ ).

related to more general cognitive development in females. For males, stronger familiarity preferences on the NAD learning task at 15 months were related to higher BSID scores, though the relation did not reach significance ( $r = 0.39$ ,  $p = .265$ , 95% CI  $[-0.32, 0.82]$ ). Note that, in males, stronger novelty preferences on the 15-month NAD task were related to larger expressive vocabulary size, which complicates interpretation of this effect.

### Materials effects

In our main analyses involving performance on the native-language NAD task, we collapsed over the two sets of materials (the *is/-ing* vs. *can/it* grammatical dependencies). An independent samples *t*-test revealed that there were no group-level differences in performance as a function of the specific dependencies ( $t(28) = 0.30$ ,  $p = .769$ ; the mean preference scores were  $.51$  ( $SE = .028$ ) and  $.50$  ( $SE = .032$ ), respectively for the *is/-ing* and *can/it* sets). It is possible, though, that effects we observed held differently within the two sets. Because sample sizes were already small, we did not want to further subdivide our sample to test whether the pattern of results held in both sets of materials. Instead, to better capture potential differences in performance as a function of the dependencies used, we re-computed the correlations using *z*-score standardized preference scores that were generated separately as a function of the dependency heard. These *z*-scores rescale the raw preference scores separately within each of the groups (i.e., independently creating a set of *z*-scores for infants who heard *is/-ing* and for those who heard *can/it*). If these groups represent different populations, using these *z*-scores results in measures that reflect variance specific to that group. Analyses using these standardized scores revealed the same patterns of significant findings as analyses using the unstandardized scores, suggesting that the results were robust to both materials sets. However, future work with larger samples will be important to directly test the possibility that there are differences in infants' sensitivity to dependencies involving *is/-ing* and *can/it*.

### Summary of findings

Summarizing our findings across males and females, we found evidence that 15-month-olds' ability to learn novel NADs in an artificial language is related to native language development. For females, there was a constellation of associations between the proportion of IDS in the speech addressed to them, their receptive vocabulary size, and their performance on both the NAD tasks. For males, however, the only relations that reached significance were between expressive vocabulary size and performance on the 15-month NAD-learning task.

### Discussion

In the current work, we used a longitudinal design to identify factors that predict infants' ability to learn non-adjacent dependencies. In particular, we tested whether infants' ability to learn novel NADs in an artificial language at 15 months, and their sensitivity to native-language NADs at 18 months, are related to their experience hearing IDS and to their receptive language skills at 12 months. We also tested whether NAD-learning scores predict sensitivity to native-language NADs. We found some evidence for the existence of these relations, though the patterns differed for males and females.

The strongest evidence for the predicted relations emerged in females, with results converging on a picture in which language input, receptive vocabulary size, and NAD learning are interconnected. Specifically, females' ability to learn novel NADs was predicted by how much IDS they heard, as well as by receptive vocabulary size. Their sensitivity to native-language NADs was predicted by their ability to learn novel NADs in an artificial language, and by their receptive vocabulary size. These results are consistent with evidence that early lexical and grammatical development are linked (Bates *et al.*, 1991) and that both are related to children's language environment (Hart & Risley, 1995; Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010). They are also consistent with evidence that infants are better able to learn statistical regularities in IDS vs. ADS (Thiessen *et al.*, 2005), that hearing more IDS from caregivers supports vocabulary development (Ramirez Esparza *et al.*, 2015), and that prosodic qualities of IDS may highlight grammatical structure (Fisher & Tokura, 1996; Christophe, Millotte, Bernal, & Lidz, 2008). Note, however, that hearing more IDS was not directly related to females' performance on the native-language NAD task. This may reflect the fact that more time had elapsed since obtaining the 12-month measure of language input at 18 vs. 15 months, and the fact that other factors, including aspects of language input that we did not capture, surely influence the emergence of sensitivity to native-language NADs.

While scores on the 15- and 18-month NAD tasks were related in females, and performance on both tasks was also related to their receptive language skills, neither NAD task was related to the cognitive subscale of the Bayley. There was a weak, non-significant relation between their Bayley scores and their performance on the NAD-learning task at 15 months. The relation between the Bayley and the native-language NAD task at 18 months was stronger but still non-significant. These findings suggest that the observed relations are not likely to have been driven by more general cognitive development, at least as captured by the Bayley, but future work will be necessary to identify additional factors that support NAD learning.

