

Acquisition as a window on the nature of NPIs

Jing Lin – *University of Amsterdam*

Fred Weerman – *University of Amsterdam*

Hedde Zeijlstra – *University of Göttingen*

Abstract. Dutch modal verb *hoeven* ‘need’ is a negative polarity item (NPI) (Zwarts 1981, Hoeksema 2000), which survives in all anti-additive, and some but not all downward entailing (DE) contexts. The aim of the paper is to explore the reason why Dutch *hoeven* is not allowed in all DE-contexts – as observed for NPIs such as *any*-terms. We answer this question by looking at acquisition. The reasoning is straightforward: the analysis underlying a linguistic phenomenon is a product of children’s acquisition of it. Data collected from a total of 132 monolingual Dutch children (2;09–5;10; M = 4;04; SD = 9.3 months) in an elicited imitation task demonstrate a learning path of *hoeven* in which children start with two lexical frames [HOEF NIET] ‘NEED NOT’ and [HOEF GEEN] ‘NEED NO’ and switch to an abstract analysis of it later on: [HOEF NEG] ‘NEED NEG’. Given this abstract analysis, emerging as a result of language acquisition, we argue that *hoeven* is an NPI because of its lexical dependency with the abstract negation NEG (cf. Postal 2000). This in turn explains the distribution of the Dutch NPI restricted to some but not all DE-contexts: *hoeven* is only allowed in those DE-contexts that incorporate the abstract negation NEG.

1. Introduction

Dutch modal verb *hoeven* ‘need’ is a negative polarity item (NPI), which occurs in negative contexts only (Zwarts 1981, Hoeksema 2000, Van der Wouden 1997). Similar to English *any*-terms, well-described in the literature, *hoeven* is restricted to downward entailing (DE) contexts (cf. Ladusaw 1979). For instance, *hoeven* is licensed by the sentential negative marker *niet* ‘not’, as in (1a), negative indefinites such as *niemand* ‘nobody’ as in (1b), semi-negative expressions just like *nauwelijks* ‘seldom’ as in (1c), exclusive adverbs such as *alleen* ‘only’ as in (1d). However, *hoeven* is not licensed in all DE-contexts that license *any*-terms. In conditional clauses or the restriction of a universal quantifier, which sanction *any*-terms, for instance, *hoeven* is ungrammatical as shown in (2a) and (2b), respectively.

- (1) a. Jan hoeft niet te koken.
John needs not to cook
‘John does not need to cook.’
- b. Niemand hoeft te koken.
nobody needs to cook
‘Nobody needs to cook.’
- c. Jan hoeft nauwelijks te koken.
John needs seldom to cook
‘John seldom needs to cook.’
- d. Jan hoeft alleen te koken.
John needs only to cook
‘John only needs to cook.’

- (2) a. *Iedereen die hoeft te koken moet nu beginnen.
 everyone that needs to cook must now start
 Intended: ‘Everyone that needs to cook must start now.’
- b. *Als Jan hoeft te koken moet hij nu beginnen.
 if John needs to cook must he now start
 Intended: ‘If John needs to cook, he must start now.’

In simple affirmative contexts, the appearance of *hoeven* is ungrammatical (Hoeksema 1994, 2000, Van der Wouden 1997, among others) – as is observed for all NPIs. See below:

- (3) *Jan hoeft te koken.
 John needs to cook
 Intended: ‘John needs to cook.’

The aim of this paper is to explore the reason why Dutch *hoeven* exhibits a distributional pattern that is restricted to some but crucially not all DE-contexts. In other words, how is *hoeven* represented in the grammar such that Dutch speakers only use it the way described above?

In order to answer this question, we will look at acquisition. The reasoning is straightforward: the analysis underlying a linguistic phenomenon is a product of children’s acquisition of it. By analysing children’s performance in an elicited imitation task (N=132; 2;09–5;10; M = 4;04; SD = 9.3 months), this paper presents a learning path of *hoeven* from two lexical frames [HOEF NIET] and [HOEF GEEN] to an abstract analysis [HOEF NEG]. Moreover, the paper shows what the acquisition data can tell us about the nature of *hoeven* – an atypical NPI in terms of distribution. Under the hypothesis that [HOEF NEG] is the analysis that emerges as a result of acquisition, *hoeven*’s distribution restricted to some but crucially not all DE-contexts is explained as a consequence of its lexical association with the abstract negation NEG (cf. Postal 2000), since NEG is incorporated in merely some but not all DE-contexts (Iatridou and Zeijlstra 2013).

The paper is organised as follows. We start out with a brief introduction to the various negative environments that may license NPIs (Section 2). Next, we introduce our experiment: the elicited imitation task (Section 3). Afterwards, we present our regression results (Section 4) and analysis (Section 5), which are followed by discussion and conclusion (Section 6).

2. Negative contexts

Ladusaw (1979) proposes that NPIs are generally licensed in DE-contexts: contexts in which the entailment relation goes from set to subset (see also Fauconnier 1975, 1978). DE-contexts can be further divided into three types, depending on their logico-semantic behaviours: anti-morphic contexts, anti-additive contexts, and DE-contexts (Zwarts 1981, 1986, 1995).¹ These contexts –

¹ For every arbitrary X, Y : iff $f(X \cap Y) \Leftrightarrow f(X) \cup f(Y)$ and $f(X \cup Y) \Leftrightarrow f(X) \cap f(Y)$, then the function f is anti-morphic; iff $f(X \cup Y) \Leftrightarrow f(X) \cap f(Y)$, then the function f is anti-additive. These definitions are adapted from Van der Wouden (1994).

as proved by Zwarts – stand in a subset relationship with each other. For instance, all anti-morphic environments are anti-additive and DE, but not vice versa. In the context of the Dutch NPI, it is relevant to know the following. First, the sentential negative marker *niet* is anti-morphic. Second, negative indefinites such as *niemand* are anti-additive but not anti-morphic. Finally, semi-negative expressions such as *nauwelijks* and exclusive adverbs such as *alleen* are merely DE.²

3. Experiment

3.1. Method

In order to access children's acquisition of the Dutch NPI *hoeven*, we carried out an elicited imitation task. In an elicited imitation task, participants are required to first listen carefully to (pre-recorded) stimuli and then repeat the stimuli as exactly as they heard it (Lust et al. 1996, Vinther 2002). When repeating a stimulus as precisely as was heard, participants are claimed to construct their own mental representation of it according to their own grammatical system established so far (Chomsky 1964, Eissenbeiss 2010, Keenan and Hawkins 1987, Panitsa 2001, Scholl and Ryan 1980). If a stimulus is in agreement with their own grammar, participants repeat the stimulus immediately after hearing it (Scholl and Ryan 1980); whereas they correct it in accordance with their own grammar, or do not repeat the stimulus if it is ungrammatical based on their grammar of the target language (Brown 1973, Keeney and Wolfe 1972, Vinther 2002).

