J Psycholinguist Res (2018) 47:1193–1217
https://doi.org/10.1007/s10936-018-9573-8

Strong and Weak Readings in the Domain of Worlds: A Negative Polar Modal and Children's Scope Assignment

Loes Koring[1,2] · **Luisa Meroni**[4] · **Vincenzo Moscati**[3]

Published online: 22 March 2018

Abstract This study investigates children's interpretation of sentences with two logical operators: Dutch universal modal *hoeven* and negation (*niet*). In adult Dutch, *hoeven* is an NPI that necessarily scopes under negation, giving rise to a NOT > NECESSARY reading. The findings from a hidden-object task with 5- and 6-year-old children showed that children's performance is suggestive of an interpretation of sentences with *hoeft niet* in which the modal scopes over negation (NECESSARY > NOT). This is in line with the Semantic Subset Principle that dictates that children should opt for the strongest possible reading in case of potential scope ambiguities. The full pattern of results, however, seems to be determined, in addition, by a particular strategy children use when facing uncertainty called Premature Closure.

Keywords Semantic Subset Principle · Scope assignment · Modality · Negation · Acquisition

Introduction

A fascinating characteristic of language is that it allows us to talk about things outside of the actual here and now (Hockett 1960). One tool that makes this possible is the use of modal expressions like *can*, *may*, or *must*. That is, modal expressions can be used to entertain

All the authors equally contributed to this article. Their names are listed in alphabetical order.

✉ Loes Koring
loes.koring@mq.edu.au

1 ARC Centre of Excellence in Cognition and its Disorders, Department of Cognitive Science, Macquarie University, Australian Hearing Hub, 16 University Avenue, Sydney, NSW 2109, Australia

2 Department of Linguistics and Philosophy, Massachusetts Institute of Technology, Cambridge, MA, USA

3 Department of Social, Political and Cognitive Sciences, University of Siena, Siena, Italy

4 Department of Language, Culture and Communication, Utrecht University, Utrecht, The Netherlands

alternative situations that do not exist right in front of the speaker's eyes. Modality is "the linguistic phenomenon whereby grammar allows one to say things about, or on the basis of, situations which need not be real" (Portner 2007, 1). Expressions of modality constitute a big challenge for children as they have to figure out the meaning of such expressions without the help of direct cues in the actual scene. Children have to understand, for example, that a sentence like *"it might rain tomorrow"* remains true even if such meteorological event will never take place. This problem has drawn considerable attention in language acquisition research and several sub-properties of modal expressions have been investigated in order to determine how they develop in early childhood. In particular, much research has focused on the acquisition of the modal's strength in relation to its base (e.g., Hirst and Weil 1982; Byrnes and Duff 1989; Moore et al. 1990; Noveck et al. 1996).

In this paper, we will instead look more closely at another factor involved in modal acquisition, namely *modal scope*, which has as of yet received much less attention. To illustrate the challenge modal scope poses for the language-acquiring child, consider the English sentences in (1) and (2), in which a modal interacts with a negative operator:

(1) John must not wear a tie
(2) John need not wear a tie

Sentences (1) and (2) constitute a minimal pair. While both sentences contain a (deontically used) modal with universal quantificational force and negation, they convey two different meanings as a result of the interaction of the modal and negation. In fact, while *must* takes wide scope over negation, resulting in a NECESSARY > NOT interpretation in (1), *need* in (2) instead takes narrow scope with respect to negation, giving rise to a NOT > NECESSARY interpretation. More informally, while sentence (1) expresses the *necessity to not* wear a tie, sentence (2) expresses that it is *not necessary* to wear a tie. An important question is whether these two alternative scope assignments are equally accessible to the child at a given stage in development, or, alternatively, if an asymmetry exists such that children allow only one particular scope assignment. Experimental work on the acquisition of scope assignment in the domain of Italian modals suggests that children start out hypothesizing that the combination of a modal and another logical operator permits one reading only (Moscati and Crain 2014).

The difficulty in acquisition related to the interaction of modals with other logical operators is not independent from (the acquisition of) more general properties of modals. An interesting property of modals is that the same modal (e.g. *must* in (1)) allows for both epistemic and deontic readings, as the modal statement needs to be evaluated against a relevant set of accessible possible worlds (Traugott 1988; Kratzer 1981, 1991; Bybee et al. 1994; Palmer 2001). This set is typically referred to as the *modal base*. The modal base relevant for interpreting a modal statement is either inferred contextually or mentioned explicitly and can also be contingent on tense and aspect marking (e.g., Hacquard 2010; Ramchand 2012). To illustrate, consider the sentences in (3) and (4) where the modal *must* appears in combination with the verb *swim*.

(3) Mary must swim to get her swimming diploma.
(4) It's 3 pm, so Mary must be swimming.

In (3), we get a reading in which Mary is obliged to swim, the modal base is deontic, given what is necessary to get a swimming diploma. In (4), however, *must* receives an epistemic interpretation as, given what I know about Mary's day schedule, Mary must be swimming.

Apart from the contents of the modal base, there is another important parameter we need to take into consideration when it comes to interpreting modals, namely the modal's *quantificational force* which can be either universal or existential. This difference is exemplified

in the examples below that show that while (5) is true if there is some world in which Mary is swimming (existential), (6) is true if and only if Mary is swimming in *all* possible worlds (universal).[1]

(5) Mary might be swimming
(6) Mary must be swimming

Existing studies on children's acquisition of modal terms return a complex picture in which high and low performance alternate. While some modal expressions are present in children's spontaneous speech from very early on, children do not yet seem to have completely mastered the subtle semantic differences between modal expressions at that age. In English, for example, some modals are attested in production before age 2, yet others seem to appear much later in children's speech (Kuczaj and Maratsos 1975; Shepherd 1982; Bliss 1988; Stromswold 1990; Stephany 1979; Wells 1979; Perkins 1983). To qualify this delay in the production of certain modal auxiliaries, the distinction between epistemic and deontic modals has often been invoked in the aforementioned studies (on the "Epistemic Gap", see also Cournane 2015). The observation is that in the same time-window in which children are able to produce deontic modals, they still avoid using epistemic modals (O'Neill and Atance 2000; Perkins 1983; Wells 1979). Comprehension studies on the acquisition of modality complicate the picture even more. Relevant for our purposes is that 5-year-olds have been shown to experience problems in evaluating epistemic possibilities (Byrnes and Overton 1986; Noveck et al. 1996; Noveck 2001; Bascelli and Barbieri 2002, for Italian among others). As will be elaborated on "in the "Premature Closure" section", this complex picture seems to be (at least partly) the result of another factor involved in testing epistemic possibility: dealing with multiple alternatives (Ozturk and Papafragou 2015). The difficulty is that part of the task for the child is to entertain more than one alternative at once; a skill they do not seem to have fully developed yet (e.g., Reinhart 2004).

The focus of the present paper is children's comprehension of the interaction between (epistemic) modal auxiliaries and another scope-bearing element, namely negation. As we will argue, children's particular way of dealing with multiple alternatives is relevant in this case too.

Interacting Logical Operators

Turning to the interaction of logical operators and negation, several studies reported that children do not behave adult-like in scope assignment (Musolino 1998; Musolino et al. 2000; Musolino and Lidz 2003; Zhou and Crain 2009). Importantly, children's difficulties with modals are exacerbated in negative sentences. That negation increases processing load has been repeatedly shown since the work of Wason (1965) and Slobin (1966) (e.g., Carpenter et al. 1999; Fischler et al. 1983; Herbert and Kübler 2011). The additional complexity of sentences with modals and negation, however, does not just seem to be the result of the cost of processing the negation. An additional challenge for the language-acquiring child in sentences with a modal *and* negation, is that the child needs to work out the scopal relations between negation and the modal (Moscati and Crain 2014; Moscati et al. 2017).

To illustrate this, let us consider again the sentences with a modal auxiliary and a negative operator in (1) and (2).

[1] Another relevant parameter when interpreting modals is the ordering source. The discussion of this parameter, however, is beyond the scope of the present paper (see Kratzer 1981).

(1)	John must not wear a tie		
	a.	*It is not necessary for John to wear a tie	NOT > NECESSARY
	b.	It is necessary that John does not wear a tie	NECESSARY > NOT
(2)	John need not wear a tie		
	a.	It is not necessary for John to wear a tie	NOT > NECESSARY
	b.	*It is necessary that John does not wear a tie	NECESSARY > NOT

The examples in (1) and (2) illustrate the two logically possible scope assignments through the lexical alternation between *need/must*. While *must* in (1) only allows for a reading in which the modal takes wide scope (NECESSARY > NOT), the reverse holds in (2).

Even though there are two logically available readings for both (1) and (2), sentences (1) and (2) are not ambiguous as shown by the ungrammaticality of (1a) and (2b). The child's task is to work out which reading is available for each of these sentences: does the modal take scope over negation, or does negation take scope over the modal? Importantly, the surface order of the modal and negation is not informative with regard to scope assignment.

