Emerging NPIs

Jing Lin j.lin@uva.nl
(University of Amsterdam)
Fred Weerman f.p.weerman@uva.nl
(University of Amsterdam)
Hedde Zeijlstra hzeijls@uni-goettingen.de
(University of Göttingen)

Negative Polarity Items (NPIs) are lexical items that can only survive in certain negative contexts. An example is the English adverb *yet* in *I have *(not) finished yet*. NPIs are cross-linguistically attested (Haspelmath 1997) and come about in different strengths (Hoeksema 2000). Whereas weak NPIs (e.g. English indefinite *any*) are fine in all Downward Entailing (DE) contexts, stronger NPIs (e.g. Dutch modal *hoeven* 'need') are only allowed in a subset of DE contexts. NPIs even weaker than *any* (e.g. Mandarin indefinite *shenme* 'a (thing)') only occur in a superset of DE contexts. The existence of NPIs in natural languages raises an important learnability problem: how do language learners acquire the corresponding licensing conditions of NPIs without being confronted by any substantial or reliable negative evidence? After all, the absence of NPIs in non-licensing conditions in the input does not necessarily indicate their ungrammaticality in such contexts. In this paper we show that this learnability problem can be solved once a *conservative widening strategy* is adopted (c.f. Van der Wal 1996). Such a strategy relies on positive evidence only and assumes a widening process in acquisition that obeys the Subset Principle (Manzini & Wexler 1987), in that the output of an analysis in an earlier stage forms the subset of the output of its reanalysis in a later stage.
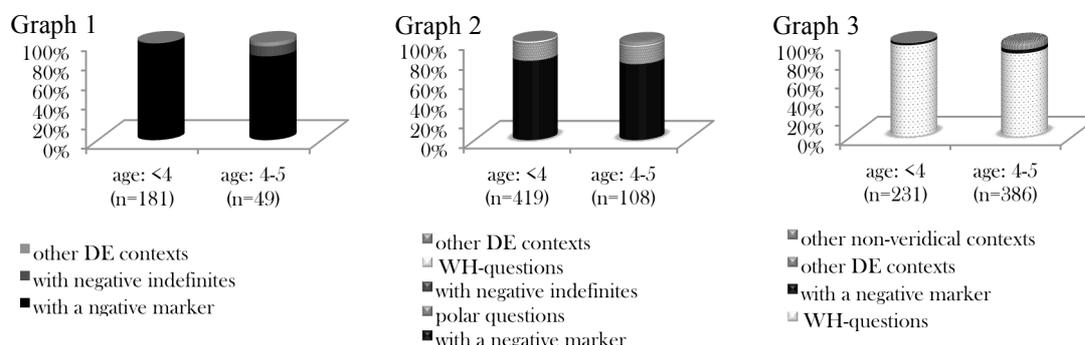
Under *a conservative widening strategy*, children start with the strictest possible analysis of a target NPI that is compatible with the limited input data available in the onset. If they receive input data that falsifies the strictness, they widen their initial analysis, making it less strict. Two logical predictions follow: (i) no overgeneralization errors should occur in NPI-licensing in acquisition; (ii) the distribution of NPIs should be broader in late child language than in early child language. These predictions have been tested by studying the spontaneous speech data of monolingual children between the ages of 1 and 5 in the CHILDES database (MacWhinney 2006). For Dutch, English and Mandarin, 744, 1822 and 734 CHAT files were investigated, respectively. The results show that the percentage of improperly licensed NPIs is zero in child Mandarin, and extremely low in child Dutch and English, too marginal to represent genuine overgeneralization errors. Hence, prediction (i) is confirmed.

Category	Proper	Improper	Unclear	TOTAL
Hoeven ('need') in child Dutch	237 (95.95%)	2 (0.81%)	4 (3.24%)	243
Any in child English	632 (99.21 %)	2 (0.31%)	3 (0.47%)	637
Shenme ('a (thing)') in child Mandarin	562 (99.8%)	0	1 (0.2%)	563

In each language, a significant difference is attested between children younger than 4 and those between 4 and 5 ($p=,000$, df=2 in Dutch; $p=,001$, df=4 in English; $p=,000$, df=6 in Mandarin), implying that older children use significantly more types of contexts to license target NPIs than younger children. This confirms prediction (ii): the NPIs' distribution is wider in later stages of child language. Since both predictions are confirmed, we conclude that all tested NPIs are acquired by means of the same conservative widening strategy, independent of their strength. However, distinct widening pathways are attested for NPIs of different strength. Dutch *hoeven* in the initial stage is only licensed by the negative marker *niet* 'not', whereas in a later stage, it is also licensed by negative indefinites (e.g. *geen* 'no',

niks 'nothing' (Graph 1). English *any*, by contrast, is only attested in early child English under the scope of the negative marker *not* or in polar questions (see also Tieu 2010 for similar results), and is not licensed in other DE contexts until the English children are at least 4 (Graph 2). *Shenme*, finally, is restricted to WH-questions in early child Mandarin, but is attested in all kinds of Non-Veridical contexts (NV) in late child Mandarin (Graph 3).