3.2. Conditions

The experiment included five DE-operators, which license the Dutch NPI *hoeven*: *niet* 'not' (four stimuli), *geen* 'no(ne)' (two stimuli), *niemand* 'nobody' (two stimuli), *weinig* 'few' (two stimuli), and *alleen* 'only' (two stimuli). All these five operators are acquired by children around age three (Van der Wal 1996: Table 4.1). The reason for this selection was that they represent different types of DE-contexts: *niet* is anti-morphic; *geen* and *niemand* are anti-additive; *weinig* and *alleen* are only DE (cf. Section 2). This manipulation enabled us to explore the contribution of the semantic knowledge of various negative contexts to the acquisition of the NPI. In order to examine whether children are aware of the ungrammaticality of *hoeven* in simple affirmative contexts, we added four ungrammatical stimuli by placing *hoeven* in sentences like (3).

In addition to the six test conditions described above, the experiment also had filler conditions containing a total of twenty fillers. As to neutralise the effect that every test stimulus contained the same modal verb *hoeven*, half of the fillers were designed with a modal verb as well, of which six involved *willen* 'will' and four involved *kunnen* 'can'. Both *willen* and *kunnen* are polarity-insensitive: they are neither NPIs like *hoeven* nor PPIs (Positive Polarity Items) like *moeten* 'must' (cf. Iatridou and Zeijlstra 2013). Since the majority of the test stimuli containing the NPI *hoeven* were negative, half of the fillers were manipulated to be negative as well.

² We follow von Stechow (1999) and analyse exclusive adverbs as a specific kind of DE-operator: Strawson-DE, which he defines by making use of presuppositions.

Moreover, four out of the twenty fillers were ungrammatical. They all contained a syntactic error due to a non-application of the V2 rule in Dutch main clauses. An example is given below. We added these ungrammatical fillers to counterbalance the (un)grammaticality of the stimuli.

- (4) *Gisteren Jan met Marie in het park wandelde.
 yesterday Johnwith Mary in the park walked
 Intended: ‘Yesterday, John walked in the park with Mary.’

3.3. Stimuli

In an elicited imitation task, the length of stimuli must be controlled (Montgomery et al. 1978, among others) as to prevent children from giving a repetition response from memory alone without first establishing their own mental representations of stimuli. Stimuli need to be long enough to override children’s working memory capacity but short enough for comprehension because children must construct their own mental representations of them without omitting too many words. All of the test and filler stimuli in the current experiment contained ten words. This represents a medium length of stimuli according to Montgomery et al. (1978), which is neither too short nor too long for participants between age four and six.

Words appearing in the stimuli are attested in daily communication with children under age five. The stimuli only contained main clauses to ensure a similar syntactic complexity. Some examples of our stimuli are given below: (5) represents the licensing conditions by *niemand*; (6) is an example of unlicensed *hoeven*. Two examples of grammatical fillers – one with a modal and the other without – are given in (7a) and (7b), respectively.

- (5) Vandaag hoeft Beer aan niemand een potje honing te geven.
 today needs Pooh to nobody one jar honey to give
 Lit. ‘Pooh needs to give nobody a jar of honey today.’
 ‘Pooh does not need to give anybody a jar of honey today.’

- (6) *Beer hoeft samen met zijn vriendjes mooie liedjes te zingen.
 Pooh needs together with his friends nice songs to sing
 Intended: ‘Pooh needs to sing nice songs together with his friends.’

- (7) a. Met het slechte weer wil Beer niet naar buiten gaan.
 with the bad weather will Pooh not to outside go
 ‘Pooh will not go outside with the bad weather.’
 b. Met het koude weer draagt Beer alleen een blauwe sjaal.
 with the cold weather wears Pooh only a blue scarf
 ‘With the cold weather, Pooh only wears a blue scarf.’

The stimuli were pre-recorded in an MP3 recorder by a female native Dutch speaker. To minimise prosodic influence, the speaker recorded the stimuli as neutrally as possible. The presentation order of the stimuli was counterbalanced.

3.4. Categorisation of responses

Children's responses to the stimuli were divided into three main categories: *no response*, *imitation response*, and *non-imitation response*. The category of *no response* referred to the instances in which the child either did not give any response at all after hearing a stimulus or gave an irrelevant response such as *Heb 'm niet gehoord* 'I didn't hear it'.

A response was categorised as *imitation* when the participants imitated the stimuli. However, as the stimuli length was controlled such that the participants needed to first establish their own mental representations of the stimuli, it was hardly ever the case that the participants were able to repeat every single word in a stimulus. We thus focused only on how the participants reacted to the licensing of *hoeven* and defined *imitation* as follows. It referred to responses in which at least both the NPI *hoeven* and the manipulated licenser were repeated in the manipulated order.

The category of *non-imitation responses* was further divided into three subcategories: *substitution*, *omission*, and *addition*. Consider the test stimulus in (5) as an example. An instance of *substitution* was counted if the child substituted the manipulated licenser *niemand* with another licenser, e.g., *niet* in (8a); substituted the NPI with another verb, e.g., *gaat* 'goes' in (8b); or substituted both the NPI and the manipulated licenser by an alternative, as shown in (8c).