One additional observation with respect to the two readings is that (1b) has more stringent truth-conditions than (1a). More precisely, the reading in (1b) asymmetrically entails the reading in (1a) (i.e. in all circumstances under which (1b) is true, (1a) is also true, but not vice versa). Following the literature we will then refer to the reading in (1a) as the weak reading, and the one in (1b) as the strong reading. This notion of *strength* has been shown to be relevant in language acquisition through work showing that weak readings are problematic for children younger than 6 years (Gualmini and Moscati 2009; Moscati and Gualmini 2007; Moscati and Crain 2014). The study reported in Moscati and Crain (2014) specifically addressed this issue by looking at children's interpretation of sentences like (7) in Italian:

(7)	Ci può non essere una mucca nella scatola			
	a.	'There might not be a cow in the box'	POSSIBLE > NOT	(WEAK READING)
	b.	*'There cannot be a cow in the box'	*NOT > POSSIBLE	(STRONG READING)

In sentence (7) the modal *potere* precedes the negative marker *non* and although the presence of two logical operators introduces a potential scope ambiguity (compare (7a) and (7b)), the interpretation is constrained: only reading (7a) is available for adult speakers of Italian. Interestingly, Moscati and Crain showed that 5-year-old Italian-speaking children consistently misinterpret (7) according to the deviant reading in (7b). That is, children assigned a strong reading to sentence (7), even though this reading is not available in adult Italian.

The experiment was set up as a reasoning scenario similar to Noveck (2001), which in turn was an adaptation of Hirst & Weil's original Hidden Object Task. As illustrated in Fig. 1 below, the child is presented with three boxes. Two of these boxes are open (Box 1 and Box 2), while a third box, the mystery box, is closed. The two open boxes contain some object(s) which are visible to the child, as the boxes are open. Since the mystery box is closed, however, the child cannot see what is inside this box. The child's task is to infer what could be in the mystery box via a reasoning rule: its content is either the same as the content of Box 1 or the content of Box 2 (e.g. Fig. 1: a horse and a cow in the top-left box, or a horse alone in the top-right box).

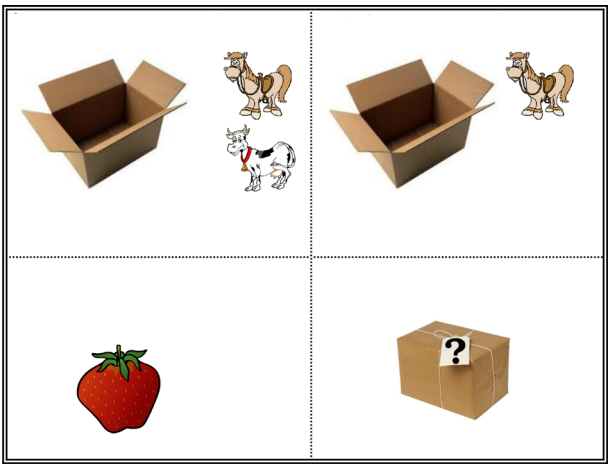

Fig. 1 A hidden object scenario as used in Moscati and Crain (2014)

Given this scenario, children judged the truth-value of sentences like (7). As we said, scope assignment in (7) follows the surface word order such that it only allows for a POSSIBLE > NOT reading. As a POSSIBLE > NOT reading is asymmetrically entailed by a NOT > POSSIBLE reading, the POSSIBLE > NOT reading is the weak reading. For adults, sentence (7) would therefore be true in the scenario in Fig. 1, as it is possible that there will not be a cow in the mystery box. Indeed, adults accepted this statement in 95% of the cases with rejections being only a tiny fraction of their responses. Children, however, incorrectly rejected sentence (7) in 65% of the cases; they seemed to ignore the surface order of the modal and negation and assigned a strong (NOT > POSSIBLE) reading to the sentence, yielding a 'false' judgment, consistent with interpretation (7b).

Moscati and Crain (2014) ascribe children's preference for the strong reading to a particular learning mechanism the language-acquiring child employs, namely the Subset Principle (Berwick 1982). More specifically, they adopt a semantic version of the Subset Principle, known as the Semantic Subset Principle (SSP) (Crain 2012; Crain et al. 1994). In short, the SSP dictates that, in case of a potential ambiguity, the child should always opt for an interpretation that makes the sentence true in the *narrowest* set of circumstances. The reason is that, if the child adopts this strategy, she can recover from possible incorrect hypotheses without having to rely on negative evidence. The SSP guarantees that learning can be based on positive evidence alone, in line with the observation that negative evidence is not useful (if existing at all) in language acquisition (Brown and Hanlon 1970; Marcus 1993).

This implies for the domain of modals and negation that the child should initially always hypothesize that a potentially ambiguous sentence allows for the strongest reading only. If it then turns out that the strong reading does not match the range of available interpretations in the target language, this erroneous hypothesis can be revised on the basis of positive evidence, i.e., hearing the sentence in a situation in which it is false under the (hypothesized) strong reading. In contrast, if a child erroneously hypothesizes a weak reading to be true, then she will never encounter any evidence suggesting that her hypothesis is incorrect, as she will never hear this sentence in a situation in which it is false (as the weak reading is entailed by the strong (target) reading). In short, in case of ambiguity, the safest option for the child is to opt for the strongest reading, as there will be ample evidence in the input to guide her towards the correct interpretation in case her initial hypothesis turns out to be incorrect.

Let us now illustrate how the SSP accounts for children's performance of sentences with negation and the possibility modal in (7). The SSP dictates that the child initially hypothesizes only the strong reading to be available when two logical operators are encountered. For the sentences in (7) this means that they will opt for the reading in which negation takes scope over the possibility modal (NOT > POSSIBLE), even if this reading is not allowed in the target grammar.

Besides providing support for the SSP as a mechanism for language development, Moscati and Crain's account also makes an interesting prediction. The claim is that children misinterpret sentences as in (7) as a result of a strategy dictated by learnability considerations. If children's misinterpretations indeed follow from properties of the language acquisition device (LAD), we expect the same pattern to occur across different modal operators and across languages. That is, we expect children to opt for the strong reading also when negation is combined with a modal of different quantificational force, or, potentially, a different modal base, even if this reading is not consistent with the target grammar. In this paper, we put (part of) this hypothesis to the test by investigating the interaction of a *necessity* modal with negation in child language. That is, we single out the parameter *quantificational force* and investigate the interaction of negation and a modal of a different quantificational force, while holding the modal base constant (epistemic). We will discuss the differences and similarities between the Italian existential modal and the Dutch universal modal in more detail "in the "Polarity Sensitive Modals: The Case of Dutch 'hoeven' (Must)" section". Universal and existential modals are expected to behave alike as far as the SSP is concerned. An important difference is, however, that the scope assignment that generates the strong reading is inverted when a universal modal and negation are considered as opposed to an existential modal and negation. Table 1 illustrates this point and the generalized prediction of the SSP.

Table 1 shows that the strongest reading for negation and an existential modal is a reading in which the existential modal scopes under negation (NOT > POSSIBLE). The strongest reading for negation and a universal modal instead arises when the modal takes wide scope over negation (NECESSARY > NOT). The SSP makes the prediction that, if any asymmetry exists between the two competing scope assignments, the child will always prefer the scope assignment that generates the strongest reading.

With this in mind, the goal of this paper is to ascertain whether the pattern found in Moscati and Crain (2014) holds cross-linguistically and across modals that differ in quantificational force, in accordance with the SSP's predictions. Specifically, we will look at children's interpretation of universal Dutch modal *hoeven* and negation. This cross-linguistic comparison would allow us to control for language specific factors active in Italian. In particular, one reason to think that the observed pattern might not hold cross-linguistically comes from a property specific to Italian that could be the culprit for children's poor performance. In Italian, it is often the case that the modal and negation are adjacent, as in the sentences in

Table 1 Expected scope assignment predicted by the Semantic Subset Principle

Reading	Modal force		Initial hypothesis on scope assignment according to the Semantic Subset Principle
	Universal	Existential	
Weak reading	NOT > NECESSARY	POSSIBLE > NOT	x
Strong reading	NECESSARY > NOT	NOT > POSSIBLE	✓

(7).[2] Since many languages employ negative auxiliaries in which a syncretic form could express negation plus other temporal, modal, and aspectual features (see Miestamo 2003), it is possible that children's difficulties in assigning the correct scope to *potere* (might) and *non* (not) are due to an incorrect parse of the two elements (negation and modal). Moreover, the scope restrictions in force on (7) are likely due to more general syntactic restrictions found in restructuring contexts in Italian (Rizzi 1982; Cardinaletti and Shlonsky 2004; Moscati et al. 2017).

Two features of Dutch make this language particularly interesting when compared to Italian:

(i) Dutch has a richer modal paradigm than Italian. In particular, the modal paradigm includes a necessity modal that encodes polarity restrictions.
(ii) Dutch word order, as opposed to Italian word order, allows negation and the modal operator to be separated.

In particular, the Dutch modal under investigation lexically encodes scope-assignment restrictions. Scope assignment for this particular Dutch modal is therefore not related to special properties of restructuring sentences, like in Italian. Furthermore, Dutch allows us to eliminate the possibility that adjacency between negation and the modal poses difficulties for the language-acquiring child, since *niet* is not a bound morpheme and, as such, can be separated from the modal. The next section will describe the properties and children's use of modal *hoeven* in comparsion to Italian *potere* in more depth.

Polarity Sensitive Modals: The Case of Dutch 'Hoeven' *(Must)*

Dutch *hoeven* is a necessity modal (i.e. universal quantificational force) that can quantify over different sets of possible worlds. It can take both a deontic (8a) and an epistemic (8b) modal base.[3] In particular, while (8a) means that according to the laws, there is no need for Ingrid to go to school today, (8b) means that, given what I know, it is not necessarily the case that Mary is at school today (she might be somewhere else).

