

# Vietnamese Children's Interpretation of Definite Noun Phrases

Ni-La Lê, Hannah Forsythe, and Cristina Schmitt<sup>1</sup>

## 1. Introduction

In order to acquire the basic meanings of different types of noun phrases, children must establish mappings between sets of features (number, gender/class, definiteness, etc.) and the individual morphological pieces that make up the noun phrase (determiners, nominal inflections, etc.). This mapping is almost never one-to-one, and the fact that children still acquire these mappings quickly and efficiently is truly impressive. It is also an argument in favor of the existence of a biologically endowed capacity to learn language, which restricts the range of hypotheses that children consider when learning a language. Since the learner must be capable of acquiring any form-meaning mapping attested in natural language, our understanding of this device can be greatly enriched by studying how children acquire languages that realize the same semantic primitives in vastly different ways.

Vietnamese is a particularly useful language in which to study how the Language Acquisition Device (LAD) handles the mapping of definiteness and number. Unlike inflectional languages, which must encode number in nearly every noun phrase using a dedicated morpheme, Vietnamese does not always require number marking, and its plural morphemes (henceforth, pluralizers) have other properties besides encoding plurality, as their distribution depends on other syntactic and semantic properties of the noun phrase. In the same way, definiteness is not marked with dedicated determiners like *a* and *the*, but rather with a combination of classifiers and pluralizers. So it is an interesting language to compare against Western languages, which are far more well studied. This paper contributes data on the interpretation of singular and plural definite noun phrases by Vietnamese children ages 3 to 7 and compares their behavior to that of children in the same age range acquiring Spanish and English.

The paper is organized as follows. We begin with a description of the acquisition problem, followed by a summary of previous findings on the acquisition of definite noun phrases and a description of how number and definiteness are realized in Vietnamese. We then present a replication of a comprehension task first designed by Munn et al. (2006) for English- and Spanish-acquiring children. Our results show early ability to use definite cues and later ability to use plural information in comprehension tasks in Vietnamese, contrasting with English- and Spanish-acquiring children who can use number early but continue to make certain definiteness

---

<sup>1</sup> Michigan State University. Corresponding author: Ni-La Le, lenila@msu.edu. This research was supported by funding from the MSU Graduate School, the MSU College of Arts and Letters, and the MSU Department of Linguistics and Languages. Parts of this paper were published in *Proceedings of the 42nd Boston University Conference on Child Language Development (BUCLD42)*. Many thanks to Ho Chi Minh City University of Education and the teachers, parents, and administrators at three kindergartens in Ho Chi Minh City, Vietnam. Thanks to the audience of BUCLD42, as well as to Alan Munn, Karthik Durvasula, and the rest of the MSU language acquisition lab for insightful feedback.

errors. We close with some thoughts about how differences in the morphological realization of number and definiteness in Vietnamese on the one hand versus Spanish and English on the other could be responsible for the two different learning paths that children take.

## **2. The acquisition problem**

Regardless of how definiteness and number features are realized morphologically across languages, children must learn to extract their information. To interpret a definite noun phrase in any given language, the child must have at least three different pieces of information: first, the meaning of the noun and any accompanying modifiers; second, whether the noun phrase is singular or plural; and third, how to identify the discourse referent that *uniquely* satisfies these properties, in accordance with the uniqueness presupposition carried by the definite feature.

The uniqueness presupposition of a definite noun phrase is satisfied differently depending on its number. For plural definite noun phrases like *the dogs next to the tree*, the uniqueness presupposition is satisfied by finding a plural set of dogs-next-to-the-tree that is large enough to be unique, i.e., the maximal set of dogs next to the tree. In contrast, to interpret a singular definite noun phrase like *the dog next to the tree*, the uniqueness presupposition is satisfied by restricting the noun phrase's domain of reference to include only a single, unique dog close to the tree. This is easily satisfied in situations where there is only one dog next to the tree in the context, but if there is more than one, it is necessary to interpret the noun phrase more strictly. In other words, they must add an implicit restriction to accommodate the definite noun phrase to mean something like *the dog closest to the tree*.