- (8) a. Vandaag hoeft Beer niet aan iemand een potje honing te geven.
 today needs Pooh not to somebody one jar honey to give
 'Pooh does not need to give a jar of honey to anybody today.'
- b. Vandaag gaat Beer aan niemand een potje honing geven.
 today goes Pooh to nobody one jar honey give
 Lit. 'Pooh is going to give nobody a jar of honey today.'
 'Pooh is not going to give anybody a jar of honey today.'
- c. Vandaag gaat Beer niet aan iemand een potje honing geven.
 today goes Pooh not to somebody one jar honey give
 Lit. 'Pooh is going to give nobody a jar of honey today.'
 'Pooh is not going to give anybody a jar of honey today.'

A *non-imitation response* was categorised as *omission* if the child omitted the NPI as in (9a); left out the manipulated licenser as in (9b); or omitted both of them as in (9c).

- (9) a. Vandaag Beer aan niemand een potje honing geven
 today Pooh to nobody one jar honey give
 'Pooh give nobody a jar of honey today'

- b. *Vandaag hoeft Beer een potje honing te geven.
 today needs Pooh one jar honey to give
 Intended: ‘Pooh needs to give somebody a jar of honey today.’
- c. Vandaag Beer een potje honing geven
 today Pooh one jar honey give
 ‘Pooh give (somebody) a jar of honey today’

A *non-imitation response* was categorised as *addition* if the child gave a grammatical response by adding a licenser for the NPI while confronted with a stimulus containing unlicensed *hoeven*. For example, an instance of *addition* was counted if the child gave (10) as a response to the ungrammatical stimulus (6), by adding *niet* to license the NPI.

- (10) Beer hoeft niet samen met zijn vriendjes mooie liedjes te zingen.
 Pooh needs not together with his friends nice songs to sing
 ‘Pooh does not need to sing nice songs together with his friends.’

3.5. Participants & Procedure

A total of 133 typically developing monolingual Dutch children (2;09–5;10; $M = 4;04$; $SD = 9.3$ months) recruited via day care centres and elementary schools in the Netherlands participated in the experiment. The experiment was conducted individually and took place at educational institutions. The procedure of the experiment was as follows. We first invited a child from a class for a game and then explained how the game would proceed and what we expected him or her to do. There were four trials for each child to become familiar with the experimenter and the experiment. If the child proved to understand what was expected of him or her after the trials, the experiment started. The experiment lasted an average of fifteen minutes for the four-year-olds, while the younger participants took five minutes more on average.

Two experimenters were present during the experiment. While one experimenter tested the child, the other experimenter filled in a score sheet, and recorded the child’s responses on an MP3 recorder for later transcription and analysis.³

4. Results

In order to model the acquisitional pathway of the Dutch NPI *hoeven*, we employed a general linear mixed-effect logistic regression analysis in R for each of the six test conditions. We assigned the value of 1 to all *imitation responses* and 0 to all *non-imitation responses* as well as in cases of *no response*. With the ages of the participants as the independent variable and their repetition scores (0 or 1) as the dependent variable, the regression analyses conducted on our cross-sectional data enabled us to generalise the developmental patterns of children’s knowledge

³ In the score sheet, the experimenter assigned the child’s responses to different categories and wrote down critical changes or corrections in the responses, when applicable.

of the NPI in different conditions. Results of the regression analyses are presented in the graph below.⁴ The x-axis represents the participants' age in terms of months, the y-axis indicates the repetition probability of the stimuli predicted by the regression models, and the interpolation line represents the mean value of the predicted repetition probabilities.

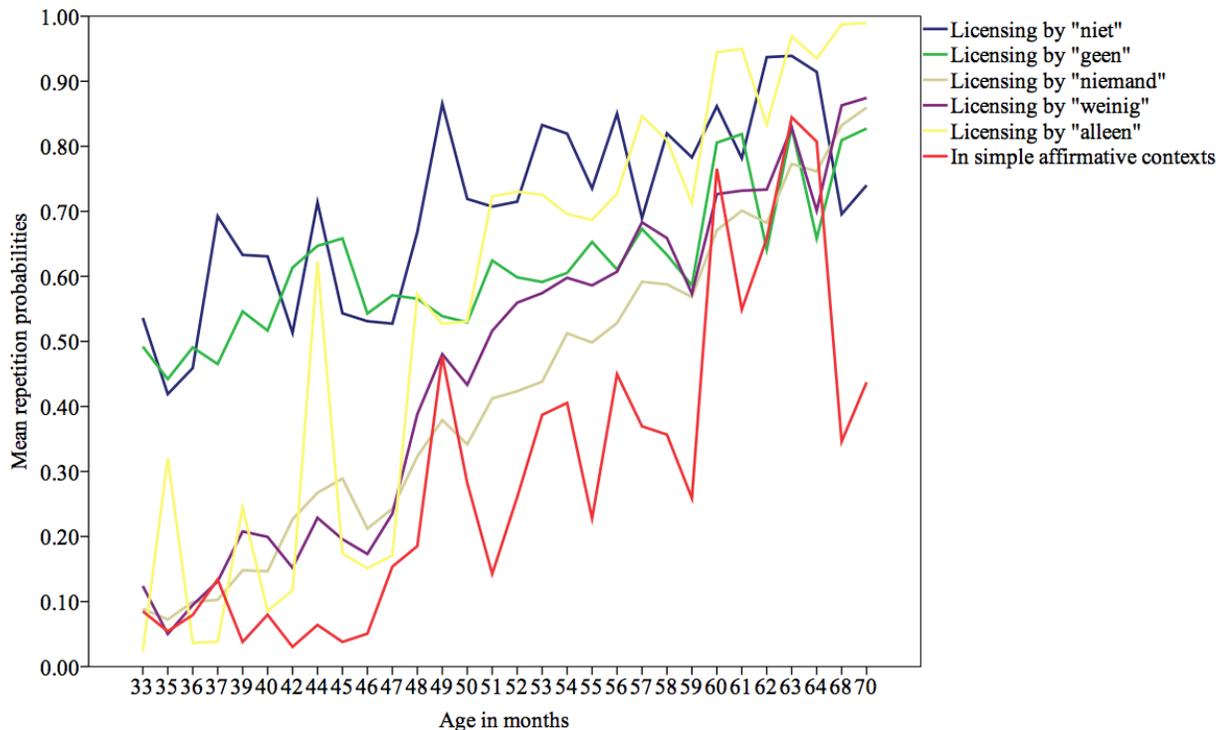


Figure 1: The development of children's knowledge of *hoeven* licensing over time