(8a) Ingrid hoeft niet naar school te gaan vandaag.
Ingrid must not to school to go today.
'Ingrid don't have to go to school today.'
(8b) Marie hoeft niet op school te zijn vandaag.
Mary must not at school to be today.
'Mary doesn't have to be at school today.'

An interesting property of *hoeven* is that it is a polarity sensitive modal such that it necessarily scopes under negation; i.e. it behaves like an NPI (Hoeksema 1997; Iatridou and Zeijlstra 2013; van der Wouden 1994; van derWouden 1997; Zwarts 1981, 1998). Sentences with *hoeven* accompanied by negation therefore always give rise to a weak NOT > NECESSARY reading, independent of the surface order of negation and the modal in the sentence. *Hoeven* requires licensing by a subclass of downward entailing operators. It is licenced in sentences

[2] Although some material could linearly occur between the two, this will result in a different meaning, with negation having narrow scope and resulting in a constituent-type of negation.
i. Il cavallo *non* sempre *può* essere nella scatola
the horse not always might be in the box
"the horse might not always be in the box" .

[3] Epistemic modals with narrow scope under negation are also recognized in von Fintel and Iatridou (2003).

(9)-(11), but it cannot be licensed by relative clauses, questions or conditionals (12) (van der Wouden 1994; Iatridou and Zeijlstra 2013).

(9a) Jan hoeft *(niet) te werken.
John need NEG to work
'John doesn't need to work.'
(9b) Ik hoorde dat Jan *(niet) hoeft te werken.
I heard that John NEG need to work.
'I heard that John doesn't need to work.'
(10) Geen/*Iedere/*Een student hoeft weg te gaan
No/*Every/*A student need away to go
(11a) Niet iedereen hoeft te werken.
NEG everybody need to work
'Not everybody needs to work.'
(11b) Alleen Jan hoeft te werken.
only Jan need to work
'Only Jan needs to work.'
(12a) *Iedereen die hoeft te werken wordt om 7:00 verwacht.
everybody who need to work is at 7:00 expected
'Everybody who needs to work is expected at 7:00.'
(12b) *Als je hoeft te werken, word je om 7:00 verwacht.
if you need to work are you at 7:00 expected
'If you need to work, you are expected at 7:00.'
(12c) *Hoeft Jan morgen te werken?
Need John tomorrow to work?
Does John need to work tomorrow?
(Examples (9) – (12b) above are taken from Iatridou and Zeijlstra 2013: 560)

The sentences in (9) show that negation can both precede and follow the NPI *hoeven* and this does not yield differences in interpretation.That is, there is no restriction such that negation must precede *hoeven* in the surface string. Yet, *hoeven* does require to be semantically interpreted in the scope of the downward entailing operator that licenses it. Thus *hoeven* necessarily reconstructs at LF to a position lower than negation (Iatridou and Zeijlstra 2013). As such, for sentences that include negation and *hoeven*, only a weak reading (NOT > NECESSARY) is available, regardless of the surface order of negation and the modal.

It is not only the case that negation can both precede and follow the modal in the surface order; the modal and its licenser can also be separated from each other by other clausal material as illustrated in (13). This intervening material could provide evidence to the child that negation and the modal are really two separate elements and as such preclude a misanalysis of the modal and negation as one syncretic head.

(13) Jan hoeft morgen niet te werken.
John need tomorrow NEG to work.
'John doesn't need to work tomorrow.'

For what concerns Dutch children, Lin et al. (2015) analyzed the spontaneous production of 53 children in corpora collected through CHILDES (MacWhinney 2000) to find out whether children produce *hoeven* in the context of a licenser, or whether they produce (ungrammatical) unlicensed *hoeven*. Out of the 117 instances of *hoeven* they found for children between 1;0 and 3;0, children's use of *hoeven* was adult-like in 95% of the cases, as in sentence (14) below:

Table 2 Comparison of the properties of Dutch modal *hoeven* and Italian modal *potere*

Modal	Force	Modal base	Syntactic category	Requires licensing?
hoeven	Universal	Epistemic, Deontic	Auxiliary Verb[a]	Yes
potere	Existential	Epistemic, Deontic	Auxiliary Verb	No

[a]Modal *hoeven* differs from other Dutch modals in that it takes a complement introduced by *te* 'to'. Other modals (e.g., *moeten* 'must', *kunnen* 'can') take a smaller VP complement without *te* 'to'. The exact status of *te 'to'* is unclear but it is not necessarily an infinitival marker (see Zwart 1993, Reuland 2003)

(14) *Nee, ik hoef niet naar de wc.*
no I need not to the toilet
'No, I do not have to go to the bathroom.'
(Wijnen and Bol 1993: abe30308.cha: line 928)

Non-adult-like uses [e.g. (15)] made up less than 5% of the data.

(15) *Ik hoef drinken.*
I need drink
'I want to drink something.'
(Wijnen and Bol 1993: mat30113.cha: line 515) (Examples are from Lin et al. (2015): 347, 345).

Even though these data reveal that children are aware of the co-occurrence of *hoeven* and negation from early on, they do not inform us about children's interpretation of *hoeven* and its licenser.[4] That is, knowing that *hoeven* co-occurs with an element that contains negation does not imply that the child knows that *hoeven* is necessarily interpreted in the scope of this element.[5] This point is crucial to us. Given the syntactic properties of sentences with *hoeven* in Dutch, this type of construction is the ideal testing ground to test the validity of the SSP in negative modal sentences and to assess whether the data discussed in Moscati and Crain (2014) are contingent upon some language-specific properties of Italian or if the pattern could be extended to other languages as well.

We already pointed out that Italian children systematically misinterpret "*puo non*" (might not) sentences according to a deviant scope assignment that yields the strong reading. However, since in Italian negation and modality are often adjacent and the rigid scope assignment between *potere* (might) and *non* (not) is arguably the result of a very specific syntactic constraint, we cannot exclude the possibility that children's preference for the strong reading is an epiphenomenon that stems from the particular grammar of Italian instead of a general learning mechanism like the SSP.

To facilitate the comparison between the Italian modal auxiliary *potere* and Dutch *hoeven*, their properties are summarized in Table 2.

Table 2 shows that the two modals differ in their modal force, but not in potential modal bases. An important additional difference is whether or not licensing is required: Dutch modal *hoeven* is an NPI, whereas Italian *potere* is not. That is, while *potere* can freely occur without a licensor and is able to take wide scope over the negative operator in particular syntactic

[4] Note that these children are younger than the age group we will test in our experiment.

[5] Note that for Lin et al. (2015) *hoeven* acquires its NPI-hood on the basis of the distribution in the input. For these authors, therefore, knowing the distributional pattern of *hoeven* would imply knowing its scope assignment. However, as pointed out in the text, co-occurrence of *hoeven* and negation is not informative about scope assignment per se.

configurations, *hoeven* is obligatorily interpreted under the negative operator, independent of its surface position with respect to negation. Scope assignment is thus lexically specified for Dutch *hoeven* and negation, but structurally determined in Italian. If children indeed prefer strong readings, like Moscati and Crain (2014) found for Italian, we would predict that children would interpret Dutch *hoeven* as taking scope over negation thus accessing an (ungrammatical) strong NECESSARY > NOT reading. Note, however, that another potential influencing factor might be the way in which polarity restrictions are encoded (lexically vs. structurally). We will come back to this point "in the "Discussion" section". In order to test whether children have a preference for strong readings, we designed an experiment that tests children's interpretation of sentences with *hoeft niet*.

Testing Weak Readings in Dutch

In order to further investigate the role of SSP in children's interpretation of potentially ambiguous sentences, we extended this line of research to epistemic uses of universal modal *hoeven* in Dutch. In the case at hand, the Semantic Subset Principle would dictate that children initially hypothesize that the necessity modal scopes over negation, independently of word order and polarity contraints. Children are therefore predicted to differ from adults in their interpretation of *hoeft niet*, as *hoeven*, for adults, is a negative polar modal that can only be interpreted in the scope of negation.

Participants

Thirty-seven monolingual Dutch-speaking children participated in the experiment. The children were recruited from two schools in Brabant, The Netherlands. Children ranged in age from 4;10 to 6;11. Children were divided in two different age groups: 19 5-year-olds (4;10 – 5;11, mean: 5;5) and 18 6-year-olds (6;0 – 6;11, mean age: 6;5). Six adults were tested as a control group.

Materials and Procedure

Our experiment employed the same hidden-object task as used in Moscati and Crain (2014). In the experiment, children were presented with three boxes on a table in front of them. One of the boxes was closed and its content was therefore unknown. The contents of this box could be inferred on the basis of the contents of the two open boxes. The Mystery Box 3 either contained exactly the same as Open Box 1 or exactly the same as Open Box 2 (see (16) and Fig. 2 for an example). As such, we set up an epistemic conversational background, as the task is trying to infer what the contents of a closed box could be given what we know about the contents of the two open boxes.