This rather complex coordination of different pieces of information makes the acquisition of definite noun phrases a non-trivial task—and an especially interesting case to study across languages that realize number and definiteness differently. The next section describes previous work on the acquisition of definiteness, while the following sections expand on that literature by looking at a new language: Vietnamese.

## **3. Acquisition background**

Previous work on the acquisition of number and definiteness markers finds that number is acquired earlier than definiteness. English-speaking children master the conceptual distinction between one and more-than-one around 20-24 months of age (Fenson et al., 1994; Barner et al., 2007), and within the next year (24-36 months) they are producing the plural marker in the correct contexts and even use it in novel words (Brown, 1973; Mervis and Johnson, 1991; Kouider et al., 2006). What little work has been done on the acquisition of number in classifier languages like Mandarin, Korean and Japanese suggests that children take a longer time to acquire plural morphology, perhaps because number is not grammaticalized and/or because number markers tend to be portmanteau morphemes (Munn et al., 2009; Nakano et al., 2009).

As for definiteness, the acquisition path appears to be much more protracted. Children as old as 5 incorrectly use the definite determiner to refer to non-unique objects, saying things like

*Give me the ball!* even when multiple identical balls are present (Maratsos, 1976; Karmiloff-Smith, 1979; Schaeffer and Matthewson, 2005). However, comprehension tasks indicate that children do at least understand the contrast between definites and indefinites; they are aware that definites maintain reference while indefinites introduce new referents (Maratsos 1976; Karmiloff-Smith 1979; Modyanova and Wexler, 2007; de Cat, 2011).

A number of different hypotheses have been advanced to explain children's misuse of definites. Early proposals by Maratsos (1976) and Karmiloff-Smith (1979) suggested that errors stem from an egocentric tendency to use definite forms for referents under the child's own focus of attention, ignoring the interlocutors' attentional state.<sup>2</sup> Observing that children overextend definites even when no elements are in focus, Wexler (2003, 2011) proposes alternatively that definiteness errors arise because children's initial representation of the definite determiner lacks the uniqueness presupposition (dubbed the 'no Maximality Hypothesis,' Wexler, 2011: p. 25).

More recent work (Drozd, 2001; Munn et al., 2006) has proposed that children have the uniqueness presupposition but are simply unable to satisfy it when doing so requires them to implicitly restrict the domain of reference. This explains an otherwise puzzling result from Munn et al. (2006), who report that English- and Spanish-acquiring preschoolers were able to associate plural definite noun phrases like *the dogs next to the tree* with a unique plural (i.e., maximal set) but were unable to associate singular definite noun phrases like *the dog next to the tree* with a unique singleton set. The difference lies in the fact that the definite singular, but not the definite plural, requires the child to restrict implicitly the noun phrase's domain of reference to mean something like *the dog **closest** to the tree*.

A major limitation of these proposals is that they have been based almost exclusively on results from Western European languages, which realize definiteness through dedicated determiners (ex. *the, a*) and which have a grammaticalized binary number distinction. In order to get a more complete picture of the cross-linguistic acquisition of definite noun phrases, we study the comprehension of definite noun phrases in Vietnamese—a language with neither of these characteristics.

#### **4. Vietnamese noun phrases**

Like other classifier languages, Vietnamese allows bare noun phrases, which are underspecified for definiteness and number. Depending on the predicate and the context in question, bare nouns can have generic, existential, indefinite and definite readings, as well as singular and plural readings. Hence, a bare noun like *chó* ('dog') can mean either 'a dog', 'the dog', 'the dogs', or just 'dogs'.

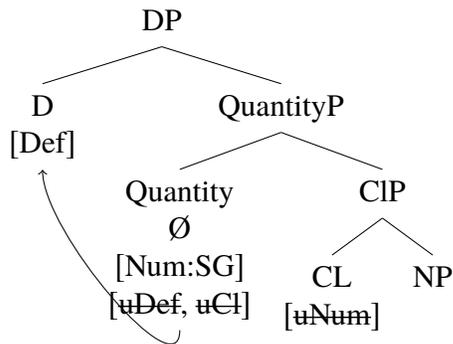
---

<sup>2</sup> These authors nevertheless assume different underlying reasons for children's 'egocentric' behavior. Maratsos suggests that children are aware that definites must signal specificity of reference for speakers but not necessarily hearers, while Karmiloff-Smith suggests that children have a more deictic representation of definite noun phrases than adults do.