We start with the licensing condition by the sentential negative marker *niet*. As illustrated in Figure 1, children are predicted to give a repetition response in this test condition around 50% of the time at younger ages, i.e. below 3;04 (i.e. 40 months). However, our regression model attests a significant age effect in the development ($p < .001$), which means that children's repetition performance improves significantly with age. In particular, at 5;00 (i.e. 60 months) and older, the predicted probabilities for children's repetition behaviour reach .90 on average.

In the licensing condition by the negative indefinite *geen*, our regression model predicts a similar developmental pattern. Children younger than 3;04 (i.e. 40 months) are predicted to be able to give imitation responses around 50% of the time and that their performance slightly improves when they are older. For instance, at 5;10 (i.e. 70 months), the predicted probabilities reach .80. Moreover, the improvement in children's repetition performance is significant ($p = .00313$).

⁴ One of the 133 children was removed from our dataset because she never repeated the NPI *hoeven* in her responses, regardless of its licensing environments.

In the licensing condition by *niemand*, *weinig*, or *alleen*, however, our regression models predict a distinct acquisitional path. Instead of a starting value at around .50 attested in the licensing condition by *niet* or *geen*, children's repetition probabilities are merely around .10 at 3;04 or younger in the licensing condition by *niemand*. Nevertheless, children's repetition performance significantly increases as they grow older ($p < .001$): at 5;10 and older, they are predicted to be able to repeat the stimuli in this licensing condition approximately 90% of the time.

Children's knowledge on *hoeven* licensing by the DE-operator *weinig* exhibits a similar growth as their knowledge on the licensing of the NPI by *niemand*. When confronted with stimuli containing *hoeven* licensed by *weinig*, children below 3;10 (i.e. 46 months) give a repetition response merely 15% of the time on average, whereas their older counterparts show better imitation performance. For instance, at 4;09 (i.e. 57 months), children's repetition probabilities are around .70; and at 5;10 (i.e. 70 months), they reach .90. In this licensing condition, children's repetition performance significantly improves with age as well ($p < .001$).

With respect to the licensing condition by the exclusive adverb *alleen*, our regression model predicts the following. Between 2;09 (i.e. 33 months) and 3;06 (i.e. 42 months), children's repetition probabilities are around .10, which significantly increase ($p < .001$) and reach 1.0 at 5;03 (i.e. 63 months). The significant increase suggests a substantial growth in children's grammatical knowledge on the licensing of *hoeven* by the exclusive adverb *alleen*.

The regression results summarised above strongly suggest two kinds of developmental patterns in acquisition, distinguishable when we consider the starting values of the predicted probabilities of children's repetition performance. One pattern covers children's development in the licensing condition by *niet* or *geen*. In both these licensing conditions, our models predict a repetition probability of at least .50 at 2;09, which increases to .90 and .80, respectively, at 5;10. Another pattern signifies the development in the licensing condition by *niemand*, *weinig*, or *alleen*. Although in these licensing conditions, our models predict a repetition probability of at least .90 at 5;10, as well as for the licensing condition by *niet* or *geen*, the starting values of the imitation probabilities are merely .10 in the licensing condition by *niemand* or *weinig* and even less than .05 in the licensing condition by *alleen*.

Recall the rationale of an elicited imitation task that children are only able to repeat a stimulus if it is in line with their own grammatical system. The two developmental patterns described above thus represent the following learning path of the Dutch NPI from the ages of approximately three to six. Children start out with a strict grammar that only generates *hoeven*'s appearance in the scope of the sentential negative marker *niet* or the negative indefinite *geen* and further develop their grammar towards an adult-like direction such that the grammar at later ages also allows the NPI to be licensed by other DE-operators, namely *niemand*, *weinig*, and *alleen*. In the next section, we will explore what early and late child grammar of the NPI may consist of such that they generate the distribution of *hoeven* in language development as observed in our experiment.

We now move on with the development predicted by the regression model for the unlicensed test condition, which is presented in Figure 1 as well. At first sight, the development in this test

condition seems to exhibit a similar path as that in the licensing condition by *niemand*, *weinig*, or *alleen*. In these four test conditions, the repetition probabilities are predicted to be extremely low at 2;09 (i.e. 33 months) but increase to at least .80 at 5;04 (i.e. 64 months). Nonetheless, the development in the licensing conditions by *niemand*, *weinig*, and *alleen* are all akin to a linear pattern, whereas the development in the unlicensed test condition appears to be much less linear but rather exhibits three stages. In particular, between 2;09 and 4;00, the predicted probabilities of the repetition performance are merely .08 on average, which nevertheless increase to approximately .47 between 4;00 and 5;00, and to around .68 after 5;00. In the discussion we will come back to this point and demonstrate that the difference with respect to the linearity observed here represents different reasons underlying the increase in children's repetition scores. We will argue for an explanation based on older children's better working memory capacity.

5. Analysis

The regression results presented in Section 4 strongly suggest a two-stage development of children's knowledge on the licensing of the Dutch NPI *hoeven*. Younger children (two- and three-year-olds) are only able to repeat stimuli in the licensing conditions by *niet* and *geen*, whereas their older counterparts also show good repetition performance in the licensing conditions by *niemand*, *weinig* and *alleen*. This section explores how the knowledge on *hoeven* licensing may be presented in the grammar of children at different ages such that it generates the developmental pattern of the NPI as observed in our experiment.