(16) Box Set I (see Figure 2)
Open Box 1: pig + elephant
Open Box 2: elephant
Mystery box: ?
Outside: kangaroo

The game was introduced to the child as a game in which we try to work out what could be in the mystery box. The box is closed, so we cannot look inside to check what is in there, but we have got some clues. Namely, what is in that mystery box is exactly the same as what is

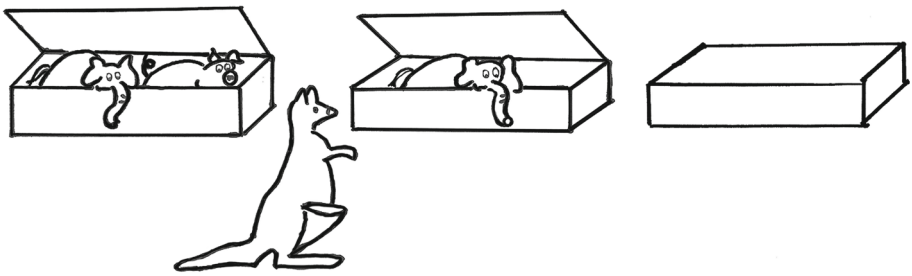

Fig. 2 This is a graphic representation of the experimental set-up

in Open Box 1. Or, the mystery box contains exactly the same as what is in Open Box 2. We also introduced a hand puppet, Miss Snail, who is particularly fond of this game. For each box set, the child, along with the hand puppet, would first inspect the contents of the open boxes. The boxes contained different types of objects (animals, fruit, toys). The hand puppet and the child would, then, have the same knowledge base. The hand puppet was used to make statements (her inferences) about the hypothesized contents of the closed box. Miss Snail, so the child was told, was a little silly, so she might not always get it right. The child's important job was to judge whether Miss Snail's statements were right or wrong. If the puppet was right, the child would reward her with a coin. If the puppet was wrong, the child would give her an "X". That is, if Miss Snail's inference was right (there must be an elephant in the box, given what we know about the contents of the two open boxes), she would get a coin. If she was wrong, however, she would get an X.

Miss Snail, so the child was told (again), was extremely shy and would therefore only whisper the statement in the experimenter's ear. The experimenter would then repeat out loud what the puppet's statement was. The reason for choosing this particular set-up is that, in this way, we could present embedded clauses as in (18) to the child. Embedded clauses were used to ensure that negation would always precede the modal operator on the surface, compare to this end (17) and (18). As such, the adult interpretation, in which the modal is interpreted in the scope of negation, would match the surface order. There is therefore another factor guiding towards the correct, weak, reading. If the child hypothesizes the strong reading, she would have to ignore the surface order of negation and the modal.

(17) Er hoeft niet een olifant in de doos te zitten.
There must not an elephant in the box to sit.
'There doesn't have to be an elephant in the box.'

(18) Mevrouw Slak zei dat er niet een olifant in de doos hoeft te zitten.
Miss Snail said that there not an elephant in the box must to sit.
'Miss Snail said that there doesn't have to be an elephant in the box.'

There were six different conditions: A POSITIVE condition (true and false) that included a universal modal, but no negation (*moeten* 'must'), a NEGATIVE STRONG condition (true and false) that contained a universal modal (*zullen* 'will') that takes wide scope over negation, and our crucial NEGATIVE WEAK condition (true and false) that included modal *hoeven* and negation. The choice for *moeten* 'must' for the universal modal in a positive declarative sentence was based on the fact that *hoeven* cannot occur without negation, hence we chose the universal *moeten*. *Moeten* can, and in our experiment did, select an epistemic modal base. The choice for *zal* 'will' in the Negative Strong Condition followed from the observation that, even though for most Dutch speakers *moeten* acts as a positive polarity item (PPI), i.e. taking wide scope over negation, this is not the case for all Dutch speakers and even speakers

for whom it is a PPI in main clauses allow *moeten* to scope under negation in embedded clauses (Iatridou and Zeijlstra 2013). *Zal* on the other hand, can only be interpreted as taking wide scope over negation (resulting in the strong NECESSARY > NOT reading). Again, in our experiment, *zal* is used epistemically.[6]

POSITIVE		
(19)	Mevrouw Slak zei dat er een olifant in de doos moet zitten. Miss Snail said that there an elephant in the box MUST sit. 'Miss Snail said that there must be an elephant in the box.'	TRUE
(20)	Mevrouw Slak zei dat er een varken in de doos moet zitten. Miss Snail said that there a pig in the box MUST sit. 'Miss Snail said that there must be a pig in the box.'	FALSE

NEGATIVE WEAK		
(21)	Mevrouw Slak zei dat er niet een varken in de doos hoeft te zitten. Miss Snail said that there not a pig in the box MUST+NPI to sit. 'Miss Snail said that there doesn't have to be a pig in the box.'	TRUE
(22)	Mevrouw Slak zei dat er niet een olifant in de doos hoeft te zitten. Miss Snail said that there not an elephant in the box MUST+NPI to sit. 'Miss Snail said that there doesn't have to be an elephant in the box.'	FALSE

NEGATIVE STRONG		
(23)	Mevrouw Slak zei dat er niet een kangoeroe in de doos zal zitten. Miss Snail said that there not a kangaroo in the box WILL sit. 'Miss Snail said that there will not be a kangaroo in the box.'	TRUE
(24)	Mevrouw Slak zei dat er niet een olifant in de doos zal zitten. Miss Snail said that there not an elephant in the box WILL sit. 'Miss Snail said that there will not be an elephant in the box.'	FALSE

Of course, these modals do differ in the frequency with which they occur. A search with CLAN (MacWhinney 2000) of the files for child Abel from the Groningen (CHILDES) corpus (Wijnen and Bol 1993) revealed a total of 24 occurrences for *hoeft*, a total of 95 occurrences for *zal*, and a whopping 1143 occurrences for *moet*. Note of course, that *hoeft* can only be used in combination with negation, in contrast to *moet* and *zal*.[7] We do not think, however, that the difference in frequency could affect the results of our experiment. For one thing, these are closed class words which occur more often than most open class lexical verbs. For instance, *hoeft* occurred much more frequently in the same corpus than the lexical verbs

[6] There is much more to say about the semantics of modal *zullen* 'will', but that is outside the scope of the present paper. See for a discussion on English *will* for instance Von Stechow (1995) and Cariani and Santorio (2017). What is crucial here is that "*niet zal*" has the strong reading of necessarily not.

[7] Furthermore, whereas *hoeft* is only used for 3rd and 2nd (only in declaratives) person singular, *moet* and *zal* are used in 1st, 2nd, and 3rd person singular.

Table 3 Percentage of correct answers per condition per age group

	Conditions					
	Positive		Negative Strong		Negative Weak	
	True (%)	False (%)	True (%)	False (%)	True (%)	False (%)
Adults	100	100	100	100	100	100
5-year-olds	97	66	91	100	53	85
6-year-olds	98	82	94	100	69	95

huilt 'cries' (3 occurrences), *springt* 'jumps' (2 occurrences), or *rent* 'runs' (0 occurrences). Moreover, as Lin et al. (2015) showed, *hoeven* is used by children from a very early age.

We used four test sets of three cardboard boxes. For each set of boxes, the hand puppet would produce six test sentences that contained modal statements, one per condition as illustrated in (19) through (24), resulting in a total of 24 sentences. Each child saw all items in all conditions. Two different lists were created that varied in order. Prior to the test phase, children were trained on two familiarization box sets (resulting in a total of six box sets). During the familiarization phase, children were asked to judge statements such as '*there is/isn't an X in the box*' relative to the closed box. In this phase (and only in this phase), children were also allowed to peek inside the closed box. Upon completion of the task, the children were rewarded with a sticker for participation.

Results

Responses were coded as right or wrong, according to adults' judgments as in (19–24). That is, correct would be the response 'true' for (19), but 'false' for (20). Incorrect would be the response 'false' for (19) and 'true' for (20). Four children were excluded from analysis, because they did not finish the task. Table 3 and Figure 3 summarize the percentage of correct responses per group per condition.

Table 3 and Figure 3 show that a sharp asymmetry emerges between the Negative Strong and the Negative Weak conditions: while in the Negative Strong conditions the proportion of correct judgments is above 90%, this proportion drops in the Negative Weak conditions. In particular, if we compare the Negative Strong True and the Negative Weak True sentences, we found that 5-year-olds accepted the former in 91% of the cases against a 53% acceptance rate in the latter. This asymmetry still holds in the 6-year-old participants, with 100% against 69%. The Negative Weak True condition is exactly the condition that was predicted to be difficult for children by the Semantic Subset Principle. This sentence (corresponding to (21) above) receives a 'true' answer under a weak (NOT > NECESSARY) interpretation, but yields a 'false' answer under a strong (NECESSARY > NOT) interpretation. Five-year-olds incorrectly rejected these sentences in 47% of the cases, suggesting they assigned a strong interpretation. At first glance then, these results lend support for the Semantic Subset Principle.

