

Running head: NON-ADJACENT DEPENDENCY ACQUISITION BY INFANTS

**\*\*DRAFT – PLEASE DO NOT CITE\*\***

Linking Infants' Distributional Learning Abilities to Natural Language Acquisition

Marieke van Heugten and Elizabeth K. Johnson

University of Toronto

Correspondence should be addressed to Elizabeth Johnson, Department of Psychology,  
University of Toronto Mississauga, 3359 Mississauga Road North, Mississauga, Ontario,  
Canada, L5L 1C6. E-mail: elizabeth.johnson@utoronto.ca. Phone: 905-569-4785. Fax: 905-569-  
4326.

## Abstract

This study examines the link between distributional patterns in the input and infants' acquisition of non-adjacent dependencies. In two Headturn Preference experiments, Dutch-learning 24-month-olds (but not 17-month-olds) were found to track the remote dependency between the definite article *het* and the diminutive suffix *-je* while no such evidence was obtained for the remote dependency between the definite article *de* and the plural suffix *-en*. In a follow-up corpus analysis, the distributional statistics in children's input (i.e. frequency, forward and backward transitional probabilities, and average distance between the two elements) were found to elegantly align with the behavioral data; distributional properties of diminutive and plural dependencies differed substantially, with more advantageous patterns for diminutive than for plural dependencies. Our results thus support the notion that there is a strong link between input distributions and the ease with which children acquire sensitivity to remote dependencies. Potential implications are discussed.

Keywords: distant dependencies; statistical learning; morphosyntax; language acquisition; linguistic input.

### Linking Infants' Distributional Learning Abilities to Natural Language Acquisition

Theories of language development often focus on how children start acquiring the meaning of words. Comprehending language, however, entails more than simply accessing the dictionary definition of consecutive words in the mental lexicon. In order to understand sentences, it is also crucial to determine how words relate to one another. Relationships between elements in sentences are plentiful in natural languages. While determiners, for example, are typically followed by a noun, pronouns or auxiliaries tend to be followed by verbs (e.g., *a book* but not *\*a is reading*; *he is reading* but not *\*he book*). The question thus arises how listeners start processing these interdependent co-occurrences.

The skills to acquire dependencies are arguably in place from early on. Using artificial languages, both adults and infants have repeatedly been shown to possess the ability to track adjacent dependencies (e.g., Aslin, Saffran, & Newport, 1998; Saffran, Aslin, & Newport, 1996; Saffran, Newport, & Aslin, 1996) and, in the presence of sufficient convergent evidence in the input, form relationships between abstract linguistic categories (Frigo & McDonald, 1998; Gómez & Gerken, 1999; Gómez & Lakusta, 2004; Mintz, 2002; Saffran, 2001). At 12 months of age, for instance, infants presented with a training language consisting of *aX* and *bY* patterns (where *a*, *b*, *X*, and *Y* formed “grammatical categories”, with *X* being bisyllabic and *Y* monosyllabic words) later generalized these patterns to novel instances of (bisyllabic) *X* following *a* and (monosyllabic) *Y* following *b* (Gómez & Lakusta, 2004). The ability to track such abstract categories could be of great value for the acquisition of natural language dependencies, in that this could be the mechanism that enables children in their second year of life to start grouping words into different word categories based on the immediately preceding element (Gerken & McIntosh, 1993; Höhle, Weissenborn, Kiefer, Schulz, & Schmitz, 2004; Johnson, 2005; Kedar, Casasola, & Lust, 2006;

Lew-Williams & Fernald, 2007; Shi & Melançon, in press; Shipley, Smith, & Gleitman, 1969; Van Heugten & Johnson, in press; Van Heugten & Shi, 2009; Zangl & Fernald, 2007).

Children thus readily acquire the relationship between frequently co-occurring adjacent elements in natural languages. These elements of a dependency, however, need not be (and are often not) adjacent. A determiner and noun, for example, may very well be separated by an adjective (e.g., *the colorful ball*). Moreover, some dependencies are never adjacent. Consider the present continuous tense, in which a form of *to be* is paired with the suffix *-ing* (e.g., *is singing*) or the English plural dependency consisting of a plural determiner and the plural suffix *-s* (e.g., *these balls*). These co-occurrences never occur adjacently, but can nonetheless potentially be greatly informative. In fact, corpus studies of child-directed speech have shown that frequent frames, combinations of two non-adjacent frequently co-occurring function words, are of considerable predictive value for determining the word category of the intervening syllables (Chemla, Mintz, Bernal, & Christophe, 2009; Mintz, 2003). The (remote) co-presence of the function words *the* and *in*, for example, is almost always interceded by a noun, while the function words *you* and *it* are typically interceded by a verb. Basic distributional co-occurrence patterns in the input could thus be used to categorically group intervening words into separate categories. Needless to say, this categorization cue can only be used once infants gain sensitivity to these frequent remote dependencies.

Headturn Preference studies have shown that the ability to track such non-adjacent co-occurrences develops early in life. Santelmann and Jusczyk (1998) tested English learners on their sensitivity to the non-adjacent present continuous *is-ing* dependency. Infants in this study were alternately presented with some passages containing grammatical (e.g., *is baking bread*) dependencies and some passages containing ungrammatical dependencies (e.g., *\*can baking*

*bread*). To create ungrammatical dependencies, the grammatical auxiliary *is* was replaced by the ungrammatical auxiliary *can*. Eighteen-month-olds listened significantly longer to grammatical as compared to ungrammatical passages, indicating that they have already acquired some sensitivity to the remote *is-ing* dependency. Similar results have been obtained for comparable verbal dependencies in German and Dutch (Höhle, Schmitz, Santelmann, & Weissenborn, 2006; Wilsenach, 2006 respectively).

While children thus possess the ability to track at least some dependencies from early on, artificial language studies have suggested that the ease with which they are learned depends on various factors. Perceptual cues such as the phonological similarity between the elements of the non-adjacent dependencies (Onnis, Monaghan, Richmond, & Chater, 2005), the use of presegmented units (Peña, Bonatti, Nespor, & Mehler, 2002), and the units the dependency consists of (Bonatti, Peña, Nespor, & Mehler, 2005; Newport & Aslin, 2004) have all been argued to play an important role in non-adjacent dependency acquisition.

In addition to the nature of the non-adjacent elements, a continuously growing body of research demonstrates that the exact distributional properties of the materials determine co-occurrence acquisition. The frequency of the dependency, its strength (both expressed in forward and backward transitional probabilities (TPs), the distance between elements in the dependency, and the variability in intervening syllables, for example, all affect how readily dependencies are acquired. The frequency of co-occurrence (e.g., Braine et al., 1990; as used by Mintz, 2003) may be the most intuitive measure of all and forms a simple count of the number of times both elements of a dependency co-occur. In general, the more frequently the two items co-occur, the easier the dependency is to learn. Related to frequency, but quantified as a relative number are the forward (e.g., Aslin et al., 1998; Morgan, Meier, & Newport, 1987; Saffran, 2001; Saffran, 2002;

Saffran et al., 1996a; Saffran et al., 1996b) and backward (Gervain, 2008; Pelucchi, Hay, & Saffran, 2009; Perruchet & Desaulty, 2008) TPs. Consider, for example, the co-occurrence of *the* and *dog*. Although the phrase *the dog* may occur relatively frequently in English, *the* can also be followed by numerous other nouns. Similarly, *dog* is not solely preceded by *the*, but also by other determiners such as *a*, *every*, or *that*. Taking into account the instances in which the individual forms occur with other words may thus prevent listeners from erroneously accepting *the dog* as one unit. Forward and backward TPs take into account this base frequency of the individual words. More specifically, forward TPs are defined as the probability of encountering the second element of a dependency (i.e. *dog* in the above example) given the first (*the*) and backward TPs are defined as the probability of encountering the first element (*the*) given the second (*dog*). The higher the TPs, the more strongly related the two elements are, and, ceteris paribus, the easier they are to track. Both frequency and TPs, as well as the nature of listeners' sensitivity to these distributional cues (Aslin et al., 1998; Mirman, Graf Estes, & Magnuson, in press; Perruchet & Peereman, 2004), have been discussed at length in the literature examining the acquisition of adjacent dependencies. Because the acquisition of adjacent and non-adjacent dependencies are strongly related (e.g., Lany & Gómez, 2008), mechanisms employed in acquiring one dependency are likely also used in the acquisition of the other dependency and should hence be taken into account. A further distributional property that may affect dependency acquisition is the average distance between the two non-adjacent elements. Infants only have a limited processing window and tracking dependencies over long distances is more demanding than tracking near-by dependencies (Santelmann & Jusczyk, 1998), causing shorter-distance relationships to be learned more easily than longer-distance relationships. A final factor considered here is the variability in intervening material. Artificial language studies have shown that non-adjacent dependencies are

learned with greater ease when the intervening material is more diverse (Gómez, 2002; Gómez & Maye, 2005). It has been argued that the less likely the intervening material will form an adjacent dependency, the more likely listeners' focus will be drawn towards longer-distance dependencies. More variability in the interceding syllables may thus facilitate non-adjacent dependency learning.