5.1. *Hoeven* in early child grammar

As our regression results show, Dutch two- and three-year-olds are only able to repeat the stimuli in the licensing conditions by *niet* and *geen*, but not those in which *hoeven* is licensed by *niemand*, *weinig*, or *alleen*. Given the rationale of elicited imitation tasks (cf. Section 3), we interpret such results as evidence that children below age four have only acquired that the NPI is allowed to appear in the scope of the sentential negative marker *niet* or the negative indefinite *geen*. Moreover, as a similar development is predicted for the licensing conditions by *niet* and *geen*, namely that children are predicted to be able to repeat the stimuli in both conditions already 50% of the time on average at 2;09 and at least 80% of the time at 5;04, or older, we further hypothesise a similar kind of knowledge underlying *hoeven*'s appearance in the scope of *niet* or *geen* in early child grammar.

Following a distributional approach proposed for category learning (Cartwright and Brent 1997, Mintz et al. 1995, 2002, Mintz 2002, Redington et al. 1998), we assume that children's analysis of their target language at initial stages is input-based only. We therefore consulted the distribution information of *hoeven* in the language input in order to explore how the NPI may be represented in grammar of Dutch children at younger ages.

As reported in Lin et al. (2015), in the language input, the NPI *hoeven* co-occurs with the sentential negative marker *niet* 80.8% of the time (299 out of 370), and with the negative indefinite *geen* 10.8% of the time (40 out of 370). More interestingly, *hoeven*'s co-occurrence

with *niet* or *geen* is either adjacent, or near-adjacent, for instance, within a distance of three syllables (see a relevant discussion in Lin et al. 2015). Adopting the distribution-based learning approach (Mintz 2002, 2003, Mintz et al. 2002), we hypothesize that Dutch children establish a lexical dependency between the NPI on the one hand and *niet* or *geen* on the other when they are confronted with the massive (near-) adjacent co-occurrence of *hoeven* with these two negative forms in the language input. We further hypothesize that this lexical dependency is represented by two lexical frames [HOEF NIET] and [HOEF GEEN] in children’s mental lexicon. Assuming that these lexical frames are part of children’s lexical knowledge and are retrieved in the same way as single lexical items, it logically follows that Dutch children at younger ages are already able to give repetition responses to the stimuli in both the licensing conditions by *niet* and *geen*. This is exactly what our experimental results show (cf. Figure 1). We therefore conclude that the early child grammar of the Dutch NPI consists of two lexical frames: [HOEF NIET] and [HOEF GEEN].

A critical reader may however raise the question of why children at younger ages, for instance, below the age of four, are only able to repeat the stimuli in the relevant licensing conditions only around 50% of the time (see again Section 4). If the two lexical frames indeed form part of children’s lexical knowledge, shouldn’t we expect (much) higher repetition probabilities in these two test conditions once children have established this knowledge? We hypothesize here a possible confounding factor that may hinder (much) better imitation performance of younger children in general but is irrelevant to their knowledge of the licensing of the NPI *hoeven*.

As to investigate this confounding factor, we compared younger children’s repetition performance in the two relevant licensing conditions with their performance when confronted with filler stimuli containing a polarity insensitive modal verb, i.e. *kunnen* ‘can’ or *willen* ‘will’, with or without negation. Such filler stimuli share the same syntactic structure as our test stimuli. In particular, both types of stimuli involve two verbs: a modal verb – *kunnen*, *willen*, or *hoeven* – and a generic lexical verb such as *geven* ‘give’, *oprapen* ‘pick up’, or *zingen* ‘sing’. The average repetition rates of children under the age of four when confronted with different test and filler stimuli containing both a modal and a lexical verb are provided in Table 1. Here the criteria mentioned in Section 3 are maintained as well: *imitation responses* refer to instances in which at least the manipulated modal verb and the manipulated negation – if applicable – were repeated.

Condition	Manipulation	Repetition rate	Number of responses
Test	<i>hoeven</i> licensed by <i>niet</i>	48.1%	108
	<i>hoeven</i> licensed by <i>geen</i>	51.9%	54
Filler	<i>kunnen</i> in affirmative contexts	43.2%	83
	<i>kunnen</i> in the scope of <i>niet</i>	59.3%	54
	<i>willen</i> in affirmative contexts	48.1%	81
	<i>willen</i> in the scope of <i>niet</i>	56.8%	81

Table 1: Average repetition rates of two- and three-year-olds in the licensing condition by *niet* and *geen*, and in the filler condition containing *kunnen* or *willen*

Data reported in the table above show that Dutch two- and three-year-olds exhibit similar rates of repetition when confronted with stimuli containing two verbs – a modal verb and a lexical verb –

irrespective of whether the manipulated modal verb was the NPI *hoeven*. A further analysis of these data confirms that there is no significant difference between younger children's repetition performance in the different test and filler conditions ($F(5,439) = .712, p = .615$). Even when we only focus on the filler stimuli containing *kunnen* or *willen* in the scope of marker *niet*, which means that the only difference between these filler stimuli and the relevant test stimuli is the modal verb, we do not find any significant difference ($F(3,285) = .446, p = .720$).

These results suggest that younger children's relatively low repetition rates in the above-listed test and filler conditions do not have any indication for how *hoeven* is represented in early child grammar or how it is retrieved from children's lexicon. Children's imitation probabilities around 0.50 at younger ages are rather explained by a factor, which is irrelevant to (children's knowledge on) the licensing of the NPI. Arguably, two- and three-year-olds' poor working memory capacity may hinder better repetition performance (cf. Montgomery et al. 1978, see also Eisenbeiss 2010). In addition, it may be that stimuli containing two verbs (a modal and a lexical verb) are difficult to process and produce for children below the age of four, resulting in their poor repetition performance in the relevant licensing and filler conditions.

5.2 *Hoeven* in late child grammar

Compared to younger children, who only show relatively good imitation performance in the licensing condition by *niet* or *geen*, their older counterparts exhibit good repetition performance in all of the five manipulated licensing conditions. In particular, older children do not only repeat the stimuli containing the NPI in the scope of *niet* or *geen*, but they also give imitation responses to the stimuli in which *hoeven* is licensed by the other manipulated DE-operators: the negative indefinite *niemand*, the semi-negative quantifier *weinig*, and the exclusive adverb *alleen*. Since these are all possible licensors for the NPI in adult language use, the results obtained with older children indicate a development of an analysis of the NPI in an adult-like direction.