In contrast to the constellation of relations we found for females, the only relations we found in males were between NAD learning at 15 months and productive language development (recall that for females, performance on that task was strongly related to receptive rather than productive language skills). Sex differences have previously been reported in the literature on NAD learning. For example, at 3 months, females display a more mature pattern of discrimination than males in a NAD-learning task (Mueller *et al.*, 2012). At 12 months, females but not males detect novel non-adjacent dependencies when given scaffolding experience with adjacent ones (Lany & Gómez, 2008), and at 24 months, females are also more likely than males to generalize newly learned non-adjacent dependencies to novel instances (Willits *et al.*, 2017). Our results are also in line with evidence that males and females differ in other potentially related aspects of language development: Recent evidence suggests that sex hormone levels (prenatally and during early infancy) are related to several aspects of language development (e.g., Beech & Beauvois, 2006). Particularly relevant here are findings that testosterone and estradiol levels (which are linked to biological sex) predict phonological discrimination and sentence comprehension (Friederici *et al.*, 2008; Schaadt, Hesse, & Friederici, 2015), and have been linked to structural differences in left-hemisphere regions involved in language processing (Witte *et al.*, 2010).

It is possible that our results reflect a female advantage, and that similar relations would emerge for males at later ages, or if the NAD tasks we used were more

sensitive to weak learning. Note, though, that our results do not necessarily suggest that females learned NADs better than males. For example, we did not find gross differences between males and females on the NAD task at 15 months. Instead, females with larger receptive vocabularies tended to show a stronger preference for grammatical strings in the NAD-learning task, and males with larger expressive vocabularies tended to prefer ungrammatical ones. Stronger preferences for ungrammatical (or novel) over grammatical (or familiar) constructions could indicate stronger learning than preferences for grammatical constructions (Hunter & Ames, 1988). Indeed, Gómez and Maye (2005) reported that 15-month-olds showed preferences for grammatical strings, while 17-month-olds showed preferences for ungrammatical ones, consistent with the possibility that better learning, likely to be evidenced by the older infants, manifests in novelty preferences. This might be taken to suggest that males in our study learned the NADs better than the females (or at least that those with larger expressive vocabularies did). However, we are reluctant to draw that conclusion from these results, especially considering that the differences between males and females were not limited to the direction of preference that was more closely associated with larger native language vocabularies: there was stronger evidence that individual differences in females' performance on the NAD tasks were related to language development. Thus, we suggest that these findings are most consistent with an interpretation in which the manifestation of NAD learning, and the factors associated with it, may be different in males and females. Future work will be necessary to uncover the nature of the differences between males' and females' NAD learning, and the factors that contribute to these differences.

Overall, then, our findings suggest that infants' performance on the NAD-learning task may be related to their native language development, and that there are sex differences in such relations, but there are several caveats to these conclusions. First, our sample was small, and not all infants contributed usable data at all three time-points. This means that we are limited in drawing conclusions about change over time. In addition, while the results of multiple analyses lined up in suggesting that there are differences in the factors that predict NAD learning in females and males, and the results from analyses on females converged in largely predicted ways, we did not predict that different relations would hold for females vs. males. Rather, we predicted that females would outperform males, and therefore that some relations might emerge more strongly in females. Finally, our findings do not precisely replicate those of Gómez & Maye (2005),<sup>6</sup> who found evidence of group-level learning of novel NADs at 15 months. In fact, we did not find evidence of sensitivity to either the novel or native-language NADs at the group level. For all of these reasons, in future work it will be important to examine these relations in larger samples to replicate and assess the strength of the effects.

However, it is important to put these caveats into a broader context. First, we address whether it is necessary to observe group-level effects in the NAD tasks to pursue questions about individual differences in learning. We suggest that is not. It is conventional to investigate individual differences only when group-level performance

---

<sup>6</sup>There were some small differences between our materials and those of Gómez & Maye (2005). Likewise, the materials we used to test NAD learning at 18 months were only loosely based on those of Santelmann and Jusczyk (1998). Thus, our results should not be considered a strict failure to replicate theirs. At a general level it may be that we did not find evidence for group-level discrimination, as they did, perhaps because using nonsense words made the task harder.

is above chance, at least when using paradigms in which chance performance clearly reflects a lack of learning. Indeed, if most participants fail on a task, then it is unlikely that observed variance in performance reflects variance in the key construct. However, in the HTPP, chance performance is much more difficult to interpret, and deviation from it in either direction can be indicative of learning. Thus, we treat these findings as suggestive, and a starting point for additional studies that address this issue more directly.