Artificial language studies have thus generated important hypothesis concerning what dependencies should be easy and what dependencies should be challenging to acquire from the input. No study that we are aware of, however, has explored the link between infants' distributional learning abilities and the acquisition of specific linguistic structures in natural languages (see Johnson & Tyler, 2010; Yang, 2004 for a related discussion). In the present study, we therefore examined whether infants' sensitivity to non-adjacent co-occurrences is influenced by distributional cues in their input. Dutch diminutive and plural dependencies, both consisting of a determiner and suffix interceded by a noun, offer an interesting approach to address this question. The precise structure and distributional properties of these dependencies make them particularly well-suited for this study. Dutch, much like French and Spanish, has two gender-marked definite articles preceding singular nouns, *de* and *het* (e.g., *de hond* 'the dog' and *het paard* 'the horse'). Diminutives (which are very productive in child-directed Dutch; Gillis, 1997), formed by adding the suffix *-je* (or one of its allomorphs *-tje*, *-pje*, *-kje*, *-etje*, all containing phonological overlap with *-je*) to a noun, can only be preceded by the definite article *het* (e.g., *de hond* 'the dog'; *het hondje* 'the doggie'). This results in a remote co-occurrence pattern of the article *het* and the suffix *-je*, interceded by a noun. Plurals, mostly ending in *-en*, in contrast, can only be preceded by the definite article *de* (e.g., *het paard* 'the horse'; *de paarden* 'the horses'), leading to a non-adjacent dependency between *de* and *-en*. Thus, Dutch diminutive and plural dependencies are not only structurally similar to each other (consisting of a definite article plus

suffix intervened by any given count noun), they are also similar to the previously tested verbal non-adjacent dependencies (both consist of a free and bound monosyllabic function morpheme interceded by a content word).

Despite this structural parallel, however, diminutive and plural dependencies do differ from one another in other aspects and may thus not be equally easy to acquire. Most importantly, the diminutive suffix *-je* only serves as a diminutive marker, while the suffix *-en* (or phonologically similar *-e*) can be used as a number or infinitive marker on verbs, gender marker on adjectives, or a plural marker on nouns. In addition, several noun stems and perfect forms of irregular verbs also end in *-en* (e.g., *eten* ‘food’ or *gelopen* ‘walked’). As a result, a word-final *-en* indicates a plural marker much less reliably than word-final *-je* indicates a diminutive marker. Such differences may be reflected in the statistical regularities of these dependencies and could, as a result, make diminutive dependencies easier to learn than plural dependencies.

The current study examines the acquisition of these nominal non-adjacent dependencies as well as the potential link to distributional properties in the input. In Experiment 1 and 2, we first test Dutch-learning 17- to 24-month-olds’ sensitivity to the non-adjacent diminutive and plural dependencies. Using nominal rather than verb dependencies also enables us to explore whether infants’ early ability to track non-adjacent dependencies generalizes to syntactic categories other than verbs. These behavioral studies are subsequently followed up by a corpus study, exploring the relationship between distributional statistics in the input and the ease of acquisition of remote dependencies.

### Experiment 1

In Experiment 1, the Headturn Preference Procedure (Kemler Nelson, Jusczyk, Mandel, Myers, Turk, & Gerken, 1995) was used to test Dutch 17- and 24-month-olds’ sensitivity to the

non-adjacent diminutive dependency (i.e. *het X-je*). Infants were presented with lists of noun phrases (NPs) consisting of a definite article (i.e. *de* or *het*) and a diminutive noun (e.g., *kagje*). Listening times to lists with the grammatical (i.e. *het kagje*) or ungrammatical NPs (i.e. *de kagje*) were measured. Previous studies have shown that infants prefer to listen to grammatical over ungrammatical discontinuous verb-based dependencies (Höhle et al., 2006; Santelmann & Jusczyk, 1998; Wilsenach, 2006). If Dutch learners have gained sensitivity to the non-adjacent diminutive dependency, they should thus display longer listening times to grammatical as opposed to ungrammatical lists.

While past studies have found a listening preference for grammatical over ungrammatical remote dependencies, these studies tested children on existing, grammatical dependencies (e.g., *is baking*). Some of these dependencies had undoubtedly been heard by the infants prior to testing. The ungrammatical dependencies (e.g. *can baking*), however, were likely never heard. Thus, children's performance in these studies may have reflected their recall of stored familiar word strings rather than their knowledge of non-adjacent dependencies. For this reason, the dependencies used in the current study were all intervened by nonsense words. When nonsense words are used, neither the grammatical nor the ungrammatical sequences could have been previously encountered in the input.

### *Method*

#### *Participants.*

Twenty-eight monolingual 24-month-old (age range: 732 - 764 days, mean: 747 days, SD: 9.29, 17 boys and 11 girls) and 28 monolingual 17-month-old (age range: 509 - 540 days, mean: 525 days, SD: 8.51, 16 boys and 12 girls) Dutch-learning infants from the Nijmegen region participated. An additional 13 24-month-olds and 12 17-month-olds were tested, but excluded

from the analysis due to fussiness or crying (22), parental interference (1), or equipment error (2). The participants in this and the following experiment were drawn from city birth records. Families received ten euros or a small gift in appreciation for their participation.

*Materials.*

A total of 20 phonotactically legal CVC-nonwords (*deet, geet, has, hies, hiet, hog, kag, kog, nas, neg, not, vas, veut, viet, vog, wag, zas, zeut, zies, ziet*) were created. These nonwords were followed by the diminutive suffix *-je* and either preceded by the grammatical definite article *het* or by the ungrammatical definite article *de*. Grammatical and ungrammatical NPs were thus identical apart from the article preceding the diminutive noun. The distance between the two function morphemes was exactly one syllable for all NPs.

A female native Dutch speaker recorded the NPs in a sound-attenuated booth (sample rate: 44.1 kHz, 16 bit). To avoid artifacts that could potentially surface as a result of producing ungrammatical utterances, both grammatical and ungrammatical nonword NPs were cross-spliced. For each NP set, four (grammatical) utterances were recorded: two tokens of the (grammatical) diminutive NP (e.g., *het kagje*) and two tokens of the (grammatical) plural dependency (e.g., *de kaggen*). Only one plural NP token was used for stimulus creation in this experiment, the other one was recorded to use in Experiment 2. The diminutive nonsense noun (e.g., *kagje*) was spliced from one of the two diminutive NP recordings. This noun token was used in both the grammatical and the ungrammatical experimental stimuli. The preceding article in the grammatical condition (i.e. *het*) was then spliced from the other diminutive NP recording, while the preceding article in the ungrammatical condition (i.e. *de*) was spliced from the plural NP recording. Since the stem of the noun was the same in the original diminutive and plural NP recordings (e.g., *kag* in both *kagje* and *kaggen*), articles and nouns were spliced after the word-initial consonant, immediately before the

onset of the vowel, thereby eliminating any disfluent phoneme transition due to cross-splicing nouns onto acoustically distinct articles. The combination of each article with the (same token of the) diminutive noun formed the grammatical (e.g., *het kagje*) and ungrammatical NPs (e.g., *de kagje*) used in this experiment.

Twelve experimental lists of NPs were created. Six of these lists consisted of grammatical NPs and six consisted of ungrammatical NPs. Each list contained ten NPs, the order of which was held constant between the grammatical and ungrammatical trials. NPs within a list were separated from each other by 800-875 ms of silence and the total duration of each list was between 17.00 and 17.33 s (mean grammatical: 17.25 s; mean ungrammatical: 17.03 s).

#### *Design.*

Participants were randomly assigned to one of four orders. Within each order, half of the lists of nonsense nouns were first presented in grammatical NPs and the other half were first presented in ungrammatical NPs. Each order presented all twelve lists in a pseudo-randomized fashion, such that no more than two grammatical or ungrammatical trials could occur consecutively, and the side of presentation could not be the same on more than two consecutive trials.