The most straightforward way to account for older children's analysis of the NPI is to follow the input-based learning approach, and to hypothesise that Dutch four- and five-year-olds have established three more lexical frames in addition to [HOEF NIET] and [HOEF GEEN] constructed at younger ages, namely [HOEF NIEMAND], [HOEF WEINIG], and [HOEF ALLEEN]. However, the input-based approach does not turn out to be an adequate learning mechanism at late stages. The investigation of the distribution of *hoeven* in the input by Lin et al. (2015) shows that the NPI is extremely infrequently attested with *niemand*, *weinig*, or *alleen*. *Alleen* licenses the NPI around 0.1% of the time (4 out of 370); *niemand* or *weinig* is even never attested as licensor of *hoeven*. Given the extremely infrequent co-occurrence of the NPI with these DE-operators, it appears unlikely that Dutch four- and five-year-olds may establish the corresponding lexical frames on the basis of the same distribution-based learning approach as their younger counterparts do.

An alternative explanation is to assume that older children have developed an analysis of the NPI via a learning mechanism that does not require massive co-occurrence of *hoeven* with the three DE-operators (*niemand*, *weinig*, and *alleen*) in the language input. But what may this analysis of *hoeven* consist of?

Here we adopt a decomposable analysis of negative indefinites in languages such as Dutch (cf. Jacobs 1980, see also Rullmann 1995, Zeijlstra 2011), namely that they are decomposed into both an abstract negation NEG and an existential quantifier. Such a decomposable analysis can be illustrated for *niemand* as follows.

(11) *niemand* ‘nobody’



The decomposable analysis as exemplified above applies to the negative indefinite *geen* (NEG-one) as well; it moreover applies to other negative expressions, for instance, the semi-negative quantifier *weinig* (NEG-many), and the exclusive adverb *alleen* (NEG-other than) (von Fintel and Iatridou 2003, Iatridou and Zeijlstra 2013, Penka 2011, Penka and Zeijlstra 2005). This means that the DE-operators that are not anti-morphic manipulated in our experiment all contain a decomposable, abstract negation NEG.

The incorporation of the abstract negation NEG into the DE but not anti-morphic operators employed in the current experiment provides us the possibility to assume that Dutch four- and five-year-olds establish a lexical dependency between the NPI *hoeven* on the one hand and the abstract NEG on the other. We further assume that this lexical dependency is realised as [HOEF NEG] (cf. Postal 2000). The analysis [HOEF NEG] demonstrates how the Dutch NPI is underlyingly represented in late child grammar, and generates *hoeven*'s occurrence in the scope of different DE-operators that contain the decomposable negation NEG. As this NEG can also be phonologically realised by the sentential negative marker *niet*, the abstract analysis [HOEF NEG] gives rise to *hoeven*'s appearance in anti-morphic contexts as well. Thus, the assumption of one single abstract analysis [HOEF NEG] explains why older children show good imitation performance in all of the five manipulated test conditions – even when confronted with *hoeven* licensed by extremely infrequently used DE-operators in the input.

The analysis that *hoeven* in late child grammar has a lexical dependency with the abstract negation NEG, represented as [HOEF NEG], in fact requires children's syntactic knowledge of the decomposable analysis of the DE-operators as exemplified in (11). This, however, raises two questions. First, how do we know that Dutch children have already acquired the decomposable analysis of the relevant DE-operators before reanalysing the NPI *hoeven* as [HOEF NEG]? Second, how do Dutch children develop the abstract analysis [HOEF NEG] after their construction of the two lexical frames based on input frequency only at younger ages?

The decomposable analysis of the DE-operators is evident when they are assigned a so-called split-scope interpretation if they are used together with a modal verb, for instance. Consider an example in this respect in (12), which has three readings. One reading is a narrow scope reading, which is marginally available, illustrated in (12a). Here the abstract negation NEG together with

The developmental similarities described above provide evidence for the hypothesis of the abstract analysis [HOEF NEG] in late child Dutch. *hoeven*'s appearance in the scope of *niemand*, *weinig*, or *alleen* is not generated by the two lexical frames [HOEF NIET] and [HOEF NEG] established at early stage. Therefore, younger children show extremely poor imitation performance in the licensing conditions by *niemand*, *weinig*, or *alleen*. After age four, children are assumed to have developed the abstract analysis [HOEF NEG], which generates *hoeven*'s occurrence in the scope of all the three relevant DE-operators since they can each be analysed as containing the abstract but decomposable negation NEG. This accounts for the significant increase in children's repetition probabilities in all these three licensing conditions, and explains why the development of children's knowledge of *hoeven* licensing by the infrequent licensors proceed simultaneously.

The assumption that the late child grammar contains merely one single abstract and generalisable analysis [HOEF NEG] is further confirmed when we look at the correlations among the repetition probabilities in each of these licensing contexts. The correlation data are given below.

Licenser	<i>niemand</i>	<i>Weinig</i>	<i>alleen</i>
<i>niemand</i>	1.00		
<i>weinig</i>	0.88 ($p < .000$)	1.00	
<i>alleen</i>	0.81 ($p < .000$)	0.83 ($p < .000$)	1.00

Table 2: Correlation coefficients among children's performance in the licensing condition by *niemand*, *weinig*, or *alleen*

As presented in the table, there are significantly strong correlations among children's repetition performance in these three licensing conditions. These correlation data suggest that *hoeven*'s appearance in the scope of *niemand*, *weinig*, or *alleen* has exactly the same status in child grammar. Given the learning path hypothesised in this subsection, this same status amounts to *hoeven*'s occurrence in the corresponding licensing conditions being generated by one and the same analysis. This provides evidence for the existence of the abstract analysis [HOEF NEG] in late child Dutch.