There is, however, another striking result that appears from Table 3 that needs to be discussed. Unexpectedly, 5-year-olds, and to some extent 6-year-olds, performed poorly in the Positive False condition (corresponding to (20)). That is, 5-year-old children judged (20) as if a pig is indeed obligatorily present in the closed box in 34% of the cases, even though it is not the case that a pig is present in both open boxes (it therefore *might* be in the closed box, but does not necessarily have to be). This suggests that there is something difficult about

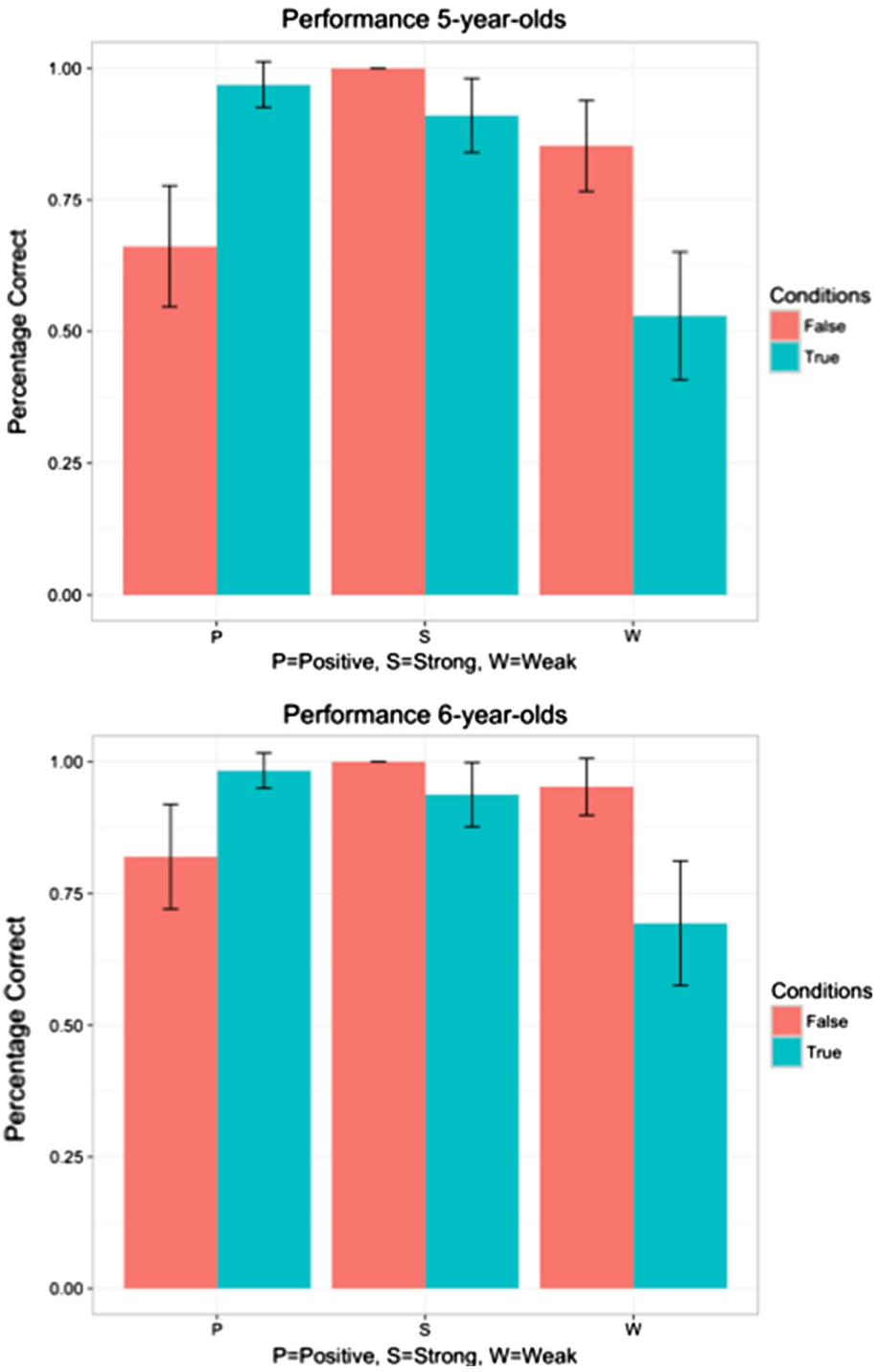

Fig. 3 Children's percentage of correct answers per condition split by age group. Error bars represent standard error of the mean

this condition that cannot be attributed to a difficulty with weak readings (there is only one logical operator, namely the modal, present in the positive sentences).

To test whether the difference in the proportion of correct responses between the Positive and the Negative Weak conditions was statistically significant, we analyzed the effect of condition on children's score in a multilevel regression analysis that included AGE (2 levels) and CONDITION (4 levels)[8] and their interaction as fixed effects as well as item, subject, and subject x condition random effects (see "Appendix A" for the final model and its output). As there were no significant interaction effects, we ran a model that included main effects of AGE and CONDITION only. The output for the main effects model showed that there was a main effect of CONDITION on score such that children, indeed, performed more poorly on Negative Weak True than on Positive True sentences ($b = -5.27(2.11)$, $p < .05$). There was also a main effect of AGE, such that 6-year-olds performed significantly better than 5-year-olds overall ($b = 1.17(0.48)$, $p < .05$).

The output of our model also supports the difficulty of the Positive False condition, showing that children scored significantly more poorly on Positive False than Positive True items ($b = -4.51(2.09)$, $p < .05$). This means that children also have difficulty with Positive False as compared to Positive True items as they do with Negative Weak True as compared to Positive True.

The idea that children's (poor) performance on the Negative Weak True items is not just the result of children assigning a strong (NECESSARY > NOT) instead of a weak (NOT > NECESSARY) interpretation is strengthened when one considers the individual data in Table 4. Specifically, under the hypothesis that children's performance is determined by incorrectly assigning a strong reading, we expect children to do so consistently, resulting in an overall score of 0% on Negative Weak True items. We could then still expect a score of 50% in case half of the children performed adult-like. When we look at the individual data in Table 4, however, it does not look like the overall score is based on children either failing or succeeding on all items. In fact, there was quite a large number of children who scored at 50%. Similarly, for the Positive False condition there were quite some children who scored 50%. Of course, one should not draw any conclusions on the basis of individual data, especially given that each child was presented with only 4 items per condition. The individual pattern does indicate, however, that the group mean is not composed of only children who consistently rejected or accepted the weak reading in the Negative Weak True condition. All in all, this suggests that there should be more going on than one would expect on the basis of the SSP alone. In the next section, we will propose an account for the observed pattern in terms of premature closure.

Premature Closure

In order to advocate that premature closure plays a role in the observed pattern, we will need to return to the asymmetry between epistemic and deontic modals observed in children's production (Kuczaj and Maratsos 1975; Shepherd 1982; Bliss 1988; Stromswold 1990; Stephany 1979; Wells 1979; Perkins 1983). A question that arises is why (if at all) children would experience more trouble with epistemic than with deontic modality.

[8] Children's overall responses in the Negative Strong Conditions showed minimal variance. This prevented convergence of GLM models. We therefore ran a model without the Negative Strong Conditions to directly compare Positive and Negative Weak conditions.

Table 4 Overview of individual response patterns

The number indicates the number of children with a particular mean score per condition. Note that the numbers do not always add up to 33 (the total number of children included in the analysis), because of missing data (and thus a score of for instance 66%)

	Conditions					
	Positive		Negative Strong		Negative Weak	
	True	False	True	False	True	False
0%	0	1	0	0	4	0
25%	0	3	0	0	6	1
50%	0	6	2	0	9	3
75%	3	7	4	0	3	4
100%	30	15	26	33	11	25

As Ozturk and Papafragou (2015) point out, the contexts in which epistemic possibility seem to be problematic for children are contexts that introduce multiple possible scenarios. Particularly, scenarios that introduce uncertainty where multiple outcomes are considered conceivable and one could not commit to one particular outcome without further evidence. Research has shown that one strategy children seem to use to overcome the difficulties associated with uncertainty is to reduce such uncertainty by eliminating competing alternatives (and thus committing to one particular alternative). The strategy of eliminating alternatives to reduce uncertainty was first described by Acredolo and Horobin (1987) (but see also Byrnes and Overton 1986; Robinson et al. 2006; a.o.) as children's propensity to offer only a single solution to any problem that, because of insufficient or ambiguous information, logically permits more than one result. This strategy, known under the name of *Premature Closure*, has been repeatedly observed in several studies, which do not necessarily involve modal expressions. In these studies, children's behaviour seems to be guided by a partial consideration of the multiple permissible outcomes (Somerville et al. 1979; Robinson and Whittaker 1986; Sophian and Somerville 1988; Taylor 1988; Beck and Robinson 2001).

One of the most recent evaluations of the premature closure hypothesis applied to the acquisition of modal expressions can be found in Moscati et al. (2017). The important contribution of this study is that it combined a measure of children's off-line truth-value judgments in a hidden-object task with an on-line measure of their eye-movements while processing sentences with modals. While the child heard the sentence, she watched a display of a hidden-object scenario that introduced uncertainty. Children's eye movements were measured to find out how different modal expressions directed children's visual exploration of the display. Figure 4 illustrates the hidden-object scenario used in Moscati et al. (2017).

Sentences with modals of different force were tested (compare (25) and (26)). In the experimental set up, the choice of the modal expression directly impacted the truth-conditions and the associated verification strategy:

(25)	A monkey might be in the orange box	(True –> verified by looking at the monkey)
(26)	A monkey must be in the orange box	(False –> falsified by looking at the pig)

In Moscati et al. (2017) two dependent measures were considered. The first was the proportion of correct truth-values at the end of the sentence; the second was the proportion of fixation to the unmentioned object (i.e. the pig) while hearing the sentence. Replicating previous findings, children's proportion of correct truth value judgments was significantly below the adult one for both conditions: some children mistakenly rejected (25) and accepted (26). Despite their poor performance in the evaluation task, children's eye-movements *were* sensitive to the

modal choice, with more fixations on the unmentioned object (the pig) in sentences like (26). Children's eye-movement pattern therefore suggests that children's difficulty in providing a truth value judgment does not lie in a lack of knowledge of the semantic difference between *must/might*. The source for children's poor performance in providing truth-values must thus lie elsewhere; most likely in children's inability to keep different alternative scenarios open and instead prematurely closing off alternatives.