#### *Procedure.*

Participants were tested using the Headturn Preference Procedure (Kemler Nelson et al., 1995). Children were seated on their caregiver's lap in the center of a dimly lit three-sided test booth. A blue light and a loudspeaker were mounted at eye level on each of the side panels. A third red light was positioned on the front panel. Underneath this front light, a camera videotaped the child's looking behavior. Trials were initiated by flashing the center red light. Once the child oriented toward this light, the experimenter, who monitored the child's looking behavior pressed a

button on a button box, which caused the red light to stop flashing and one of the two blue lights to start flashing. As soon as the child looked toward this blue light, an additional button press started a trial. All recordings were played at a comfortable listening level (approximately 68-70 dB). Trials were either played until the end of the speech stream or until the child looked away for two consecutive seconds. Both the parent and the experimenter wore sealed noise cancellation headphones playing masking music to prevent them from biasing the study.

### *Results*

Mean listening times to the grammatical and ungrammatical lists were calculated for each subject. As the initial trials in infant preference studies are often unstable (e.g., Marquis & Shi, 2008), the first two test trials (one grammatical and one ungrammatical) were excluded from data analysis. Across all 24-month-olds, the average listening time was 9.50 s (SEM = .58) for the grammatical trials and 7.65 s (SEM = .56) for the ungrammatical trials (see Figure 1), with 19 infants listening longer to grammatical lists. Across all 17-month-olds, the average listening time was 7.84 s (SEM = .63) for the grammatical trials while the average listening time to the ungrammatical trials was 7.93 s (SEM = .61), with 15 infants listening longer to grammatical than to ungrammatical lists. A mixed 2 (Age) x 2 (Grammaticality) ANOVA revealed a marginally significant main effect of Grammaticality ( $F(1, 54) = 3.070; p = .085; \eta_p^2 = .054$ ), but not Age ( $F(1, 54) = 1.018; p = .318; \eta_p^2 = .018$ ). More importantly, however, the interaction between Age and Grammaticality approached significance ( $F(1,54) = 3.729; p = .059; \eta_p^2 = .065$ ),<sup>1</sup> suggesting that the listening duration toward grammatical and ungrammatical trials is age-dependent. Planned comparisons indicated that only 24-month-olds listened longer to grammatical as opposed to ungrammatical lists ( $t(27) = 2.775, p = .01, \eta_p^2 = .222$ ; mean difference = 1.84, 95%-confidence interval (CI) = .48 - 3.20; all t-tests reported in this paper are two-tailed).<sup>2</sup> No such differences

were found for 17-month-olds ( $t(27) = -.120, p = .906, \eta_p^2 = .001$ ; mean difference = .09, 95%-CI = -1.63 - 1.45).<sup>3</sup>

### *Discussion*

Dutch 24-month-olds, but not 17-month-olds listen longer to grammatical than ungrammatical NPs, indicating that they start tracking the non-adjacent dependency between the definite article *het* and the diminutive suffix *-je* sometime between 17 and 24 months of age, at least when the dependency is intervened by only one syllable. This extends previous work with verbal dependencies to noun dependencies, indicating that early sensitivity to remote dependencies not restricted to verbs and emerges without much delay compared to auxiliary-verb dependencies in Dutch (Wilsenach, 2006). Moreover, it shows that children are indeed able to track generalized dependencies and are not merely remembering frozen sequences of previously encountered words.

Nominal non-adjacent dependencies are, however, not uniform. Dutch plurals, like diminutives, consist of a definite article (i.e. *de*) and a suffix (mostly *-en*). As discussed in the Introduction, however, the suffix *-en* (or phonologically similar *-e*) is not only used on nouns, but also as an infinitive or perfect marker on verbs and gender marker on adjectives. The diversity of uses of this suffix may confuse infants, potentially leading to delayed acquisition of the plural compared to the diminutive dependency. It is therefore important to examine whether sensitivity to one non-adjacent dependency in a natural language generalized to other (structurally similar) non-adjacent dependencies.

Examining plural dependencies is also beneficial because it allows us to entertain an alternative explanation for the results of Experiment 1. We found that Dutch-learning 24-month-olds listened longer to grammatical than to ungrammatical trials and interpreted this to mean that children are sensitive to the non-adjacent diminutive dependency. Alternatively, these results

could indicate that children simply prefer listening to *het*-initial as opposed to *de*-initial NPs. Although *de* and *het* are approximately equally frequent in child-directed speech (Van Heugten & Johnson, in press), it is possible that phonological characteristics make *het* sound more pleasant and lead infants to prefer listening to *het* over *de*. Past studies ruled out this type of alternative explanation by showing that infants had no preference for grammatical over ungrammatical items when the number of intervening syllables was increased (Santelmann & Jusczyk, 1998) or when the syntactic structure of the intervening material was changed (Höhle et al., 2006). Although this explanation is also an unlikely one for our results given the absence of any preference for *het*-initial NPs at 17 months of age, it is nonetheless important that it be ruled out as the determining factor.

## Experiment 2

In Experiment 2, Dutch-learning 24-month-olds were tested on their sensitivity to the non-adjacent plural dependency. The goal of this experiment is twofold: to determine whether different nominal non-adjacent dependencies are acquired at similar rates and to rule out the possibility infants' preference to listen to the grammatical dependencies in Experiment 1 was merely due to a preference for *het*- over *de*-initial NPs. Like the diminutive dependency tested in the previous two experiments (i.e. *het X-je*), plurals consist of an article (*de*) and a suffix (*-en*) interceded by a noun (i.e. *de X-en*). Listening times were measured to lists of plural nonword NPs preceded by grammatical (i.e. *de*) or ungrammatical definite articles (i.e. *het*). Note that the grammatical definite article in this experiment was ungrammatical in the previous experiments and vice versa. If children listen longer to the grammatical than to the ungrammatical lists of NPs, this would indicate that they gain sensitivity to the plural and diminutive dependency around the same time.

In addition, it would exclude the possibility that the effect obtained in Experiment 1 resulted from a basic preference of *het*-initial over *de*-initial NPs.

### *Methods*

#### *Participants.*

Twenty-eight Dutch-learning 24-month-olds participated (age range: 729-762 days, mean: 747 days, SD: 9.51, 16 boys and 12 girls). An additional 11 children were tested, but excluded from the analysis due to fussiness or crying (10) or experimenter error (1).

#### *Materials.*

Materials were created from the same recordings used in Experiment 1. The only difference involved the use of plural rather than diminutive NPs. The nonword stems were identical to the nonword stems in Experiment 1 as was the distance between the two function morphemes (one syllable). As before, only the definite article differed between conditions and the final lists of NPs were cross-spliced. One plural noun token (e.g., *kaggen*) was spliced from the yet unused plural NP recording and was used in both the grammatical and the ungrammatical conditions. The preceding articles were spliced from a realization of a second plural NP for the grammatical condition (e.g., *de kaggen*) and from a realization of a diminutive NP in the ungrammatical condition (e.g., *het kaggen*). The total list length was between 17.42 and 17.83 s (mean grammatical: 17.46 s; mean ungrammatical: 17.68 s).

#### *Design and Procedure.*

The design and procedure were identical to Experiment 1.

### *Results*

Analyses were comparable to those in Experiment 1. The mean listening times for the grammatical and ungrammatical condition were calculated excluding the first two trials. The

average orientation time was 7.42 s (SEM = .50) for the grammatical and 7.55 s (SEM = .51; see Figure 1) for the ungrammatical trials, with 14 infants listening longer to the grammatical lists. A two-tailed paired sample t-test revealed that the difference between these two conditions was not significant ( $t(27) = -.236, p = .815, \eta_p^2 = .002$ ; mean difference = .12, 95%-CI = -1.12 - .95).<sup>4</sup>

To compare whether 24-month-olds' performance differed depending on whether they heard diminutives or plurals, a mixed 2 (Dependency) x 2 (Grammaticality) ANOVA was carried out. A marginally significant Dependency effect was found ( $F(1, 54) = 2.916; p = .093; \eta_p^2 = .051$ ), indicating that children tend to listen longer to diminutive than plural dependencies. There was also an effect of Grammaticality ( $F(1, 54) = 4.147; p = .047; \eta_p^2 = .071$ ), which implies that children overall listened longer to grammatical than to ungrammatical dependencies. Most importantly, however, these effects were modulated by a Dependency x Grammaticality interaction ( $F(1,54) = 5.420; p = .024; \eta_p^2 = .091$ ), indicating that the differential listening pattern found for 24-month-olds in Experiment 1 only held for the diminutive dependency. Thus, while 24-month-olds listen longer to grammatical than to ungrammatical diminutive NPs, no evidence was obtained that they differentiate between grammatical and ungrammatical plural NPs.

### *Discussion*

Contrary to the results obtained for the diminutive dependency, Dutch 24-month-olds listened equally long to grammatical and ungrammatical plural NPs, thereby eliminating the alternative explanation of Experiment 1. If the grammaticality effect in that experiment were mainly caused by a plain preference for *het* over *de*, this should have resulted in longer orientation times for the (ungrammatical) *het*-initial trials than for the (grammatical) *de*-initial trials in Experiment 2. The absence of such looking behavior, however, rules out this possibility.