Nouns preceded by a classifier (1a) are interpreted as singular and definite.<sup>3</sup> This fact might make it seem like Vietnamese classifiers are portmanteau morphemes, encoding both singularity and definiteness, but in fact, classifiers can also be found in indefinite noun phrases and in plural noun phrases. For example, adding the numeral *một* ‘one’ to the CL-N sequence forces a singular indefinite interpretation (1b), while adding the pluralizer *các* triggers a plural, definite interpretation (1c).<sup>4</sup> In Lê and Schmitt (2016) we argued for the structures in (2a-c) for the singular definite, singular indefinite, and plural definite noun phrases.

- (1) a. con chó  
 CL dog  
 ‘the dog’  
 b. một con chó  
 a/one CL dog  
 ‘a dog’  
 c. các con chó  
 CAC-PL CL dog  
 ‘the dogs’

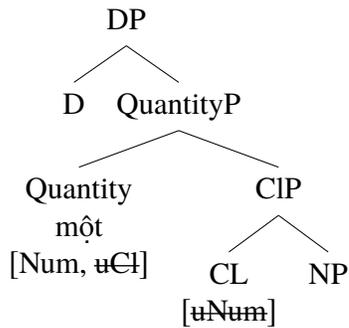
- (2) a. Structure of a Vietnamese Noun Phrase containing CL-N (singular, definite):



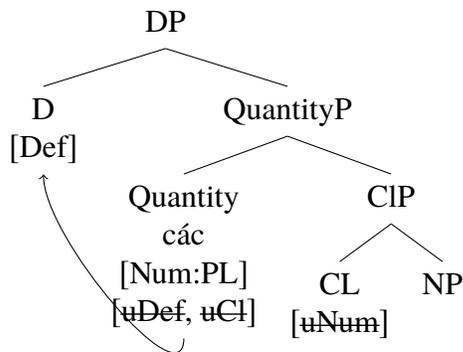
<sup>3</sup> This is different from most other classifier languages (ex. Japanese, Korean, Thai) which do not allow a CL-N sequence to appear in isolation.

<sup>4</sup> Again, Vietnamese differs from other classifier languages, in that the pluralizer requires the presence of a classifier.

b. Structure of a Vietnamese Noun Phrase containing *một*-CL-N (singular, indefinite):



c. Structure of a Vietnamese Noun Phrase containing *các*-CL-N:

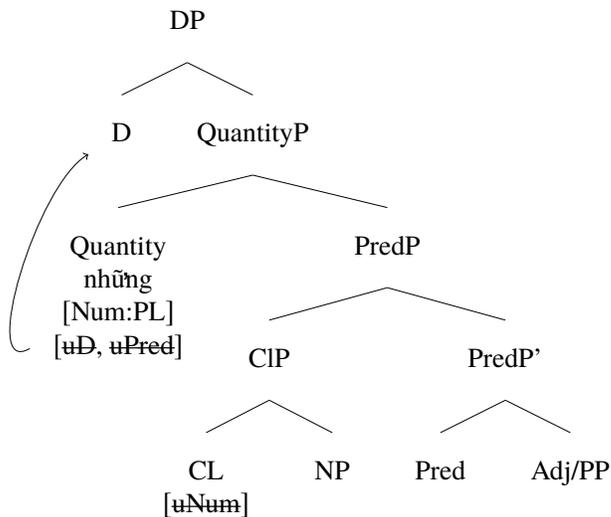


Vietnamese also has a second pluralizer *những*, which triggers a plural interpretation but whose definiteness status remains unclear. One of its striking properties is that *những* imposes the additional requirement that the noun be modified, as illustrated in (3a). While all studies agree that the noun phrases pluralized by *các* are definite (Nguyen, Tai Can, 1975; Nguyen, Hung Tuong, 2004; among others), most authors claim that the pluralizer *những* indicates only a subset of a given set (Thompson, 1965; Nguyen, Tai Can, 1975) and that its interpretation is of a specific indefinite (Nguyen, Hung Tuong, 2004). However, others argue that the interpretation of *những* varies on a continuum somewhere between indefinite and definite (Cao, 1998; Bui, 2000), while in Lê and Schmitt (2016), we argue that noun phrases with *những* are not inherently definite, but rather vary with the context, being able to appear also in indefinite contexts such as question words, as illustrated in (3b).