Let us now consider if and how a premature closure explanation would account for the difficulties children had in our study. Children performed significantly more poorly on sentences (20) and (21) as compared to (19) (repeated here), but not differently on (22) as compared to (19). Children had to evaluate the truth of these sentences with regard to a closed box on the basis of the box set in (16).

(16) Box Set 1
Open Box 1: pig + elephant
Open Box 2: elephant
Outside: kangaroo

POSITIVE STRONG		
(19)	Mevrouw Slak zei dat er een olifant in de doos moet zitten. Miss Snail said that there an elephant in the box MUST sit.	TRUE
(20)	Mevrouw Slak zei dat er een varken in de doos moet zitten. Miss Snail said that there a pig in the box MUST sit.	FALSE

NEGATIVE WEAK		
(21)	Mevrouw Slak zei dat er niet een varken in de doos hoeft te zitten. Miss Snail said that there not a pig in the box MUST+NPI to sit.	TRUE
(22)	Mevrouw Slak zei dat er niet een olifant in de doos hoeft te zitten. Miss Snail said that there not an elephant in the box MUST+NPI to sit.	FALSE

The main tenet of the premature closure hypothesis is that in a context that introduces indeterminacy, children will arbitrarily select one possibility over the other. Applied to our

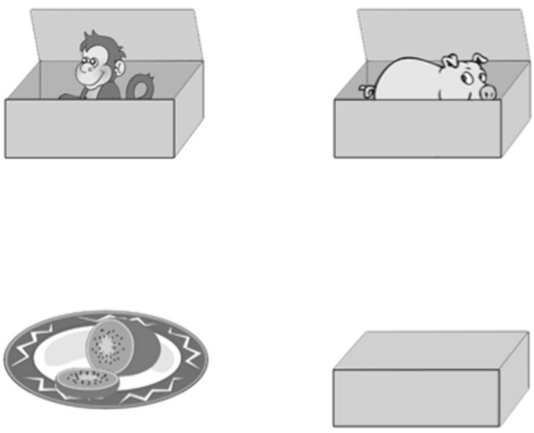

Fig. 4 A hidden object scenario (Moscati et al. 2017)

Table 5 Children's choice under premature closure

Children	Choice
Group A	Closed box = Pig + Elephant = Open Box 1
Group B	Closed box = Elephant = Open Box 2

experiment, this means that instead of inferring the contents of the mystery box in light of the two alternatives in (16), children will a priori choose either Open Box 1 *or* Open Box 2 to determine its contents. Children can thus be divided into two groups A and B, according to the option they chose (Table 5):

Given this initial choice, let us consider the truth-values of the sentences in (19–22). It is easy to verify that sentence (19) is true for children in both groups. Regardless of the fact that some children choose an elephant alone (Open Box 2) while some others an elephant and a pig together (Open Box 2), sentence (19) is always true. Premature closure, then, predicts that children will correctly accept sentence (19), as we observed. We now turn to sentence (20). When considering both alternatives, this sentence is false, as it is possible that a pig is not inside the mystery box. Children, however, only consider one alternative, either what is in Open Box 1 or what is in Open Box 2, as they close off the other alternative. This sentence is therefore false only for children in Group B. These children base their judgment on evaluating (20) in light of Open Box 2 that only contains an elephant, which makes sentence (20) false. For children in Group A, however, this sentence is true, as a pig is necessarily in the mystery box when evaluated only in light of Open Box 1. We therefore expect that the right truth-value is accessible by only half of the young participants. Following this reasoning, the asymmetry between (19) and (20) could be accounted for by premature closure. In particular, premature closure will only affect sentences for which the different boxes would give rise to a different response.

Similarly, premature closure predicts a contrast in evaluating sentences (21) and (22). Regardless of children's scope assignment and regardless of children's choice of box, sentence (22) will always be false as an elephant is present in both open boxes. For sentence (21), however, we do predict children to respond differently depending on which open box they chose (i.e. whether they belong to Group A or Group B). For children in Group A, a pig is necessarily present in the mystery box. For children in Group B, however, a pig is necessarily not part of the mystery box.

Importantly, premature closure is also in line with the individual patterns that we found. Assuming that children randomly close off one alternative on each separate trial, we would expect to get a 50% (individual) score over the four combined trials. Indeed, as discussed in the Results section, we found that there were quite some children who scored 50% (and also a number of children who performed adult-like). If the child is using a strategy of premature closure, this would thus affect children's performance exactly in the conditions they had difficulty with and in a way that we seem to observe, i.e. resulting in chance performance per child.[9]

A crucial question at this point is whether premature closure could serve as the sole explanation for children's poor performance in the Negative Weak True condition. This is a difficult question to answer. There are reasons why it does not appear to be the case that the findings can be attributed to premature closure entirely. One observation is that the

[9] Of course, an individual child might always prefer to close off one particular alternative. As mentioned, it is impossible to draw firm conclusions on the basis of the individual data. At the same time, the results are in line with what one would expect on the basis of a premature closure account.

Negative Weak True condition still seems to cause more trouble for the child participants than the Positive False condition. Twenty-two children scored 75–100% for Positive False as compared to fourteen on Negative Weak True. The same main effects model as presented in the results section, now with Positive False as the baseline, showed that there is a trend for an effect of Negative Weak True, such that children perform worse on Negative Weak True than Positive False sentences ($b = -0.76(0.40)$, $p = .053$). An additional important observation is that there are *no* children who made one or more mistakes on Positive False, but not on Negative Weak True sentences. Yet there are children (five in total) who made one or more mistakes on Negative Weak True, yet performed at 100% on Positive False sentences. Furthermore, there seem to be children who in fact *do* consistently reject the weak reading, resulting in 10 children who scored 0–25% on Negative Weak True compared to 4 children on Positive False. In the discussion, we will further elaborate on this point, also in relation to the NPI status of modal *hoeven*.

Discussion

This study tested the SSP in the domain of two logical operators: Dutch universal modal *hoeven* and negation. The SSP predicted that 5-year-old children will assign a strong (NECESSARY > NOT) reading to *hoeft niet*, as this reading is true in a narrower set of circumstances than its weak NOT > NECESSARY counterpart. In adult Dutch, however, this reading is ruled out as *hoeven* is an NPI and therefore necessarily scopes under negation, yielding the NOT > NECESSARY reading as the only possible one.

The results from the hidden-object task displayed that children indeed have particular difficulties with *hoeft niet*, in line with Moscati and Crain (2014). That is, like in Italian, children performed poorly on the condition that requires a weak interpretation, even though we are dealing with a different language and a different type of modal. Crucially, in Dutch, in contrast to Italian, scope assignment is independent of the order of modal and negation in the surface string. Should we consider children's result on the *hoeft niet* condition only, we could conclude that children are led by the SSP in interpreting potentially ambiguous sentences containing a modal operator and negation (as do Moscati and Crain (2014)).

The children in our experiment, however, did not only have difficulties with sentences in this particular condition. Children displayed poor performance in another type of sentence as well, namely the sentences belonging to the Positive False condition. The SSP does not predict any particular difficulty in this condition; in fact, it does not make any predictions about this type of sentences, as they do not involve negation (so only one logical operator, 'must'). To account for this result, we suggested that the source of children's difficulty was a strategy children adopt when dealing with situations in which multiple alternative outcomes are possible, known as premature closure. This would not only account for difficulties with Positive False sentences, but could also be a source for children's difficulty with Negative Weak True sentences. That is, Positive False and Negative Weak True sentences share the same level of indeterminacy. At the same time, however, the Negative Weak True condition seemed to pose an extra challenge for 5- and 6-year-olds.

A question that remains then is whether the additional difficulty in the Negative Weak True condition is simply the effect of the extra complexity negation introduces (as compared to positive sentences) or whether the SSP does still play a role. One way to find out is to run an experiment in which children's performance on Negative Weak True sentences is compared to their performance on negative sentences that only allow a strong interpretation of the

modal and negation (NECESSARY > NOT), yet introduces the same level of indeterminacy.[10] The premature closure hypothesis would predict these sentences to be equally difficult, as the level of indeterminacy is the same and both sentences involve negation. If SSP still plays a role, however, children should perform better on these sentences than Negative Weak True sentences.

One reason to think that premature closure should crucially be coupled with the SSP to account for the observed pattern comes from Negative Strong sentences with possibility modal *can* as tested by Moscati and Crain (2014). These authors tested sentences as in (27) in which negation precedes the modal, yielding a strong NOT > POSSIBLE interpretation.

(27)	NEGATIVE STRONG	
	a. Non ci può essere una fragola nella scatola 'There cannot be a strawberry in the box'	True
	b. Non ci può essere una mucca nella scatola 'There cannot be a cow in the box'	False

They observed that children had relatively few difficulties with the Negative Strong sentences in (27), with similar proportions of correct answers for (27a) and (27b). The SSP correctly predicts this pattern as in both sentences children would select the stronger reading. Under the premature closure account, however, there should have been an asymmetry between (27a) and (27b). Specifically, (27b) should have been more problematic than (27a) in a scenario in which a strawberry is present in neither of the two open boxes. As such, regardless of which box children choose to represent the contents of the closed box, they will derive the correct answer. (27b), however, will be true or false, depending on whether the child evaluates the sentence in light of an open box that contains a cow vs. an open box that does not contain a cow. Therefore, a lower rate of correct responses is predicted for (27b), contrary to facts. Thus, a premature closure account cannot account for the similarity in performance between sentences (27a) and (27b).