More importantly, these results suggest that sensitivity to the remote plural dependency is not acquired until after 24 months of age, implying different rates of acquisition across structurally similar non-adjacent dependencies intervened by words from the same syntactic category. This raises the question of what factors make the diminutive dependency easier to learn than the plural dependency. Diminutives and plurals, despite being structurally similar, vary in other ways. Unlike the diminutive suffix *-je*, the plural suffix *-en* (or phonologically similar *-e*) surfaces on many syntactic categories such as verbs (e.g., *lopen* ‘to walk’, *gelopen* ‘walked’), adjectives (e.g., *mooie*, ‘nice’), or non-plural nouns (e.g., *eten* ‘food’). In other words, the word-final presence of the diminutive suffix form is a more reliable indicator of a diminutive than the word-final presence of the plural form is of a plural. It stands to reason that this difference in reliability between diminutive and plural nouns is reflected in the distributional patterns of those dependencies. To explore whether such statistics can be linked to differences in rate of acquisition between non-adjacent dependencies (as observed in our behavioral experiments), distributional differences between these two dependencies were analyzed from a corpus of infant-directed speech.

#### Corpus study

In this corpus analysis, we examine potential factors determining the rate of diminutive and plural dependency acquisition in Dutch. Both artificial and natural language studies examining adjacent or non-adjacent dependencies have put forward a number of statistics that may affect the ease with which dependencies are learned. As far as we are aware, however, no study to date has yet explored the link between these statistics and differences in rate of acquisition of specific natural language dependencies. In this experiment, we therefore carried out an analysis on a corpus of child-directed speech to examine the differences in infants’ input between diminutive and plural dependencies on a number of those statistics - the frequency of occurrence, the strength

of the relationship expressed in both forward and in backward TPs, the distance between the two non-adjacent elements, and the variability in the intervening syllables.

### *Method*

#### *Corpus.*

The Groningen corpus (Wijnen & Verrips, 1998) in the ChILDES database (MacWhinney, 2000) was used to explore the differences between the diminutive and plural dependencies in infant-directed Dutch. Only parental input received by children between 18 and 36 months of age was analyzed, leading to a total number of 305,413 words across all 13 parents.

#### *Analysis.*

In Experiment 1 and 2, the orientation time to grammatical dependencies was compared to the orientation time to ungrammatical dependencies. 24-month-olds listen longer to grammatical (i.e. *het X-je*) than to ungrammatical diminutive dependencies (i.e. *de X-je*), but no such effect was evidenced for plural dependencies and we concluded that diminutive dependencies are learned more easily than plural dependencies. When examining the regularities in the input, it is thus crucial to maintain these exact same comparisons used in the behavioral experiments: grammatical versus ungrammatical diminutive NPs on the one hand and grammatical versus ungrammatical plural NPs on the other hand. It may seem intuitive to use the actual diminutive and plural dependencies. Words in running speech, however, are (obviously) not tagged for their grammatical categories and such analysis would therefore not do justice to all non-plural words ending in *-en*. Rather than artificially eliminating spurious dependencies present in the input, we thus take into account as diminutives all words ending in *-je* and as plurals all remaining words ending in *-en* or phonologically similar *-e*, regardless of the actual syntactic category of the word.

Orthographic transcriptions of the sessions were used to compute the frequency of each of the two definite article forms, *de* and *het*. Although *het* can also surface as an expletive pronoun or a provisional subject, these forms are phonologically identical and are thus also included in our counts. In addition, an exhaustive overview and frequency count of all *-je*, *-en*, or *-e*-final words in the corpus was extracted, as well as an exhaustive overview of all (non-adjacent) combinations of *de* or *het* with *-je*-, *-en*-, or *-e*-final words within an utterance. Note that because words and co-occurrences were selected independent of their word categories, these lists include any spurious diminutive (e.g., *oranje* ‘orange’) or plural form (e.g., *lopen* ‘to walk’) as well as any spurious grammatical or ungrammatical diminutive or plural dependency between *de* or *het* and word-final *-je* or *-en* (e.g., *het eten* ‘the food’ consists of *het* plus word-final *-en* and is thus selected even though this co-occurrence does not involve a plural dependency) that may be present in the input. A native Dutch speaker (the first author) then eliminated all cases where grapheme-to-phoneme mismatches had caused words to be included that did not end in [jə] or [ə] (the typical realization for diminutive and plural suffixes, respectively), when produced in a colloquial fashion. The Dutch word *meteen* ‘immediately’, for example, orthographically ends in *-en* but is pronounced as [møten], with a full vowel rather than a schwa in the second syllable. In addition, because diminutive and plural markers only surface as suffixes attached to a root morpheme, any monosyllabic word that may have been in the lists was also excluded from the analyses. The Dutch pronoun *je* ‘you’ and the indefinite article *een* ‘a’ were thus not taken into account. Two different lists were created, one for *-je*-final and one for *-e* and *en*-final words (henceforth: *-en*-final words), with *-je*-final words only occurring in the former list.

Next, the number of syllables in between the definite articles and suffixes was counted. Children are known to have a limited processing window, and by 18 months of age, English learners only compute distant verbal dependencies if the two elements are separated by no more than three intervening syllables (Santelmann & Jusczyk, 1998). Participants in Experiment 1 and 2 were slightly older and hence we used a slightly longer window - only combinations with elements separated by more than five intervening syllables were discarded.<sup>5</sup> Subsequently, for all diminutive and plural dependencies, a list was formed containing all unique combinations of intervening syllables as to create a measure of variability of the interceding material.

Finally, in the absence of any morphological tags, in order to confirm that word-final *-je* is a more reliable predictor for a diminutive than word-final *-en* is for a plural, all words in the *-je*-final list were coded for whether they were a diminutive and all words in the *-en*-final list were coded for whether they were a plural. Where necessary, context was used to determine whether a word was or was not a diminutive or plural.

## *Results*

### *Suffix reliability.*

As a first step, the consistency of the diminutive and plural suffix was computed. Given that the plural ending *-en*, but not the diminutive suffix *-je* can be used on a wide variety of word categories, one may expect the diminutive suffix to reliably indicate the presence of a diminutive, but the plural suffix to fail to reliably do so for a plural. Indeed, while 99% (4,932 out of 4,987) of the words ending in *-je* are diminutives, the probability of words ending in *-en* being a plural is only 9% (2,533 out of 28,915). Similarly, the likelihood that a diminutive dependency is being formed when *het* and *-je* co-occur within a domain of five intervening syllables is 82% (675 out of 825) while the co-occurrence of *de* and *-en* within the same domain forms a plural dependency

only 18% (413 out of 2,339) of the time. The exclusivity of the diminutive suffix may thus make it a more reliable indicator of a diminutive than the plural suffix is of a plural. This difference in suffix reliability may further affect the diminutive and plural dependencies to be different on the measures discussed below.

*Frequency.*

For the frequency measure, the frequency of each of the grammatical and ungrammatical diminutive and plural dependencies was calculated. As for the diminutive, a *-je*-final word was preceded by the diminutive determiner (*het*) 825 times and by the plural determiner (*de*) 50 times. With regard to the plural, a *-en*-final word was preceded by the plural determiner (*de*) 2,339 times and by the diminutive determiner (*het*) 1,704 times (see Table 1). A Two-Way Chi-square analysis with Dependency Type (diminutive vs. plural) and Grammaticality<sup>6</sup> (grammatical vs. ungrammatical determiner) as independent factors revealed a significant association between these two factors ( $\chi^2(1) = 416.117$ ;  $p < 0.001$ ), indicating that the relative frequency of grammatical diminutive dependencies (odds ratio:  $825/50 = 16.5$ ) is different from the relative frequency of grammatical plural dependencies (odds ratio:  $2,339/1,704 = 1.37$ ). Odds ratios can be seen as a measure of the strength of the frequency discrepancies between grammatical and ungrammatical dependencies. More specifically, higher odds ratios indicate the presence of relatively more grammatical than ungrammatical co-occurrences in the input. If the frequency of the dependency plays a role when two non-adjacent dependencies are pitted against each other (as we did in Experiment 1 and 2, where we presented children with grammatical and ungrammatical dependencies), higher odds ratios are thus considered advantageous. Indeed, the odds ratio associated with the diminutive dependency is larger than the odds ratio associated with the plural dependency, illustrating that the proportion of grammatical diminutive dependencies is higher than

the proportion of grammatical plural dependencies. Based on input frequencies alone, the grammaticality effect for the diminutive dependency would thus be expected to be greater than any effect found for the plural dependency, thereby providing a potential explanation for the results obtained from the behavioral experiments.