Next, while we generally urge caution in interpreting our findings, given the limitations in the study, it is important to explicitly consider the findings that were strongest and most consistent within our sample vs. those that were weaker. Multiple measures suggested that there are sex differences in NAD learning at 15 months of age: sex interacted with the IDS and receptive vocabulary size measures in predicting performance on the NAD-learning task, and follow-up analyses revealed strong and consistent correlations between IDS, receptive vocabulary size, and 15-month NAD learning only in females. Thus, we are relatively confident that there were meaningful sex differences in NAD learning in our task. Males' performance on the NAD-learning task at 15 months was related to their productive vocabulary size at both 15 and 18 months. However, these were the only significant relations for males across all of the measures, including sensitivity to native-language NADs, and thus we are less confident of these findings.

Considering the results of analyses predicting infants' sensitivity to native-language NADs at 18 months, there was more modest evidence that sex interacts with the same set of predictors, or with 15-month NAD learning: only the interaction between sex and receptive vocabulary size approached significance. When considering the results from the females on their own, receptive vocabulary size at 12 months predicted native-language NAD knowledge, as did performance on the 15-month NAD-learning task. These results are consistent with the results of the 15-month analyses. However, somewhat surprisingly, sensitivity to native-language NADs was not related to the IDS measure. Thus, the pattern of findings for females at 18 months is somewhat less strong and consistent than at 15 months, and there were also fewer infants contributing data to the analyses. Combined with the weak results from the moderation analyses, we are more cautious about the conclusion that sex is a moderator of key predictors of sensitivity to native-language NADs.

If the sex differences in NAD learning described above are reliable, it will be important to determine how NAD learning differs in males and females. As noted above, it is possible that males and females recruit a different constellation of learning mechanisms when tracking NADs, perhaps due to developmental differences in their maturity, or to more enduring differences in brain structure and function (Witte *et al.*, 2010). Another possibility is suggested by evidence that the acoustic characteristics of IDS used with males and females growing up in Australian-English-speaking homes differ, with IDS to females having a greater pitch range than IDS to males, and more likely to serve the function of encouraging infants' attention (Kitamura & Burnham, 2003; Kitamura, Thanavishuth, Burnham, & Luksaneeyanawin, 2002). These differences in speech to female and male infants are especially pronounced at 12 months, which is also the age at which we collected the language samples. There were no differences in the amount of ambient speech potentially heard by males and females, as assessed by LENA estimates, or in the proportion of speech to them in the IDS register. In fact, for both males and females, a substantial amount of the speech addressed to them was in the ID register. However, it is possible that the features of IDS spoken to males and females learning American English differ, and that the IDS to females contains more cues that support

learning NADs. Thus, analyses targeting specific features of IDS might be more sensitive to sex-related input differences. Another possibility is that females and males respond to key features of IDS differently. We are not aware of evidence that there are sex differences in infants' preferences for IDS at a gross level, but it may be that female infants are more likely to encode or attend to specific features of IDS that are particularly relevant to NAD learning. This possibility is consistent with evidence that sensitivity to pitch changes is related to NAD learning in 3-month-olds (Mueller *et al.*, 2012).

In sum, these findings provide evidence that infants' performance on lab-based artificial-language-learning tasks can be used to shed light on the mechanisms supporting acquisition of their native language. Here we found evidence that infants' ability to track non-adjacent statistical regularities is related to their native language development. Furthermore, this approach revealed that there may be differences in the factors that support NAD learning for males and females, as well as in the underlying processes that are recruited for learning. Finally, these findings suggest that studying the interaction between individuals and their environments can yield insights into the nature of language-learning mechanisms.