On the other hand, an open question is what the effect of NPI-hood is in acquisition. That is, the particular modal we tested in our experiment was a modal that lexically encodes polarity restrictions, different from the Italian case. The modal obligatorily combines with negation (some other downward entailing (DE) contexts are also acceptable) and the result is never ambiguous for the adult Dutch speaker: the modal necessarily reconstructs at LF to be interpreted within the scope of negation. The question is whether children at this stage have already mastered the polarity constraint on *hoeven:* if they haven't, the SSP comes into play as it guides the child towards the stronger interpretation. If, on the contrary, they do know the polarity restriction on *hoeven*, there is no ambiguity to begin with (and so no reason to allude to the SSP). A discussion of children's acquisition of polarity goes beyond the goal of the present paper, but we will nevertheless discuss different possibilities in light of the existing literature.

From Lin et al.'s study, we know that children know something about *hoeven* and negation – namely that they co-occur. Of course, for the language-acquiring child, the scope relations do not immediately follow from the input: a co-occurrence of the modal and negation does not entail that the modal is always interpreted in the scope of negation and so a potential

[10] Note that the Negative Strong items in the present experiment did not introduce the same uncertainty. These sentences would give rise to the same judgment, regardless of which box the child considered.

ambiguity can arise. The crucial question is thus: has the child already acquired the (semantic) polarity restriction?

Under semantic approaches to polarity, NPIs occur in the particular environments they occur in, *because of their semantics* (e.g., Kadmon and Landman 1993; Krifka 1994; Chierchia 2013). As a result of NPIs such as *any* being domain-wideners (activating alternatives), they must occur in DE environments in order to be informative. That is, their contribution is strengthened in DE environments by exhaustification of the subdomain alternatives (whereas exhaustification does not work in non-DE environments). The licensing conditions, then, follow from the particular semantics, and thus evidence that the child knows the distributional restrictions could be taken to show that they have acquired the right semantics for the NPI (see Tieu and Lidz 2016). As Tieu and Lidz point out, however, this conclusion would be too hasty. There is very little evidence in the input to the child for the semantics of a word like any for instance. Instead, they raise the possibility that children use the distribution information in their input as a guide towards the target semantics.

Now semantic approaches seem to cope well with NPIs in the nominal domain. Semantic approaches do, however, not fare well in the domain of universal modals that behave like NPIs (Iatridou and Zeijlstra 2013). In particular, NPI *hoeven* obligatorily scopes under negation, like NPIs in the nominal domain such as *any*, but, crucially (and different from *any*), *hoeven* has universal force. As a result, we derive the *weaker* NOT > NECESSARY interpretation instead of a strengthened interpretation under negation. A semantic account in terms of exhaustification of alternatives can thus not account for the distribution of *hoeven* in negative environments.

Syntactic approaches to NPI licensing are potentially better equipped in accounting for negative polar universal modals (Iatridou and Zeijlstra 2013). Under syntactic approaches, NPIs bear a negative feature that needs to be checked (e.g., Progovac 2005; Postal 2000; Guerzoni 2006). Hoeksema (2008) observed that licensing of universal model *hoeven* is more restricted than licensing of *any* for instance, such that not all DE operators can license the NPI modal. Iatridou and Zeijlstra (2013) suggest that it is in fact *only* environments that spell out negation separately that are acceptable licensing environments (see also Lin et al. 2017). This, as Iatridou and Zeijlstra (2013) suggest, is possibly the result of a syntactic constraint on NPI licensing for universal modals (the NEG feature must be checked by a separate negation). The co-occurrence of *hoeven* with negation in the child's input could then perhaps guide the child to postulate a NEG feature that requires checking as part of the lexical entry for *hoeven* (cf. Lin et al. 2017). Once the child has postulated a NEG feature that requires to be checked at LF, *hoeven* must scope under negation for the child.[11] Figuring out NPI-hood (adding a NEG feature), then, entails fixing the scope relations. The informationally structural weaker interpretation is allowed (and required) for the child, overriding the default setting that provides the informationally strongest interpretation, in order for the derivation to converge. Of course, this does not mean that the SSP does not come into play in the Italian case where the scope relations are not lexically fixed.

As a final remark before concluding this work, we would like to point out that there is an additional factor that makes universal modal NPIs difficult. This is the observation that they come with an additional inference, like modal *might*. The additional inference is the (scalar) inference that proposition p is at least possible (negating the stronger NECESSARY > NOT). That is, sentence (28) would be odd if uttered in a situation in which it is not even possible that there is a kangaroo in the box.

[11] Lin et al. (2017) hypothesize that the child first postulates that *hoeft* must always co-occur with *niet* and only later on (after age 4) reanalyses this as an abstract NEG.

(28) There doesn't have to be a kangaroo in the box.

As such, evaluating whether or not p is not necessary is a 2-step process. The first step is to figure out whether it is possible. Then one needs to figure out whether it is not the case that it is necessary (and so a different instruction for evaluation per box within the set). In this, it differs from evaluating necessarily not, for which we only need to establish whether [not p] in all cases. Whether children compute this scalar inference or not (perhaps they accept (28) in the context where there cannot be a kangaroo in the box) and how this would affect their behaviour is another question for future research.

Conclusion

This paper presented the results of an experiment investigating children's scope assignment in sentences with two logical operators: Dutch universal modal *hoeven* and negation. Our goal was to shed light on children's resolution of scope ambiguities in sentences with *hoeven* and negation, as such sentences provide an ideal testing ground for the Semantic Subset Principle (Crain et al. 1994; Crain 2012).

Concerning the role played by the SSP, our results are twofold. On the one hand, our findings support the SSP, as children have particular difficulties with sentences that are 'true' under a weak reading only. On the other hand, however, the observed response pattern in our experiment cannot be accounted for by the SSP alone. We suggested that children's performance in this particular task is affected by using a strategy in the face of uncertainty, namely prematurely closing off possible alternatives. We then continued to ask the question whether the SSP comes into play at all with items that lexically encode polarity restrictions (i.e. do children know it is an NPI?). Further research is required to tease apart the separate contributions of the SSP, negation, premature closure, and NPI-hood.

Finally, our experimental findings further show that children do not have any preference for surface scope interpretations, contra Musolino (2006a, b). Children's scope assignment is determined by different considerations than simple surface word order.

Acknowledgements We would like to thank the teachers and children at Montessori school Oosterhout and De Kring, Rijen for their enthusiastic participation in this study.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Appendix A

Output for the mixed-effects model. The baseline is the Positive True condition. Estimates indicate performance as compared to this baseline. This model converges with optimizer bobyqa and an increased number of iterations.

Fixed effects	Estimate (SE)	$p <$
Condition Positive False	− 4.51 (2.09)	.05
Condition Negative Weak True	− 5.27 (2.11)	.05
Condition Negative Weak False	− 2.37 (1.97)	n.s.
Age	1.17 (0.48)	.05

Final model: m.main = glmer(Correct ~ 1 + Condition + Age + (1|Subject) + (1|Item) + (1 + Condition|Subject), data = mydata, family = binomial, control = glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 100000))).

References

Acredolo, C., & Horobin, K. (1987). Development of relational reasoning and premature closure. *Developmental Psychology*, *23*(1), 13–21.

Beck, S. R., & Robinson, E. J. (2001). Children's ability to make tentative interpretations of ambiguous messages. *Journal of Experimental Child Psychology*, *79*, 95–114.

Berwick, R. C. (1982). *Locality principles and the acquisition of syntactic knowledge doctoral dissertation*. Cambridge: MIT.

Bliss, L. S. (1988). Modal usage by preschool children. *Journal of Applied Developmental Psychology*, *9*(3), 253–261.

Brown, R., & Hanlon, C. (1970). Derivational complexity and order of acquisition on child speech. In J. Hayes (Ed.), *Cognition and the development of language* (pp. 11–53). New York: Wiley.

Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press.

Byrnes, J. P., & Duff, M. A. (1989). Young children's comprehension of modal expressions. *Cognitive Development*, *4*(4), 369–387.

Byrnes, J. P., & Overton, W. E. (1986). Reasoning about certainty and uncertainty in concrete, causal, and propositional contexts. *Developmental Psychology*, *22*, 793–799.

Cardinaletti, A., & Shlonsky, U. (2004). Clitic positions and restructuring in Italian. *Linguistic Inquiry*, *35*(4), 519–557.

Cariani, F., & Santorio, P. (2017). Will done better: Selection semantics, future credence and indeterminacy. *Mind*, *127*(505), 129–165.

Carpenter, P. A., Just, M. A., Keller, T. A., Eddy, W. F., & Thulborn, K. R. (1999). Time course of fMRI-activation in language and spatial networks during sentence comprehension. *NeuroImage*, *10*(2), 216–24. https://doi.org/10.1006/nimg.1999.0465.

Chierchia, G. (2013). *Logic in grammar: Polarity, free choice, and intervention. Oxford studies in semantics and pragmatics* (Vol. 2). Oxford: Oxford University Press.

Cournane, A. (2015). Revisiting the epistemic gap: Evidence for a grammatical source. In *Proceedings of the 39th annual Boston University conference on language development*.

Crain, S. (2012). The emergence of meaning. Cambridge University Press.

Crain, S., Ni, W., & Conway, L. (1994). Learning, parsing and modularity. In C. Clifton, K. Rayner, & L. Frazier (Eds.), *Perspectives on sentence processing* (pp. 443–467). Hillsdale, NJ: Erlbaum.