*Forward transitional probabilities.*

Next, forward TPs were examined. The frequency of each of the grammatical and ungrammatical dependencies was compared to the number of cases the determiner was followed by anything but the suffix. Although a strict definition of TPs would take into consideration the total number of occurrences of the determiner as a baseline, in these adjusted TP analyses only the number of cases where the determiner was not followed by the suffix within five syllables was used as for the comparison. These two measures are closely related (the total number of occurrences of the determiner can easily be found by summing the number of cases where the determiner is followed by the suffix and the number of cases where the determiner is not followed by the suffix), but using this latter statistic prevents violating Chi-square's assumption of independence of each individual data point.

With regard to the diminutive dependency, the grammatical determiner *het* was followed by a *-je*-final word 825 times and by anything but a *-je*-final word 5,814 times (adjusted TP: .142). The plural determiner *de*, in contrast, co-occurred with the diminutive suffix 50 times and with anything but the diminutive suffix 6,518 times (adjusted TP: .008; see Table 2 for an overview). A two-way Chi-square analysis with Grammaticality (grammatical vs. ungrammatical determiner) and Suffix (present vs. absent) revealed a significant association between the two factors ( $\chi^2(1) = 726.257; p < 0.001$ ). If we compare this to the plural dependency, *de* was followed by a *-en*-final word 2,339 times and by anything but a *-en*-final word 4,229 times (adjusted TP: .553), while *het*

was followed by a *-en*-final word 1,704 times and by anything but a *-en*-final word 4,935 times (adjusted TP: .345; see Table 2 for an overview). Comparable to the diminutive dependency, the two-way Chi-square indicated that there was a significant association between the two factors ( $\chi^2(1) = 153.748$ ;  $p < 0.001$ ). The odds ratio for the diminutive dependency ( $0.142/0.008 = 18.50$ ) was, however, much larger than the odds ratio for the plural dependency ( $0.553/0.345 = 1.60$ ).

These analyses show that while the TP between a grammatical definite article and suffix is higher than the TP between an ungrammatical definite article and suffix for both the diminutive and the plural dependency, this (relative) difference is much more pronounced for the diminutive than for the plural dependency. Higher odds ratios indicate a higher probability of encountering the suffix given the article for the grammatical than for the ungrammatical article. Thus, if forward TPs affect the likelihood that non-adjacent dependencies are being tracked, the higher the odds ratio, the more advantageous the grammatical dependency would be. Much like our frequency analysis, the forward TP analysis therefore suggests that the input favors diminutive dependencies over plural dependencies. Both frequency and forward TPs are thus consistent with our behavioral results.

*Backward transitional probabilities.*

Another statistic determining the strength of a relationship is the backward TP. Backward TPs are highly similar to forward TPs, but with the second rather than the first element as the baseline (i.e. the suffix and not the determiner). Similar to the forward TPs the number of times a suffix is preceded by the (grammatical or ungrammatical) definite article should be compared to the number of times the suffix is preceded by anything but that definite article. One challenge this analysis poses, however, is that the occurrences are no longer independent of each other, thus violating the assumption of independence needed to perform a Chi-square analysis. The

occurrences in which the diminutive suffix *-je* is preceded by the grammatical definite article *het*, for example, partially overlap with the occurrences of *-je* being preceded by anything but *de*. Similarly, all cases in which either suffix is preceded by anything but a definite article overlap between the grammatical and ungrammatical dependencies. In order to avoid this problem, we will compare the number of times a suffix is preceded by the definite article *het* (or *de*) to the number of times it is preceded by anything but a definite article. Although this may make the comparison slightly less transparent, the changes in TPs are only minimal (the largest difference involves the grammatical plural dependency where the backward TP changes from .088 to .094, a difference of .006).

The grammatical diminutive co-occurrence of *het* and word-final *-je* of 825 times is now compared to the number times word-final *-je* is preceded by anything but a definite article (4,112; adjusted TP: .201) and so is the number of ungrammatical diminutive occurrence of *de* and word-final *-je* of 50 times (adjusted TP: .012; see Table 3 for an overview). With both the grammatical and the ungrammatical diminutive dependency being compared to the same number (i.e. the number of times the suffix is not preceded by any definite article), we can bypass a two-way Chi-square and directly compare the number of grammatical and ungrammatical co-occurrences. A binomial test shows that these numbers are different ( $z = 26.17; p < .001$ ). For the plural dependency, a *-en*-final word was preceded by *de* 2,339 times, by *het* 1,704 times, and by anything but a definite article 24,872 times (adjusted TPs: .094 and .069 respectively; see Table 3 for an overview). As was the case with the diminutive dependency, the difference between occurrences of grammatical and ungrammatical definite articles reaches significance ( $z = 9.97; p < .001$ ). The ratios of grammatical versus ungrammatical TPs, however, once again reveal major differences between the diminutive ( $.210/.012 = 16.50$ ) and plural dependencies ( $.094/.068 = 1.37$ ).

Similar to the forward TPs, these analyses demonstrate that the difference in backward TPs between the diminutive and plural dependencies is captured in the size of the effect. That is, for both dependencies the backward TP between the suffix and the grammatical definite article is higher than the TP between the suffix and the ungrammatical definite article. This (relative) difference is, however, much more prominent in the diminutive than in the plural dependency. Given that higher ratios express higher probabilities of encountering the grammatical than the ungrammatical definite article given the suffix, if backward TPs affect the ease of acquisition of remote co-occurrences, higher ratios should enhance a grammaticality effect. Comparing the two ratios reveals that there is a great distributional advantage for diminutive over plural dependencies in infants' input, suggesting that the strength of the effect of backward TPs may be greater for diminutive over plural dependencies. All three types of distributional statistics examined so far thus reveal statistical patterns present in the input that could be driving the difference the effect of grammaticality between diminutive and plural dependencies in Experiment 1 and 2.

*Average distance between the non-adjacent elements.*

Another distributional cue potentially influencing remote dependency learning is the distance between the two non-adjacent elements in an utterance. With respect to the diminutive dependencies, the mean distance between elements of a grammatical dependency (i.e. *het X-je*) was 1.81 syllables (SEM = 0.039) and between elements of an ungrammatical diminutive dependency (i.e. *de X-je*) 3.90 syllables (SEM = .167). In the plural dependencies, the elements were separated by an average of 2.24 syllables (SEM = 0.025) when the dependency was grammatical (i.e. *de X-en*) and by an average of 2.60 syllables (SEM = 0.034) when the ungrammatical was ungrammatical (i.e. *het X-en*; see Table 4). A two-way ANOVA with Dependency Type (diminutive vs. plural) and Grammaticality (grammatical vs. ungrammatical

determiner) as independent factors revealed main effects of both Dependency Type ( $F(1,4729) = 21.994; p < .001$ ) and Grammaticality ( $F(1,4729) = 172.64; p < .001$ ), but more importantly, these main effects were modulated by the interaction between these two factors ( $F(1,4729) = 85.543; p < .001$ ). Planned comparisons indicated that the difference in number of intervening syllables differed between grammatical and ungrammatical dependencies in both the diminutive and the plural dependencies ( $t(873) = -12.69; p < .001$  and  $t(3274) = -8.578; p < .001$ , respectively).

Regardless of whether the dependency was a diminutive or a plural, the number of syllables intervening the two distant elements was higher in the ungrammatical than in the grammatical co-occurrences. The size of the effect, however, was different, with the discrepancy being larger in the diminutive than in the plural dependency. Given that children may initially have difficulty tracking non-adjacent dependencies over a larger domain (Santelmann & Jusczyk, 1998), fewer intervening syllables would be advantageous for the sensitivity to non-adjacent dependencies. Thus, the greater the relative difference between grammatical and ungrammatical co-occurrences (with the average distance being longer for the ungrammatical than for the grammatical dependencies), the easier the larger a potential grammaticality effect is expected to be. This is therefore the fourth converging statistic showing that distributional cues may explain why infants display an effect of grammaticality for diminutive but not plural dependencies.