We take these differences in acquisitional pathways to represent different (re)analyzing processes in the acquisition of the selected NPIs. We show that the frequency of the input data is of crucial importance in establishing the initial analysis of the target NPIs, and that certain linguistic knowledge and the Subset Principle jointly guide the subsequent process of reanalysis. Moreover, we argue that what kind of an NPI (strong, weak, superweak) emerges at the end of acquisition, is determined not only by the most dominant type of NPI licensers in the input (i.e., negation in Dutch, negation and polar questions in English and WH-questions in Mandarin) but also the reanalyzing and widening process.

Graph 1

100%
80%
60%
40%
20%
0%
age: <4 (n=181) age: 4-5 (n=49)

▪ other DE contexts
▪ with negative indefinites
▪ with a ngative marker

Graph 2

100%
80%
60%
40%
20%
0%
age: <4 (n=419) age: 4-5 (n=108)

▪ other DE contexts
⌣ WH-questions
▪ with negative indefinites
▪ polar questions
▪ with a negative marker

Graph 3

100%
80%
60%
40%
20%
0%
age: <4 (n=231) age: 4-5 (n=386)

▪ other non-veridical contexts
▪ other DE contexts
▪ with a negative marker
⌣ WH-questions

When children are confronted with *hoeven* in combination with its licenser *niet,* which is a combination that is massively present in child-directed Dutch, they start analyzing this NPI as being lexically connected with the negative marker ([HOEVEN NIET]). But this initial analysis is not compatible with the input data in which *hoeven* is licensed by other licensers than *niet*, in particular negative indefinites. For this reason, *hoeven* is reanalyzed as being lexically associated with a negation that is also present in all negative indefinites (cf. Jabobs 1980, Zeijlstra 2011), suggesting *hoeven* is an NPI in the sense of Postal (2000). The reanalysis of [HOEVEN NEG] predicts *hoeven*'s grammatical appearance only in those DE contexts containing a decomposable negation, which is confirmed by Iatridou & Zeijlstra (t.a.).

In the acquisition of *any* a distinct (re)analyzing process is attested. Whereas children's initial analysis is that *any* may only appear in exactly those contexts that may trigger exhaustification themselves, in casu negative expressions introduced by *not* and polar questions (cf. Guerzoni & Sharvit 2007), *any* is later on reanalyzed as an element that is capable itself in triggering exhaustification. The reanalysis of *any* then amounts to assigning it a syntactic feature [uEXH]/[σ] that must be checked by a higher covert exhaustifier. If *any* is obligatorily exhaustified, it follows immediately why its occurrence is restricted to all DE contexts: only in DE-contexts does an exhaustified *any* not give rise to a semantic contradiction (cf. Kadmon & Landman 1993, Chierchia 2006, 2011). The fact that *any* is a weak NPI thus follows from the way it is acquired.

Mandarin children, finally, start with a narrow assumption of *shenme* being a WH-quantifier, as the most dominant licensing conditions for *shenme* is that of WH-questions in child-directed speech. Being confronted with input in which *shemne* is not always licensed in WH-questions, but also in a variety of NV-contexts, children must weaken down their initial analysis and reanalyze *shenme* as a non-referential existential quantifier. This is because only such quantifiers may survive exclusively in NV-contexts, as they cannot give rise to

existential import (Giannakidou 1998). Non-referential existential quantifiers form a supercategory of WH-quantifiers. Thus, in this reanalyzing and widening process, the Subset Principle holds. The reanalysis of *shenme* predicts that *shenme* must be a superweak NPI that may only survive in NV contexts, a prediction that is borne out (Lin 1998, Lin 2011, 2012).

The investigation of the learnability of NPIs leads to the following conclusions: (i) NPIs are acquired via the same *conservative widening learning strategy* irrespective of their strength; and (ii) different types of NPIs emerge as a result of language acquisition.

Selected References: Chierchia, G. 2006. 'Broaden your Views. Implicatures of Domain Widening and the Spontaneous Logicality of Language' *Linguistic Inquiry* 37: 535-590. | Hoeksema, J. 2000. 'Negative Polarity Items' In L. Horn & Y. Kato (eds) *Negation and Polarity: syntactic and semantic perspectives.* 115-146. Oxford: Oxford University Press. 837-859. | Manzini, R & Wexler, K. 1987. 'Parameters, Binding Theory, and Learnability'. *Linguistic Inquiry*. 18:413-444. | Postal, P. 2000. The Ohio Lectures on *squat*. Manuscript. New York University, New York. | Van der Wouden, T. 1994. *Negative contexts*. PhD Dissertation, University of Groningen. | Zwarts, F. 1981. 'Negatief polaire uitdrukkingen 1' *Glot* 4: 35–132.