Fischler, I., Bloom, P. A., Childers, D. G., Roucos, S. E., & Perry, N. W. (1983). Brain potentials related to stages of sentence verification. *Psychophysiology*, *20*(4), 400–409.

Gualmini, A., & Moscati, V. (2009). The early steps of modal and negation interactions: Evidence from child Italian. In E. Aboh, E. van der Linden, J. Quer, & P. Sleeman (Eds.), *Romance languages and linguistic theory 2007*. Amsterdam: John Benjamins.

Guerzoni, E. (2006). Intervention effects on NPIs and feature movement: Towards a unified account of intervention. *Natural Language Semantics*, *14*(4), 359–398.

Hacquard, V. (2010). On the event relativity of modal auxiliaries. *Natural Language Semantics*, *18*(1), 79–114.

Herbert, C., & Kübler, A. (2011). Dogs cannot bark: Event-related brain responses to true and false negated statements as indicators of higher-order conscious processing. *PloS One.*, *6*(10), e25574. https://doi.org/10.1371/journal.pone.0025574.

Hirst, W., & Weil, J. (1982). Acquisition of epistemic and deontic meaning of modals. *Journal of Child Language*, *9*, 659–66.

Hoeksema, J. (1997). Negation and negative concord in Middle Dutch. In D. Forget, et al. (Eds.), *Negation and polarity: Syntax and semantics* (pp. 139–158). Amsterdam: John Benjamins.
Hoeksema, J. (2008). Distributieprofielen van negatief-polaire uitdrukkingen: een vergelijking van het Nederlands, Engels en Duits. *TABU*, *37*(3/4), 111–195.
Iatridou, S., & Zeijlstra, H. (2013). Negation, polarity, and deontic modals. *Linguistic Inquiry*, *44*(4), 529–568.
Kadmon, N., & Landman, F. (1993). Any. *Linguistics and Philosophy*, *16*(4), 353–422.
Kratzer, A. (1981). The notional category of modality. In H.-J. Eikmeyer & H. Rieser (Eds.), *Words, worlds, and contexts. New approaches in word semantics* (pp. 38–74). Berlin: Mouton de Gruyter.
Kratzer, A. (1991). Modality. In A. von Stechow & D. Wunderlich (Eds.), *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung* (pp. 639–650). Berlin: Walter de Gruyter.
Krifka, M. (1994). The semantics and pragmatics of weak and strong polarity items in assertions. In M. Harvey & L. Santelmann (Eds.), *Proceedings from semantics and linguistic theory IV* (pp. 195–219). Ithaca: Cornell University, Department of Modern Languages and Linguistics.
Kuczaj, S. A., & Maratsos, M. P. (1975). What children can say before they will. *Merrill- Palmer Quarterly*, *21*, 89–111.
Lin, J., Weerman, F., & Zeijlstra, H. (2015). Emerging NPIs: The acquisition of Dutch hoeven 'need'. *The Linguistic Review*, *32*(2), 333–374.
Lin, J., Weerman, F., & Zeijlstra, H. (2017). Acquisition of the Dutch NPI hoeven 'need': From lexical frames to abstract knowledge. Language Acquisition, 1–28.
MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
Marcus, G. F. (1993). Negative evidence in language acquisition. *Cognition*, *46*(1), 53–85.
Miestamo, M. (2003). C*lausal negation: A typological study*. Doctoral dissertation, University of Helsinky.
Moore, C., Pure, K., & Furrow, D. (1990). Children's understanding of the modal expression of speaker certainty and uncertainty and its relation to the development of a representational theory of mind. *Child Development*, *61*(3), 722–730.
Moscati, V., & Crain, S. (2014). When negation and epistemic modality combine: the role of information strength in child language. *Language Learning and Development*, *10*(4), 345–380.
Moscati, V., & Gualmini, A. (2007). More facts that Isomorphism cannot explain. *In Semantics and Linguistic Theory*, *17*, 202–209.
Moscati, V., Zhan, L., & Zhou, P. (2017). Children's on-line processing of epistemic modals. *Journal of Child Language*, *44*(5), 1025–1040.
Musolino, J. (1998). *Universal grammar and the acquisition of semantic knowledge: An experimental investigation of quantifier-negation interaction in English*. Doctoral dissertation, University of Maryland.
Musolino, J. (2006a). Structure and meaning in the acquisition of scope. In V. van Geenhoven (Ed.), *Semantics in acquisition* (pp. 141–166). New York: Springer.
Musolino, J. (2006b). On the semantics of the subset principle. *Language Learning and Development*, *2*(3), 195–218.
Musolino, J., Crain, S., & Thornton, R. (2000). Navigating negative quantificational space. *Linguistics*, *38*(1), 1–32.
Musolino, J., & Lidz, J. (2003). The scope of isomorphism: Turning adults into children. *Language Acquisition*, *11*(4), 277–291.
Noveck, I. A. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, *78*(2), 165–188.
Noveck, I. A., Ho, S., & Sera, M. (1996). Children's understanding of epistemic modals. *Journal of Child Language*, *23*(3), 621–643.
O'Neill, D. K., & Atance, C. M. (2000). 'Maybe my Daddy give me a big piano': The development of children's use of modals to express uncertainty. *First Language*, *20*(58), 29–52.
Ozturk, O., & Papafragou, A. (2015). The acquisition of epistemic modality: From semantic meaning to pragmatic interpretation. *Language Learning and Development*, *11*(3), 191–214.
Palmer, F. R. (2001). *Mood and modality* (2nd ed.). Cambridge: Cambridge University Press.
Perkins, M. R. (1983). *Modal expressions in English*. London: Frances Pinter.
Postal, P. M. (2000). *The Ohio lectures on squat*. New York: Manuscript, New York University.
Progovac, L. (2005). *A syntax of Serbian*. Bloomington: Slavica Publishers.
Ramchand, G. (2012). *Indexical vs. anaphoric modals*. Tromsø: Ms, University of Tromsø.
Reinhart, T. (2004). The processing cost of reference-set computation: Acquisition of stress shift and focus. *Language Acquisition*, *12*(2), 109–155.
Rizzi, L. (1982). *Issues in Italian syntax*. Dordrecht: Foris.

Robinson, E. J., Rowley, M., Beck, S. R., Carroll, D. J., & Apperly, I. A. (2006). Children's sensitivity to their own relative ignorance: Handling of possibilities under conditions of epistemic and physical uncertainty. *Child Development*, *77*, 1642–1655.

Robinson, E. J., & Whittaker, S. J. (1986). Children's conceptions of meaning message relationships. *Cognition*, *22*, 41–60.

Shepherd, S. C. (1982). From deontic to epistemic: An analysis of modals in the history of English, creoles, and language acquisition. In A. Ahlquist (Ed.), *Papers from the fifth international conference on historical linguistics*. Amsterdam: Benjamins.

Slobin, D. I. (1966). Grammatical transformations and sentence comprehension in child and adulthood. *Journal of Verbal Learning and Verbal Behavior*, *5*(3), 219–227.

Somerville, S. C., Hadkinson, B. A., & Greenberg, C. (1979). Two levels of inferential behavior in young children. *Child Development*, *50*(1), 119–131.

Sophian, C., & Somerville, S. C. (1988). Early developments in logical reasoning: considering alternative possibilities. *Cognitive Development*, *3*, 183–222.

Stephany, U. (1979). Modality. In P. Fletcher & M. Garman (eds.), *Language Acquisition*, 375-400. Cambridge: Cambridge University Press. 2nd ed. 1986.

Stromswold, K. (1990). *Learnability and the acquisition of auxiliaries*. Ph.D. Doctoral dissertation, MIT, Cambridge.

Taylor, M. (1988). Conceptual perspective taking: Children's ability to distinguish what they know from what they see. *Child Development*, *59*, 703–718.

Tieu, L., & Lidz, J. (2016). NPI licensing and beyond: Children's knowledge of the semantics of any. *Language Acquisition*, *23*(4), 311–332.

Traugott, E. (1988). Pragmatic strengthening and grammaticalization. In *Proceedings of the fourteenth annual meeting of the Berkeley Linguistics Society* (pp. 406–416)

van der Wouden, T. (1994). *Negative Contexts*. Ph.D. dissertation, Groningen, The Netherlands.

van der Wouden, T. (1997). *Negative contexts: Collocation, polarity and multiple negation*. London: Routledge.

Von Stechow, A. (1995). On the proper treatment of tense. In V. T. Galloway & M. Simons (Eds.) *Proceedings of SALT* (Vol. 25). Cornell University.

Wason, P. C. (1965). The contexts of plausible denial. *Journal of Verbal Learning and Verbal Behavior*, *4*(1), 7–11.

Wells, G. (1979). Learning and using the auxiliary verbs in English. In V. Lee (Ed.), *Language development*. London: Open University Press.

Wijnen, F., & Bol, G. W. (1993). The escape from the optional infinitive stage. In A. de Boer, J. de Jong, & R. Landeweerd (Eds.), *Language and cognition 3*. Groningen: University of Groningen, Centre for Language and Cognition.

Zhou, P., & Crain, S. (2009). Scope assignment in child language: Evidence from the acquisition of Chinese. *Lingua*, *119*(7), 973–988.

Zwarts, F. (1981). Negatief Polaire Uitdrukkingen I'. *GLOT*, *4–1*, 35–132.

Zwarts, F. (1998). Three types of polarity. In Fritz Hamm & Erhard Hinrichs (Eds.), *Plurality and quantification* (pp. 177–238). Dordrecht: Kluwer.