#### *Variability in the intervening syllables.*

The final distributional statistic on which we compare diminutives and plurals is the variability in the intervening syllables. For both the grammatical and the ungrammatical realizations of each of the dependency types, the number of unique and non-unique instances were examined to create a measure of the heterogeneity of variance (see Soderstrom, Conwell, Feldman, & Morgan, 2009 for discussion). With regard to the diminutive dependency, a total of

307 unique and 518 non-unique intervening syllables (uniqueness ratio:  $307/518 = .593$ ) were found for the grammatical co-occurrence pattern versus a total of 46 unique and 4 non-unique intervening syllables (uniqueness ratio:  $46/4 = 11.5$ ; see Table 5 for an overview) for the ungrammatical co-occurrence pattern. A two-way Chi-square analysis with Grammaticality (grammatical vs. ungrammatical determiner) and Uniqueness (unique vs. non-unique) revealed a significant association between the two factors ( $\chi^2(1) = 58.80$ ;  $p < 0.001$ ). As for the plural dependency, there were a total of 1138 unique and 1201 non-unique intervening syllables (uniqueness ratio:  $1138/1201 = .948$ ) in the grammatical co-occurrences versus a total of 1168 unique and 536 non-unique syllables (uniqueness ratio:  $1168/536 = 2.179$ ) in the ungrammatical co-occurrence pattern. Again, the two-way Chi-square indicated that there was a significant association between the two factors ( $\chi^2(1) = 159.17$ ;  $p < 0.001$ ). The odds ratio for the diminutive dependency ( $.539/11.5 = .05$ ) was, however, much smaller than the odds ratio for the plural dependency ( $0.948/2.179 = .44$ ).

These analyses indicate that the intervening material is more variable in the ungrammatical than in the grammatical dependencies. Given that more variability has been shown to increase the ease of acquisition of a dependency in artificial languages (Gómez, 2002), the reduced variability in grammatical compared to ungrammatical dependencies may impede the acquisition of the (grammatical) diminutive and plural dependencies. Moreover, the smaller odds ratio for the diminutive than for the plural dependency suggests that this disadvantage is especially exhibited in the diminutive dependency. Based on the level of variability of the interceding materials alone, infants' sensitivity to the (grammatical) diminutive dependency should be delayed compared to their sensitivity to the (grammatical) plural dependency. This is thus the only distributional cue not convergent with the behavioral data.

*Discussion*

Our corpus study reveals crucial differences in the distributional statistics between diminutive and plural dependencies. Four out of the five examined distributional cues indicate an advantage for the diminutive over the plural dependency. More specifically, the frequency, forward and backward TPs, and the average distance between the non-adjacent elements all suggest that an effect of grammaticality should be expected for diminutive before plural dependencies. Although the degree of variability in the intervening material does not follow this same pattern, the combination of these four factors in favor of the diminutive dependency may outweigh any effect the intervening syllable may have. Alternatively, as will be discussed in more detail in the General Discussion, the role of variability in the intervening material for non-adjacent dependency acquisition in natural languages may be different from its role in artificial language studies. Nonetheless, consistent with the idea that the plural *-en* is less informative than diminutive *-je*, the distributional properties of infants' typical input provide more reliable cues for the diminutive than for the plural dependency.

## General discussion

This study is the first to show convergence between distributional properties of the input and differences in rate of acquisition of non-adjacent dependencies in natural languages. Experiment 1 demonstrated that some time between 17 and 24 months of age, Dutch learners start tracking the non-adjacent diminutive dependency, an age close to when English learners first exhibit sensitivity to non-adjacent dependencies involving verbs (Santelmann & Jusczyk, 1998). Since infants in our study were presented with nonwords and had hence never heard the specific word sequences before, they must have abstracted away from any stored familiar word strings and created a dependency between the determiner and suffix, independent of the intervening material.

Experiment 2 examined the acquisition pattern of the plural dependency. While the remote diminutive and plural dependency in Dutch are structurally similar, they differ in other aspects, such as the variability in syntactic functions of the plural word-final *-en*, but not the diminutive *-je*. In contrast to the diminutive dependency, Dutch learners do not show evidence of having developed sensitivity to the remote plural dependency by 24 months of age, indicating that sensitivity to one non-adjacent dependency does not necessarily imply sensitivity to another non-adjacent dependency.

In the follow-up corpus study, we examined whether this difference in ease of acquisition is related to distributional patterns in infants' input. Artificial language studies have proposed that dependencies are acquired more readily when they are more frequent, when the TPs are higher, and when the distance between the elements is shorter. The relative frequency of grammatical versus ungrammatical dependencies was indeed found to be higher for diminutives than for plurals, as were both types of TPs. In addition, the ratio of the average distance between grammatical and ungrammatical diminutive dependencies is more advantageous for diminutives than plural dependencies. Taken together, these distributional statistics in the input show an overall benefit for the diminutive over the plural dependencies. If infants made use of these statistics, they would be expected to acquire the diminutive dependency more readily than the plural dependency. The observation of this exact pattern in Experiments 1 and 2 indicates that statistical regularities in the input could quite elegantly explain our obtained behavioral results and suggests that there may be a close link between distributional patterns in the input and the ease of acquisition of (non-adjacent) dependencies.

The only statistic that does not reveal an advantage for diminutive over plural dependencies is the degree of variability within the intervening words. The syllables interceding

plural dependencies were much more variable than the syllables interceding diminutive dependencies. Although not conforming to the same pattern of results as the other four distributional statistics, the higher degree of variability within the interceding words for the plural rather than the diminutive dependency is, in fact, not surprising. The remote combination of *de* and *-en* is, in the vast majority of the cases, not indicative of a plural dependency; it stands to reason that in such cases, when two non-adjacent elements do not form a dependency, the variability in the intervening materials logically increases compared to those cases where the two elements do form a dependency. How does the variability of these middle elements then affect non-adjacent dependency learning? Greater variability in the intervening materials has been shown to help infants track non-adjacent dependencies (Gómez, 2002). With a wider range of middle syllables, the TPs between adjacent elements decrease, causing infants to redirect their attention towards the more informative non-adjacent dependencies. Compared to natural languages, however, artificial languages greatly reduce the overall variability and thereby oversimplify infants' typical input (cf. Johnson & Tyler, 2010). While *some* variability may indeed be required for generalization, the presence of *excessive* variability may impede this generalization process, allowing a U-shaped relationship between variability in the input and infants' capacity to track distant dependencies.<sup>7</sup> Alternatively, increased variability might be beneficial when the variability is relatively low, but a threshold could, quite plausibly, limit this advantage. Once this threshold is reached, children shift their attention towards distant dependencies regardless of the extent to which the threshold is exceeded. The variability within the intervening materials for Dutch diminutive and plural dependencies may be of such magnitude that the benefit of having a wider range of intervening materials is masked under the high overall level of variability. This raises the question of whether this type of variability in the intervening material does play a role in

dependency acquisition in natural languages or whether its observed effect in artificial language experiments is mainly caused by the small lexicons used in these studies.

The finding that the acquisition of natural language dependencies may possibly emerge as a function of the distributional statistics in infants' typical language input raises questions regarding the nature of infants' early sensitivity to the remote diminutive dependency. First, the effect could be grammatical. Under this view, Dutch learners may have acquired that word-final *-je* (or, abstracted even more, a diminutive) can be preceded not only by the definite article *het*, but also by other neuter-gender determiners (e.g. *dit* 'this' or *dat* 'that'). Second, the effect could be statistical in nature. According to this view, it is the statistical regularity between the specific two elements (i.e. *het* and *-je*) that forms the dependency and sensitivity to this co-occurrence does not imply generalization to other neuter determiners (see Dahan, Swingley, Tanenhaus, & Magnuson, 2004; Van Heugten & Shi, 2009 for discussion). In addition, the corpus study allows us to consider a further closely related possibility. In line with previous studies (Höhle et al., 2006; Santelmann & Jusczyk, 1998), we considered 24-month-olds' longer orientation times to grammatical as opposed to ungrammatical trials in Experiment 1 to mean that infants have gained sensitivity to the diminutive dependency. Similarly, the absence of differences in orientation times in Experiment 2 was taken as evidence for infants' insensitivity to plural dependencies. However, these results may not speak to the acquisition of the specific structures per se; rather, the behavioral results may merely reflect the greater distributional advantage for grammatical (relative to ungrammatical) dependencies in the diminutives than in the plurals. That is, because the extent to which the distributional cues favor the grammatical over the ungrammatical dependency is so substantial for diminutives, infants may display a listening preference to these grammatical dependencies that is solely driven by a preference for those patterns that best mirror the input. The

distributional differences between grammatical and ungrammatical plural dependency, in contrast, are less extreme and hence infants are equally attentive to grammatical and ungrammatical dependencies. This latter interpretation still allows distributional statistics to play a crucial role, while remaining agnostic over the acquisition of the (grammatical) diminutive and plural dependencies.<sup>8</sup> Thus, independent of which interpretation may be correct, our findings provide clear evidence for infants' sensitivity to the distributional patterns in their input.

Although this study does not allow us to form strong conclusions regarding the type of knowledge driving the effect, the overwhelming advantage for diminutive over plural dependencies in the input suggests that the distributional statistics may at least be a first cue assisting infants tune into the co-occurrence between *het* and *-je*. Whether or not this knowledge has developed into more abstract grammatical knowledge by two years of age is an empirical question and should be examined in future work. If infants, for example, would prefer listening to frequently co-occurring yet ungrammatical items over less frequently co-occurring grammatical items, this would suggest the effect is closely tied to the statistical distribution of the surface form of language and would hence be statistical rather than grammatical in nature.