The above-reported correlation data also has an implication for Zwarts' theory of polarity licensing (cf. Zwarts 1981, 1986, 1995). As introduced in Section 2, negative contexts – which are DE-contexts in terms of Ladusaw (1979) – are categorised into three types depending on their logico-semantic behaviours: anti-morphic, anti-additive, and DE-contexts. Such a categorisation, however, does not turn out to be crucial or necessary for the acquisition of the Dutch NPI, given what we have observed in the current experiment. The negative indefinite *niemand* is an anti-additive operator, whereas *weinig* and *alleen* are both only DE. This categorial difference is nevertheless not reflected in the pace or pattern of the development of children's knowledge on *hoeven*'s appearance in the scope of these licensors. The development predicted by our regression models for the licensing conditions by *niemand*, *weinig*, and *alleen* is rather strongly correlated. In spite of the logico-semantic difference between the anti-additive operator *niemand* and the DE but not anti-additive operators *weinig* and *alleen*, children show a similar learning path in all three licensing conditions (cf. Figure 1).

On top of this, we also find that the correlation between the repetition behaviour in the licensing conditions by *geen* and *niemand* is much weaker ($r = .48, p < .000$). Since both *geen* and *niemand* are anti-additive, we would expect a much stronger correlation – if the notion of anti-additivity indeed played a crucial role in the acquisition of the NPI *hoeven*.

Taken together, the correlation results lead to the conclusion that the distinction between notions such as anti-additivity or downward entailment is irrelevant to the acquisition of the Dutch NPI. Lin et al. (2015: Appendix 2) already illustrate that the abstract negation NEG gives rise to a restricted distributional pattern of the NPI as is empirically observed with Dutch native speakers. We therefore conclude that semantic notions such as anti-additivity or downward entailment are irrelevant to the licensing of this particular NPI either. Given the abstract analysis [HOEF NEG], emerged as a result of language acquisition, we argue that *hoeven* is only allowed to appear in DE-contexts that incorporate this abstract negation, because of its lexical dependency with the abstract negation NEG (cf. Postal 2000). This in turn may explain the distributional difference between the Dutch NPI and English *any*-terms as introduced at the beginning of the paper.

6. Discussion & Conclusion

Above we hypothesised a two-staged development of how Dutch children acquire the NPI modal verb *hoeven*, in which they start out with two lexical frames [HOEF NIET] and [HOEF GEEN] at initial stages, and develop one single abstract analysis [HOEF NEG] later on. Before drawing any conclusion, we would like to first discuss children's repetition performance attested in the test condition in which *hoeven* appears in simple affirmative contexts. We will argue for an explanation based on children's working memory capacity.

Recall the repetition performance in the unlicensed test condition predicted by our regression model: between 2;09 and 4;00, the repetition probabilities are merely .08, which increase to approximately 0.47 between 4;00 and 5;00, and further increases to around 0.68 after 5;00. At first sight, the improvement in children's repetition performance when confronted with *hoeven* in simple affirmative contexts seems to suggest a development towards a non-adult-like direction. In particular, it seems that children are developing a tolerant grammar, which even allows *hoeven* to appear in the absence of a licenser – although they start out with a much narrower analysis of *hoeven* that restricts it to co-occur with *niet* or *geen* only.

We argue here that the increase in children's repetition probabilities in the unlicensed test condition does not represent a change in children's knowledge on *hoeven* licensing towards a non-target-like direction but is rather explained as a consequence of older children's better working memory capacity. As mentioned in Section 3, the length of stimuli is crucial to children's behaviours in an elicited imitation task. To ensure that children (re)construct their own mental representation of stimuli based on their own grammar but do not give a repetition of stimuli from memory alone, stimuli must be sufficiently long to override children's memory capacity. Nevertheless, to be able to compare the performances of our participants of different ages, we opted for a unified stimuli length of ten words – a medium length of stimuli according

to Montgomery et al. (1978). If we assume that the working memory capacity of our participants – who were all typically developing – increases with age, it is not impossible that the length of ten words was just too short for the four- and five--year-olds to override their better working memory capacity compared to their younger counterparts. This may result in the unexpected improvement in their repetition performance in the unlicensed test condition.

The above hypothesised explanation may account for the difference with respect to the linearity of the development observed for the licensing conditions by *niemand*, *weinig*, and *alleen* on the one hand, and that attested for the unlicensed test condition on the other, although the developments predicted by our regression models for all four test conditions seem to have a similar starting value at the age of 33 months. The unlicensed test condition demonstrates a more stage-like development, whereas the development in the licensing conditions by *niemand*, *weinig*, and *alleen* are all akin to a linear growth pattern (cf. Figure 2). The difference with respect to the linearity may represent different reasons underlying the improvement attested in children's repetition performance. However, our experiment did not contain any procedure for examining the participants' working memory. This calls for further research in this respect.

To conclude, the experimental results obtained with 132 monolingual Dutch children in an elicited imitation task suggest an acquisitional path as follows. Children start with a strict grammar that only allows *hoeven* to appear in the scope of either the sentential negative marker *niet* or the negative indefinite *geen*, represented by two lexical frames [HOEF NIET] and [HOEF GEEN] in early child grammar, but later switch to a less strict grammar that allows *hoeven* to appear in a wider set of DE-contexts, namely those introduced by *niemand*, *weinig*, or *alleen*, represented by [HOEF NEG] in late child language. Since [HOEF NEG] is the analysis of the Dutch NPI emerged as the result of language acquisition, we conclude that *hoeven* is an NPI because it has a lexical dependency with the abstract negation NEG (cf. Postal 2000). Our experimental results also lead to the conclusion that semantic notions such as anti-additivity or downward entailment is irrelevant to the acquisition of NPIs such as *hoeven*, which exhibit a narrower distribution than NPIs like *any*-terms. Moreover, our exploration of the acquisition of the Dutch NPI strongly suggest that the acquisition of NPIs like *any* cannot and must not show the same learning pathway as that detected for the Dutch NPI *hoeven*. The reason is twofold: on the one hand, input evidence differs from language to language, and from NPI to NPI; on the other hand, an analysis such as [ANY NEG] does not generalise the target distribution of *any*.