**Acknowledgements.** This work was supported by funds from NSF BCS-1352443 to J. L.

## References

- Bates, E., Bretherton, I., & Snyder, L. S. (1991). *From first words to grammar: individual differences and dissociable mechanisms* (Vol. 20). Cambridge University Press.
- Bayley, N. (2006). *Bayley scales of infant and toddler development: administration manual* (3rd ed.). Psychorp.
- Beech, J. R., & Beauvois, M. W. (2006). Early experience of sex hormones as a predictor of reading, phonology, and auditory perception. *Brain and Language*, 96(1), 49–58.
- Cartwright, T. A., & Brent, M. R. (1997). Syntactic categorization in early language acquisition: formalizing the role of distributional analysis. *Cognition*, 63, 121–70.
- Christiansen, M. H., Onnis, L., & Hockema, S. A. (2009). The secret is in the sound: from unsegmented speech to lexical categories. *Developmental Science*, 12, 388–95.
- Christophe, A., Millotte, S., Bernal, S., & Lidz, J. (2008). Bootstrapping lexical and syntactic acquisition. *Language and Speech*, 51(1/2), 61–75.
- Culbertson, J., Koulaguina, E., Gonzalez-Gómez, N., batesdre, G., & Nazzi, T. (2016). Developing knowledge of nonadjacent dependencies. *Developmental Psychology*, 52, 2174–83.
- Fisher, C., & Tokura, H. (1996). Acoustic cues to grammatical structure in infant-directed speech: cross-linguistic evidence. *Child Development*, 67, 3192–218.
- Friederici, A. D., Mueller, J. L., & Oberecker, R. (2011). Precursors to natural grammar learning: preliminary evidence from 4-month-old infants. *PLoS One*, 6(3), e17920.
- Friederici, A. D., Pannekamp, A., Partsch, C., Ulmen, U., Oehler, K., Schmutzler, R., & Hesse, V. (2008). Sex hormone testosterone affects language organization in the infant brain. *Cognitive Neuroscience and Neuropsychology*, 19, 283–6.
- Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, 13, 431–6.
- Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, 70(2), 109–35.
- Gómez, R. L., & Lakusta, L. (2004). A first step in form-based category abstraction in 12-month-old infants. *Developmental Science*, 7, 567–80.
- Gómez, R. L., & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy*, 7, 183–206.
- Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: P.H. Brookes.
- Höhle, B., Schmitz, M., Santelmann, L. M., & Weissenborn, J. (2006). The recognition of discontinuous verbal dependencies by German 19-month-olds: evidence for lexical and structural influences on children's early processing capacities. *Language Learning and Development*, 2, 277–300.

- Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research*, 5, 69–95.
- Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, 61(4), 343–65.
- Jusczyk, P. W., Hirsh-Pasek, K., Nelson, D. G. K., Kennedy, L. J., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, 24(2), 252–93.
- Kitamura, C., & Burnham, D. (2003). Pitch and communicative intent and mother's speech: adjustments for age and sex in the first year. *Infancy*, 4, 85–110.
- Kitamura, C., Thanavishuth, C., Burnham, D., & Luksaneeyanawin, S. (2002). Universality and specificity in infant-directed speech: pitch modifications as a function of infant age and sex in a tonal and non-tonal language. *Infant Behavior and Development*, 24, 372–92.
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255, 606–8.
- Lany, J. (2014). Judging words by their covers and the company they keep: probabilistic cues support word learning. *Child Development*, 85, 1727–39.
- Lany, J., & Gómez, R. L. (2008). Twelve-month-old infants benefit from prior experience in statistical learning. *Psychological Science*, 19, 1247–52.
- Lany, J., & Saffran, J. R. (2010). From statistics to meaning: infant acquisition of lexical categories. *Psychological Science*, 21, 284–91.
- Lany, J., Shoaib, A., Thompson, A., & Estes, K. G. (2018). Infant statistical-learning ability is related to real-time language processing. *Journal of Child Language*, 45(2), 368–91.
- Legendre, G., Barrière, I., Goyet, L., & Nazzi, T. (2010). Comprehension of infrequent subject–verb agreement forms: evidence from French-learning children. *Child Development*, 81(6), 1859–75.
- Lidz, J., & Gleitman, L. R. (2004). Argument structure and the child's contribution to language learning. *Trends in Cognitive Sciences*, 8(4), 157–61.
- Liu, H.-M., Kuhl, P. K., & Tsao, F. M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science*, 6, F1–F10.
- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82, B101–B111.
- Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition*, 90, 91–117.
- Mintz, T. H., Newport, E. L., & Bever, T. G. (2002). The distributional structure of grammatical categories in speech to young children. *Cognitive Science*, 26, 393–424.
- Monaghan, P., Christiansen, M. H., & Chater, N. (2007). The phonological-distributional coherence hypothesis: cross-linguistic evidence in language acquisition. *Cognitive Psychology*, 55, 259–305.
- Mueller, J. L., Friederici, A. D., & Männel, C. (2012). Auditory perception at the root of language learning. *Proceedings of the National Academy of Sciences*, 109, 15953–8.
- Nazzi, T., Barrière, I., Goyet, L., Kresh, S., & Legendre, G. (2011). Tracking irregular morphophonological dependencies in natural language: evidence from the acquisition of subject–verb agreement in French. *Cognition*, 120(1), 119–35.
- Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I: statistical learning of non-adjacent dependencies. *Cognitive Psychology*, 48(2), 127–62.
- Ngon, C., Martin, A., Dupoux, E., Cabrol, D., Dutat, M., & Peperkamp, S. (2013). (Non)words, (non)words, (non)words: evidence for a protolexicon during the first year of life. *Developmental Science*, 16(1), 24–34.
- Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical learning in a natural language by 8-month-old infants. *Child Development*, 80, 674–85.
- Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science*, 298, 604–7.
- Pinker, S. (1989). *Learnability and cognition: the acquisition of argument structure*. Cambridge, MA: MIT Press.
- Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, 36(4), 717–31.
- Ramírez-Esparza, N., García-Sierra, A., & Kuhl, P. K. (2014). Look who's talking: speech style and social context in language input to infants are linked to concurrent and future speech development. *Developmental Science*, 17(6), 880–91.