A major strength of the current study is the use of both behavioral and corpus data to examine the link between infants' input and the acquisition of specific linguistic structures. However, because the corpus transcription is only a simplified representation of the input received by the child, this approach is by no means perfect. An analysis of the orthographic transcriptions of infants' input cannot capture all factors that may play a role in non-adjacent dependency learning. The prosody of the utterance a dependency is embedded in, for example, quite plausibly affects the rate of acquisition. Prosodic phrases have been considered as the main processing units, even at the very early stages (Christophe, Guasti, Nespor, Dupoux, & Van

Ooyen, 1997; Christophe, Mehler, & Sebastián-Gallés, 2001; Gout, Christophe, & Morgan, 2004; Johnson & Seidl, 2008; Soderstrom, Kemler Nelson, & Jusczyk, 2005; see Van Heugten & Shi, 2010, however, for evidence that infants are able to compute non-adjacent dependencies across minor prosodic boundaries). If a prosodic phrase boundary limits the domain over which statistics are computed (see Shukla, Nespor, & Mehler, 2007), then the lack of prosodic information in the corpus might have skewed our findings to some extent. Another factor that orthographic transcriptions cannot capture concerns the lack of acoustic information in our corpus analyses. Infants, for example, seem to be sensitive to subphonemic differences between phonologically identical nouns and verbs (Conwell & Morgan, 2007), which may potentially help infants distinguish between grammatical and spurious dependencies. Sensitivity to acoustic information could also affect non-adjacent dependency learning for a further reason. Where word-final consonants are often fully resyllabified when a plural suffix is added, this resyllabification process is much more subtle for diminutives. Given that infants are sensitive to subtle subphonemic differences (Johnson, 2008b; McMurray & Aslin, 2005) and that extracting noun stems from suffixed nouns is more challenging when resyllabification occurs (Johnson, 2008a), it seems reasonable that degree of syllabification affects rate of non-adjacent dependency acquisition. A combination of many convergent measures, including acoustical analyses of the input and a better understanding of children's ability to perceive and process acoustic variation would thus be tremendously useful to better understand infants' acquisition of non-adjacent dependencies. Although the corpus does not allow us to do (and was not meant for doing) prosodic and acoustical analyses, technological advances could lead to future developments of corpora that do retain these types of information. Nonetheless, the finding that distributional patterns in the input nicely predict the order of acquisition even when prosodic and acoustic

factors are not taken into account is an impressive display of the powerful role that distributional factors may play in the acquisition process.

In sum, this study is not the first to examine distributional patterns in infants' language input (Mintz, 2003; Chemla et al., 2009), nor is it the first to examine infants' behavioral responses to dependencies embedded in natural speech (Höhle et al., 2006; Santelmann & Jusczyk, 1998) or in artificial languages (e.g., Aslin et al., 1998; Bonatti, et al., 2005; Gómez, 2002; Newport & Aslin, 2004; Saffran et al., 1996a, 1996b). As far as we are aware, however, this study *is* the first to combine all of these methodologies in a single study (i.e. natural language dependencies intervened by nonsense words, linked to infants' typical input), which enables us to demonstrate that even in natural language development, a more advantageous distributional pattern may assist infants' early dependency acquisition. That is, although infants have been shown to form relationships between distant function morphemes at a young age, this ability appears to be modulated by a variety of distributional cues, present in the ambient language. Where artificial languages provide a well-controlled approach to studying specific learning mechanisms and constraints thought to be relevant to the acquisition of distributional dependencies, only a combination of behavioral experiments and a corpus analysis allows us to examine the use of this information in natural language acquisition. Further combining these tools in future studies holds great potential for better understanding the link between distributional patterns in the input and children's acquisition of natural language dependencies.

## References

- Aslin, R. N., Saffran, J. R. and Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, 9, 321-324.
- Bonatti, L. L., Peña, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on statistical computations: The role of consonants and vowels in continuous speech processing. *Psychological Science*, 16, 451-459.
- Braine, M. D. S., Brody, R. E., Brooks, P. J., Sudhalter, V., Ross, J. A., Catalano, L., et al. (1990). Exploring language acquisition in children with a miniature artificial language: Effects of item and pattern frequency, arbitrary subclasses, and correction. *Journal of Memory and Language*, 29, 591-610.
- Chemla, E., Mintz, T. H., Bernal, S., & Christophe, A. (2009). Categorizing words using ‘frequent fames’: what crosslinguistic analyses reveal about distributional acquisition strategies. *Developmental Science*, 12, 397-407.
- Christophe, A., Guasti, T., Nespor, M., Dupoux, E., & Van Ooyen, B. (1997). Reflections on phonological bootstrapping: Its role for lexical and syntactic acquisition. *Language and Cognitive Processes*, 12, 585-612.
- Christophe, A., Mehler, J., & Sebastián-Gallés, N. (2001). Perception of prosody boundary correlates by newborn infants. *Infancy*, 2, 385-394.
- Conwell, E., & Morgan, J. L. (2007). Resolving grammatical category ambiguity in acquisition. In H. Caunt-Nulton, S. Kulatilake, & I. Woo (Eds.), *Proceedings of the 32th annual Boston University Conference on Language Development: Vol. 1* (pp. 117-128). Somerville, MA: Cascadilla Press.

- Dahan, D., Swingley, D., Tanenhaus, M.K., & Magnuson, J.S. (2000). Linguistic gender and spoken word recognition in French. *Journal of Memory and Language*, *42*, 465-480.
- Frigo, L., & McDonald, J. L. (1998). Properties of phonological markers that affect the acquisition of gender-like subclasses. *Journal of Memory and Language*, *39*, 218-245.
- Gerken, L. A., & McIntosh, B. (1993). The interplay of function morphemes and prosody in early language. *Developmental Psychology*, *29*, 448-457.
- Gervain, J. (2008). *From the richness of the signal to the poverty of the stimulus: Mechanisms of early language acquisition*. Unpublished doctoral dissertation, SISSA, Trieste, Italy.
- Gillis, S. (1997). The acquisition of diminutives in Dutch. In W. Dressler (Ed.), *Studies in pre- and protomorphology* (pp. 165-179). Vienna: Verlag der Österreichischen Akademie der Wissenschaften.
- Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*, 431-436.
- Gómez, R. L., & Gerken, L. A. (1999). Artificial grammar learning by one-year-olds leads to specific and abstract knowledge. *Cognition*, *70*, 109-135.
- Gómez, R. L., & Lakusta, L. (2004). A first step in form-based category abstraction in 12-month-old infants. *Developmental Science*, *7*, 567-580.
- Gómez, R. L., & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy*, *7*, 183-206.
- Gout, A., Christophe, A., & Morgan, J. L. (2004). Phonological phrase boundaries constrain lexical access: II. *Infant data*. *Journal of Memory and Language*, *51*, 547-567.
- 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.

Höhle, B., Weissenborn, J., Kiefer, D., Schulz, A., & Schmitz, M. (2004). Functional elements in infants' speech processing: The role of determiners in segmentation and categorization of lexical elements. *Infancy*, 5, 341-353.

Johnson, E. K. (2005). Grammatical gender and early word recognition in Dutch. In A. Brugos, M.R. Clark-Cotton & S. Ha (Eds.), *Proceedings of the 29th annual Boston University Conference on Language Development: Vol. 1* (pp. 320-330). Somerville, MA: Cascadilla Press.

Johnson, E. K. (2008a). *Dutch infants' segmentation of diminutive word forms: An exploration of morpheme stripping and lexical embedding*. Paper presented at the 16th International Conference on Infant Studies, Vancouver, Canada.

Johnson, E. K. (2008b). Infants use prosodically conditioned acoustic-phonetic cues to extract words from speech. *Journal of the Acoustical Society of America*, 123, EL144-EL148.

Johnson, E. K., & Seidl, A. (2008). Clause segmentation by 6-month-old infants: A cross-linguistic perspective. *Infancy*, 13, 440-455.

Johnson E. K., & Tyler, M. D. (2010). Testing the limits of statistical learning for word segmentation. *Developmental Science*, 13, 339-345.

Kedar, Y., Casasola, M., & Lust, B. (2006). Getting there faster: 18- and 24-month-old infants' use of function words to determine reference. *Child Development*, 77, 325-388.

Kemler Nelson, D. G., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A., & Gerken, L. A. (1995). The head-turn preference procedure for testing auditory perception. *Infant Behavior and Development*, 18, 111-116.