References

- Brown, R. (1973). *A First Language: The Early Stages*. Cambridge, MA: Harvard University Press.
- Cartwright, T. A. and M. R. Brent (1997). Syntactic categorization in early language acquisition: Formalizing the role of distributional analysis. *Cognition* 63(2). 121–170.
- Chomsky, N. (1964). Formal discussion. In Bellugi, U. and Brown, R. (Eds.). *The Acquisition of Language*. Monographs of the Society for Research in Child Development. 29 (1, Serial No. 92) (pp. 37–39).
- Eisenbeiss, S. (2010). Production Methods. In Blom, E. and Unsworth, S. (Eds.), *Experimental Methods in Language Acquisition Research* (pp. 11–34). Amsterdam: John Benjamins.

- Fauconnier, G. (1975). Pragmatic scales and logical structure. *Linguistic Inquiry* 6(3), 353–375.
- Fauconnier, G. (1978). Implication reversal in a natural language. *Formal semantics and pragmatics for natural languages* (pp. 289–301). The Netherlands: Springer.
- von Fintel, K. (1999). NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics* 16(2), 97–148.
- von Fintel, K. and Iatridou, S. (2003). Epistemic containment. *Linguistic Inquiry* 34(2), 173–198.
- Hoeksema, J. (1994). On the grammaticalization of negative polarity items. *Annual Meeting of the Berkeley Linguistics Society, Berkeley*. 20, 273–282.
- Hoeksema, J. (2000). Negative polarity items: Triggering, scope and c-command. In Horn, L. and Kato, Y. (Eds.), *Negation and polarity: Syntactic and semantic perspectives* (pp. 115–146). Oxford: Oxford University Press.
- Iatridou, S. and Zeijlstra, H. (2013). Negation, Polarity, and Deontic Modals. *Linguistic Inquiry* 44(4), 529–568.
- Jacobs, J. (1980). Lexical decomposition in Montague Grammar. *Theoretical Linguistics* 7, 121–136.
- Keenan, E. L. and Hawkins, S. (1987). The psychological validity of the Accessibility Hierarchy. In Keenan, E. L. (Ed.), *Universal Grammar: 15 Essays* (pp. 60–85). London: Croom Helm.
- Keeney, T. and Wolfe, J. (1972). The acquisition of agreement in English. *Journal of Verbal Learning and Verbal Behavior* 11, 698–705.
- Ladusaw, W. (1979). Negative polarity items as inherent scope relations. Unpublished doctoral dissertation, University of Texas at Austin.
- Lin, J., Weerman, F. and Zeijlstra, H. (2015). Emerging NPIs: the acquisition of Dutch *hoeven* ‘need’. *The Linguistic Review* 32(2). 333–374.
- Lust, B., Flynn, S. and Foley, C. (1996). What children know about what they say: Elicited imitation as a research method for assessing children's syntax. In McDaniel, D., McKee, C. and Cairns, H. S. (Eds.), *Methods for assessing children's syntax* (pp. 55–76). Cambridge, MA: MIT Press.
- MacWhinney, B. (2009). The CHILDES project. Tools for Analyzing Talk – electronic edition 2.
- Mintz, T. H. (2002). Category induction from distributional cues in an artificial language. *Memory & Cognition* 30(5). 678–686.
- Mintz, T. H., Newport, E. L. and Bever, T. G. (1995). Distributional regularities of grammatical categories in speech to infants. *The North East Linguistics Society* 25. 43–54.
- Mintz, T. H., Newport, E. L. and Bever, T. G. (2002). The distributional structure of grammatical categories in speech to young children. *Cognitive Science* 26(4). 393–424.
- Montgomery, M. M., Montgomery, A. A. and Stephens, M. I. (1978). Sentence repetition in preschoolers: Effects of length, complexity and word familiarity. *Journal of Psycholinguistics Research* 7, 435–452.
- Panitsa, G. (2001). Aspects of aspect: Acquiring grammatical aspect in Modern Greek: The interaction of aspect with aspectual adverbials in children's grammars. UCL Working Papers in Linguistics 13, 405–424.
- Penka, D. (2011). Negative indefinites. Oxford: Oxford University Press.
- Penka, D. and Zeijlstra, H. (2005). Negative indefinites in Dutch and German. Paper presented at the 20th Comparative Germanic Syntax Workshop, Tilburg. <http://ling.auf.net/lingbuzz/000192>

- Postal, P. M. (2000). The Ohio lectures on squat. Manuscript, *New York University*.
- Redington, M., Chater, N. and Finch, S. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science* 22(4), 425–469.
- Rullmann, H. (1995). Geen eenheid [No Unit]. *TABU* 25(2), 194–197.
- Scholl, D. M. and Ryan, E. B. (1980). Development of metalinguistic performance in the early school years. *Language and Speech* 23, 199–211.
- Van der Wal, S. (1996). Negative polarity items and negation: Tandem acquisition. Doctorial dissertation, University of Groningen.
- Van der Wouden, T. (1997). Negative Contexts: Collocation, Polarity and Multiple Negation. Unpublished doctorial dissertation, University of Groningen.
- Vinther, T. (2002). Elicited imitation: a brief overview. *International Journal of Applied Linguistics* 12(1), 54–73.
- Zeijlstra, H. (2011). On the syntactically complex status of negative indefinites. *Journal of Comparative Germanic Linguistics* 14(2), 111–138.
- Zwarts, F. (1981). Negatief polaire uitdrukkingen I [Negative polarity expressions I]. *Glott* 4(1), 35–132.
- Zwarts, F. (1986). *Categoriale Grammatica En Algebraïsche Semantiek: Een Onderzoek Naar Negatie En Polariteit in Het Nederlands* [Categorical grammar and algebraic semantics: a study of negation and polarity in Dutch]. Doctorial dissertation, University of Groningen.
- Zwarts, F. (1995). Nonveridical contexts. *Linguistic Analysis* 25, 286–312.