- Ramirez-Esparza, N., Garcia-Sierra, A., & Kuhl, P. K. (2015). Look who's talking: speech style and social context in language input to infants are linked to concurrent and future speech development. *Developmental Science*, 17, 880–84.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926–8.
- Saffran, J. R., & Wilson, D. P. (2003). From syllables to syntax: multilevel statistical learning by 12-month-old infants. *Infancy*, 4(2), 273–84.
- Santelmann, L. M., & Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: evidence for limitations in processing space. *Cognition*, 69(2), 105–34.
- Schaadt, G., Hesse, V., & Friederici, A. D. (2015). Sex hormones in early infancy seem to predict aspects of later language development. *Brain and Language*, 141, 70–6.
- Shoib, A., Wang, T., Hay, J. F., & Lany, J. (2018). Do infants learn words from statistics? Evidence from English-learning infants hearing Italian. *Cognitive Science*, 42, 3083–99.
- Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, 7(1), 53–71.
- Van Heugten, M., & Johnson, E. K. (2010). Linking infants' distributional learning abilities to natural language acquisition. *Journal of Memory and Language*, 63(2), 197–209.
- Van Heugten, M., & Shi, R. (2010). Infants' sensitivity to non-adjacent dependencies across phonological phrase boundaries. *Journal of the Acoustical Society of America*, 128(5), EL223–EL228.
- Wang, F. H., & Mintz, T. H. (2018). Learning nonadjacent dependencies embedded in sentences of an artificial language: when learning breaks down. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(4), 604–14.
- Weisleder, A., & Fernald, A. (2013). Talking to children matters: early language experience strengthens processing and builds vocabulary. *Psychological Science*, 24, 2143–52.
- Willits, J., Saffran, J., & Lany, J. (2017). Toddlers can use semantic cues to learn difficult nonadjacent dependencies. doi.org/10.31234/osf.io/4ca78
- Witte, A. V., Savli, M., Holik, A., Kasper, S., & Lanzenberger, R. (2010). Regional sex differences in grey matter volume are associated with sex hormones in the young adult human brain. *Neuroimage*, 49(2), 1205–12.
- Xu, D., Yapanel, U., & Gray, S. (2009). Reliability of the LENATM language environment analysis system in young children's natural home environment. Retrieved from <<http://www.lenafoundation.org/TechReport.aspx/Reliability/LTR-05-2>>.
- Yang, C. D. (2004). Universal Grammar, statistics or both? *Trends in Cognitive Sciences*, 8(10), 451–6.

## Appendix

Materials used in the NAD tasks

15-Month NAD task materials

	Grammatical frames	Ungrammatical frames	X-words
G1	pel X jic	pel X rud	wadim
	vot X rud	vot X jic	kicey
			puser
G2	pel X rud	pel X jic	fengle
	vot X jic	vot X rud	coomo
			loga
			gople

(Continued)

(Continued.)

Grammatical frames	Ungrammatical frames	X-words
		taspu
		hifam
		deecha
		vamey
		skiger
		benez
		gensim
		feenam
		laeljeen
		chila
		roosa
		plizet
		balip
		malsig
		suleb
		nilbo
		wiffle

18-Month NAD task materials

Grammatical frames	Ungrammatical frames	X-words
She is Xing	She can Xing	bav
She can X it	She can Xing	boof
		chan
		cheel
		chup
		doove
		dut
		gip
		giz
		goot
		jaf

(Continued)

(Continued.)

Grammatical frames	Ungrammatical frames	X-words
		jope
		kag
		kiv
		koom
		paj
		pif
		puz
		teej
		tep
		tig

**Cite this article:** Lany J, Shoaib A (2020). Individual differences in non-adjacent statistical dependency learning in infants. *Journal of Child Language* 47, 483–507. <https://doi.org/10.1017/S0305000919000230>