- (3) a. *những con chó \*(bự)*  
 NHUNG-PL CL dog big  
 ‘the big dogs’
- b. *Những quyển sách nào cần chuyển đi?*  
 NHUNG-PL CL book which need move go  
 ‘Which books need to be moved?’

In Lê and Schmitt (2016: p.170), we propose a simplified structure, as in (4), which could account for the properties of the pluralizer *những*. Crucial here is that the pluralizer has a D feature but this feature is not specified in the morpheme itself and that *những* selects not for a NP but rather for a small clause of sorts.

(4) Structure of a Vietnamese Noun Phrase containing *những*-CL-N-Modifier:



Finally, the quantifier *tất cả* ('all') can be added to a pluralized noun phrase to produce a maximal reading. Once again, if the pluralizer used is *những*, the nominal must be modified (5b).

- (5) a. *tất cả các con chó*  
 all CAC-PL CL dog  
 'all the dogs'
- b. *tất cả những con chó \*(bự)*  
 all NHUNG-PL CL dog big  
 'all the big dogs'

We begin by asking whether children can associate definite noun phrases like those in (1) and (3a) to sets with the correct number and definiteness properties.

## 5. Research questions

We choose to replicate Munn et al.'s (2006) task testing the comprehension of singular and plural definite noun phrases because this will allow us to make cross-linguistic comparisons with English- and Spanish-acquiring children. In particular, we will be able to address three questions:

- Q1:** Do Vietnamese-acquiring children know number, i.e., do they associate CL-N sequences to singleton sets and *các/những*-CL-N sequences to plural sets?

- Q2:** Do Vietnamese-acquiring children know definiteness, i.e., do they associate both CL-N and *các*-CL-N sequences to unique sets? What about *những*-CL-N sequences?
- Q3:** How does their behavior compare to children acquiring other languages such as English and Spanish?

## 6. Hypotheses and predictions

Since the interpretation of definiteness is dependent on number properties, a logical hypothesis is that number will be acquired before definiteness, across languages. If so, we would expect Vietnamese children to have the same behavior as English- and Spanish-acquiring children who participated in this task, showing adult-like interpretation of number morphology but committing at least some definiteness errors. However, the ‘No Maximality’ and ‘No Implicit Domain Restriction’ hypotheses make different predictions about what those definiteness errors should look like. If children lack the uniqueness presupposition (per ‘No Maximality,’ Wexler, 2003; 2011) then they should associate singular definite noun phrases also to non-unique singleton sets and plural definite noun phrases to non-maximal plural sets. Alternately, if children do have the uniqueness presupposition but instead have trouble with some form of domain restriction (per ‘No Domain Restriction,’ Drozd, 2001; Munn et al., 2006), then they should produce more definiteness errors in the singular condition.

On the other hand, it is also reasonable to hypothesize that the ability to use number and definiteness in comprehension tasks is dependent on the morphological realization of these features in the target language. For Vietnamese, this could lead to one of two scenarios. One possibility is that the general lack of a one-to-one correspondence between individual morphemes and individual number and definiteness features delays Vietnamese children’s acquisition of both features. If so, they should fail to distinguish between singular and plural definite noun phrases and show no tendency to associate either one to a unique/maximal set. The other possibility is that children initially associate the classifier with definiteness. If so, they should treat any noun phrase with a classifier as definite and ignore the singular-plural distinction until later on in acquisition.

## 7. Experiment 1

### 7.1. Subjects

99 children were recruited from three kindergartens in Ho Chi Minh City, Vietnam, with 34 exclusions for refusal to participate or failure to name relevant animals and landmarks during the pretest. Of the remaining 65 children who finished the test, 7 were excluded from analysis for failing to provide at least 3 out of 8 correct answers during training and control conditions. The data presented here is from 58 subjects, including 11 3-year-olds (M=3;7, range: 3;2-3;11), 12 4-year-olds (M=4;5, range: 4;0-4;10), 15 5-year-olds (M=5;4, range: 5;0-5;11), 13 6-year-olds (M=6;6, range: 6;0-6;11), and 7 7-year-olds (M=7;4, range: 7;0-7;7). 8 native Vietnamese-

speaking adults (ages 25 to 47), who were studying or working at Michigan State University, also participated as controls.