Lany J., & Gómez, R. L. (2008). Twelve-Month-Old Infants Benefit From Prior Experience in Statistical Learning. *Psychological Science*, 19, 1247-1252.

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, 18, 193-198.

Marquis, A., & Shi, R. (2008). Segmentation of verb forms in preverbal infants. *Journal of the Acoustical Society of America*, 123, EL105-EL110.

McMurray, B., & Aslin, R.N. (2005). Infants are sensitive to within-category variation in speech perception. *Cognition*, 95, B15-B26.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk. Third Edition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Mintz, T. H. (2002). Category induction from distributional cues in an artificial language. *Memory & Cognition*, 30, 678-686.

Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition*, 90, 91-117.

Mirman, D., Graf Estes, K. & Magnuson, J. S. (in press). Computational modeling of statistical learning: Effects of transitional probability versus frequency and links to word learning. *Infancy*.

Morgan, J. L., Meier, R. P., & Newport, E. L. (1987). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology*, 19, 498-550.

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, 48, 127-162.

- Onnis, L., Monaghan, P., Richmond, K., & Chater, N. (2005). Phonology impacts segmentation in online speech processing. *Journal of Memory & Language*, *53*, 225-237.
- Peña, M., Bonnatti, L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science*, *298*, 604-607.
- Pelucchi, B., Hay, J. F. & Saffran, J. R. (2009). Learning in reverse: Eight-month-old infants track backwards transitional probabilities. *Cognition*, *113*, 244-247.
- Perruchet, P., & Desauty, S. (2008). A role for backward transitional probabilities in word segmentation? *Memory & Cognition*, *36*, 1299-1305.
- Perruchet, P., & Peereman, R. (2004). The exploitation of distributional information in syllable processing. *Journal of Neurolinguistics*, *17*, 97-119.
- 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*, 172-196.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996a). Statistical learning by 8-month old infants. *Science*, *274*, 1926-1928.
- Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996b) Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*, 606-621.
- Santelmann, L. M., & Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition*, *69*, 105-134.
- Shi, R. & Melançon, A. (in press). Syntactic organization in French-learning infants. *Infancy*.
- Shipley, E., Smith, C., & Gleitman, L. (1969). A study in the acquisition of language. *Language*, *45*, 322-342.

- Shukla, M., Nespors, M., & Mehler, J. (2007). An interaction between prosody and statistics in the segmentation of fluent speech. *Cognitive Psychology*, 54, 1-32.
- Soderstrom, M., Conwell, E., Feldman, N., & Morgan, J. L. (2009). The learner as statistician: three principles of computational success in language acquisition. *Developmental Science*, 12, 409-411.
- Soderstrom, M., Kemler Nelson, D. G., & Jusczyk, P. W. (2005). Six-month-olds recognize clauses embedded in different passages of fluent speech. *Infant Behavior and Development*, 28, 87-94.
- Van Heugten, M., & Johnson, E. K. (in press). Gender-marked determiners help Dutch learners' word recognition when gender information itself does not. *Journal of Child Language*.
- Van Heugten, M., & Shi, R. (2009). French-learning toddlers use gender information on determiners during word recognition. *Developmental Science*, 12, 419-425.
- Van Heugten, M. & Shi, R. (2010). *Infants track non-adjacent dependencies across phonological phrase boundaries*. Poster presented at the 17<sup>th</sup> International Conference on Infant Studies, Baltimore, MD.
- Wijnen, F., & Verris, M. (1998). The acquisition of Dutch syntax. In S. Gillis & A. De Houwer (Eds.), *The Acquisition of Dutch* (pp. 223-299). Amsterdam: Benjamins.
- Wilsenach, A. C. (2006). *Syntactic Processing in Developmental Dyslexia and in Specific Language Impairment: A study on the acquisition of the past participle construction in Dutch*. Unpublished Ph.D. dissertation, University of Utrecht, the Netherlands.
- Yang, C. D. (2004). Universal grammar, statistics, or both? *Trends in Cognitive Science*, 8, 451-456.
- Zangl, R., & Fernald, A. (2007). Increasing flexibility in children's online processing of

grammatical and nonce determiners in fluent speech. *Language Learning and Development*,  
3, 199-231.

### Author Notes

This research was funded by an NWO Spinoza Grant awarded to Anne Cutler, University of Toronto startup funds to EKJ, and a Prins Bernhard Cultuurfondsbeurs to MvH. We thank Angela Khadar, Judith Hanssen, and the rest of the Baby Research Centre, as well as participating families, for enabling the completion of this study, Bruce Schneider for statistical assistance, and Cara Tsang and three anonymous reviewers for insightful comments on an earlier draft of this paper. A portion of this work was previously presented at the 16th and 17th International Conference on Infant Studies (Vancouver, 2008; Baltimore, 2010) and the 11th International Congress for the Study of Child Language (Edinburgh, 2008).

## Footnotes

<sup>1</sup> Fatigue potentially prevented this interaction to be significant at the .05-level. Without the last two test trials, the interaction is significant with respect to the traditional standards ( $F(1, 54) = 7.615; p = .008$ ).

<sup>2</sup> Without excluding the first two trials, the two-tailed t-test is significant at the .031 level.

<sup>3</sup> Including the first two trials in the analysis does not change the (absence of the) effect.

<sup>4</sup> Neither including the first two trials nor excluding the last two trials in the analysis changes the (absence of the) effect.

<sup>5</sup> The pattern of results was independent of the number of intervening syllables. That is, results were similar regardless of whether we excluded all dependencies with more than three, four, or five intervening syllables.

<sup>6</sup> For the sake of simplicity, we refer to the dependencies as grammatical versus ungrammatical. Note however, that the observed patterns do not necessarily involve dependencies (the co-occurrence of a determiner and a suffix does not intrinsically mean that the two belong together), and that what is referred to as ungrammatical in terms of a dependency may, in fact, be grammatical (based on the co-occurrence of *het* and word-final *-en*, for example, *het eten*, ‘the food’) would be classified as an ungrammatical plural dependency even though it does, in fact, not involve a plural and forms a grammatical NP).

<sup>7</sup> We would like to thank an anonymous reviewer for this suggestion.

<sup>8</sup> We would like to thank an anonymous reviewer for this suggestion.

Table 1

*Frequency: frequency of grammatical and ungrammatical co-occurrences broken down by diminutive and plural dependencies.*

---

Diminutives (suffix <i>-je</i> )	frequency
grammatical article ( <i>het</i> )	825
ungrammatical article ( <i>de</i> )	50

---

Plurals (suffix <i>-en</i> )	
grammatical article ( <i>de</i> )	2,339
ungrammatical article ( <i>het</i> )	1,704

---

Table 2

*Forward TPs: frequency of occurrence of each of the articles followed and not followed by the dependency suffix broken down by diminutive and plural dependencies.*

Diminutives (suffix <i>-je</i> )		
	dependency	no dependency
grammatical article ( <i>het</i> )	825	5,814
ungrammatical article ( <i>de</i> )	50	6,518
Plurals (suffix <i>-en</i> )		
grammatical article ( <i>de</i> )	2,339	4,229
ungrammatical article ( <i>het</i> )	1,704	4,935

Table 3

*Backward TPs: frequency of occurrence of the dependency suffix preceded by the grammatical definite article, the ungrammatical definite article, or no definite article in both the diminutive and the plural dependency.*

---

	grammatical article	ungrammatical article	other
diminutives (suffix <i>-je</i> )	825	50	4,112
plurals (suffix <i>-en</i> )	2,339	1,704	24,872

---

Table 4

*Distance: the average number (standard error of the mean in brackets) of unique and non-unique grammatical and ungrammatical co-occurrences broken down by diminutive and plural dependencies.*

Diminutives (suffix <i>-je</i> )	Mean (SEM) number of intervening syllables
grammatical article ( <i>het</i> )	1.81 (.039)
ungrammatical article ( <i>de</i> )	3.90 (.167)
Plurals (suffix <i>-en</i> )	
grammatical article ( <i>de</i> )	2.24 (.025)
ungrammatical article ( <i>het</i> )	2.60 (.034)

Table 5

*Variability: the total number of unique and non-unique material intervening grammatical and ungrammatical co-occurrences broken down by diminutive and plural dependencies.*

Diminutives (suffix <i>-je</i> )	unique	non-unique
grammatical article ( <i>het</i> )	307	518
ungrammatical article ( <i>de</i> )	46	4
Plurals (suffix <i>-en</i> )		
grammatical article ( <i>de</i> )	1138	1201
ungrammatical article ( <i>het</i> )	1168	536

*Figure 1.* Mean orientation time (and SEM) towards the lists of grammatical and ungrammatical NPs in Experiment 1, 2, and 3.

Figure 1