## 7.2. Materials

Our experiment was a replication of the act-out task reported in Munn et al. (2006), illustrated in Figure 1. For logistic and cultural reasons, we used a toy tree (rather than a toy barn) as one of the two landmarks, and we used animal types familiar to Vietnamese children: *mèo* ('cat'), *chó* ('dog'), *gà* ('rooster'), and *cá* ('fish'). Additionally, we used groups of four animals per side, rather than the original three.<sup>5</sup>

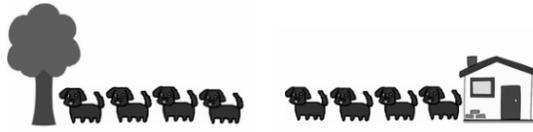


Figure 1. *Experimental setup*

Participants were instructed to choose the appropriate animal or animals using prompts as in (6). Three test conditions used definite singular and definite plural noun phrases (the latter containing either the pluralizer *các* or *những*). Three control conditions used noun phrases which either lexically specified a singleton set using the numeral *one* or lexically specified the maximal set with the quantifier *all*, accompanied by a pluralizer (*các* or *những*). Target responses are shown in Table 1.

### (6) *Example test items*

- a. Đưa cho cô { $\emptyset$  /*các* /*những*} con chó đứng kế cái cây  
 Give for aunt { $\emptyset$  /CAC-PL /NHUNG-PL} CL dog stand next CL tree  
 'Give me the dog/dogs next to the tree.'

### *Example control items*

- b. Đưa cho cô {*một* /*tất cả* các /*tất cả* những} con chó đứng kế cái cây  
 Give for aunt {one /all PI /all PI} CL dog stand next CL tree  
 'Give me one/all the dogs next to the tree.'

<sup>5</sup> The rationale for this change was that some studies show an effect of set size on children's comprehension of definites (Modyanova & Wexler 2007; de Cat, 2011). A replication of this task, using only three animals per side, as did Munn et al. (2006), will be discussed in section 8 below.

**Table 1.** *Noun phrase types used in Experiment 1*

	Condition:	Noun phrase type	Target referent
Experimental	SG-def	<b>CL N</b>	closest dog
	PL-def	<i>các</i> CL N	all the dogs
	PL-def (?)	<i>những</i> CL N	all the dogs (?) <sup>6</sup>
Control	SG-indef ‘one’	<i>một</i> CL N	any single dog
	all + <i>các</i>	<i>tất cả các</i> CL-N	all the dogs
	all + <i>những</i>	<i>tất cả những</i> CL-N	all the dogs

The full set of prompts was generated by crossing these 6 noun phrase types with 4 animal types and 2 landmark types to produce 48 items. Each child was tested on 12 items (2 of each noun phrase type) plus 4 fillers.

Participants were randomly assigned to one of 4 different versions of the 12-item test, each with a different order. In all versions, control items were presented **after** experimental items to prevent children from developing a contrast strategy for interpreting the experimental items. In 3 of the 4 versions, experimental items were presented in blocks, with either the *những*-CL-N block first (version 1), the *các*-CL-N block first (version 2), or the CL-N block first (version 3), and items ordered randomly within each block. In the remaining version, each block contained a *những*-CL-N item, a *các*-CL-N item, and a CL-N item in a random order. No significant differences were found between versions; therefore, we collapse them when reporting the results.

### 7.3. Procedure

The task included three phases: (i) a pretest, in which children were asked to name the animals and landmarks and to demonstrate their understanding of the prepositional phrase *đứng kế* ‘standing next to’ and the overall setup by answering the question *Đứng kế con chó/mèo/etc. là con gì?* ‘What is standing next to the dog/cat/etc.?’; (ii) the training, in which children responded to prompts using the numerals 2 through 4, and (iii) test and control items. The entire procedure lasted approximately 30 minutes in total. Adults were tested, either individually or in a group, using a pencil and paper version of the task without the pretest or training phase.

### 7.4. Results

Overall, adults performed as expected, producing 100% target responses in all control and test conditions. Importantly however, in the singular control condition (ex. *Give me **one CL dog next to the tree***), which had multiple potential target responses, adults chose the closest singleton dog for all trials, which is also the answer expected in the CL-N condition. In the plural test condition with pluralizer *những*, whose definiteness status was unclear, adults chose the

<sup>6</sup> If *những* is indefinite, subjects could pick out non-maximal sets. However, as discussed in Lê and Schmitt (2016), there is a near-categorical preference for interpreting *những* noun phrases as definite, contra previous claims made in the literature.

plural definite response (maximal set of dogs) for all trials. We present children’s results below, beginning with control items and then presenting responses to test items. As there was no significant difference across the four versions of the test, we collapsed the results below into a single group.

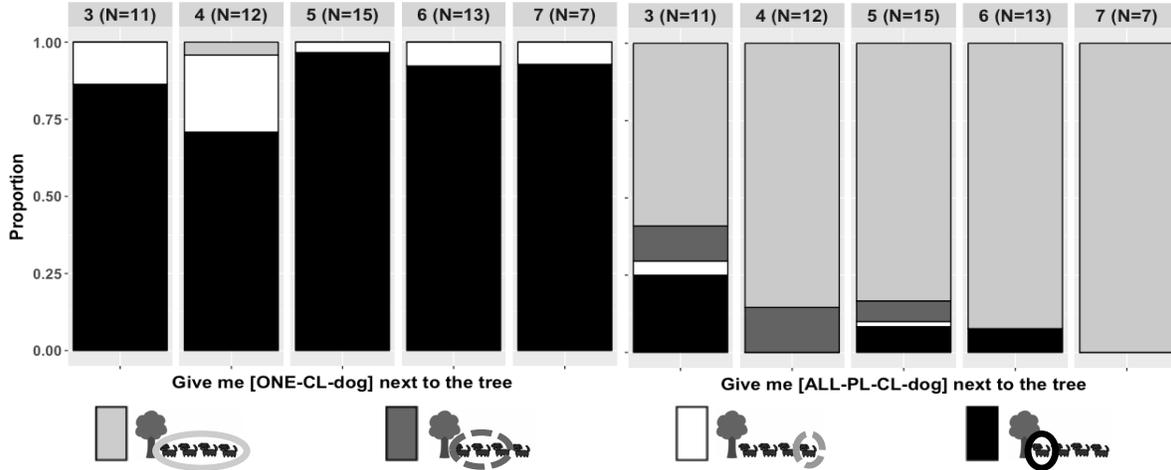
#### 7.4.1. Control sentences

Table 2 gives the percentage and frequency of children’s responses in control conditions, grouped by number (singular vs. plural responses) and definiteness (closest/maximal vs. non-closest/non-maximal responses). Expected responses are in shaded cells. Across all three conditions, children provided mostly expected responses, showing that they understood the task. In the singular control condition, children showed a distinct preference for the closest animal, similarly to adults, something which should be taken into account when assessing children’s behavior in the ‘one’-CL-N experimental condition.

**Table 2.** Percentage (frequency) of children’s response types in control conditions

		Condition: one-CL-N	all- <i>các</i> -CL-N	all- <i>những</i> -CL-N
Response type				
Singular	Closest	87.93% (102)	7.76% (9)	9.48% (11)
Singular	Non-closest	11.21% (13)	4.31% (3)	0% (0)
Plural	Maximal	0.86% (1)	81.90% (95)	84.48% (98)
Plural	Non-maximal	0% (0)	7.76% (9)	6.03% (7)

Figure 2 divides children’s responses by age group, collapsing across the two plural control conditions (all-*các*-CL-N and all-*những*-CL-N sequences). Even children in the youngest age group produced majority target responses.



**Figure 2.** Proportion of children’s response types, by age group, in singular control condition (left) and plural control conditions (right). Response types include singleton closest to the landmark (black), singleton but not the closest one (white), plural proper subset (medium gray) and plural maximal set (light grey).

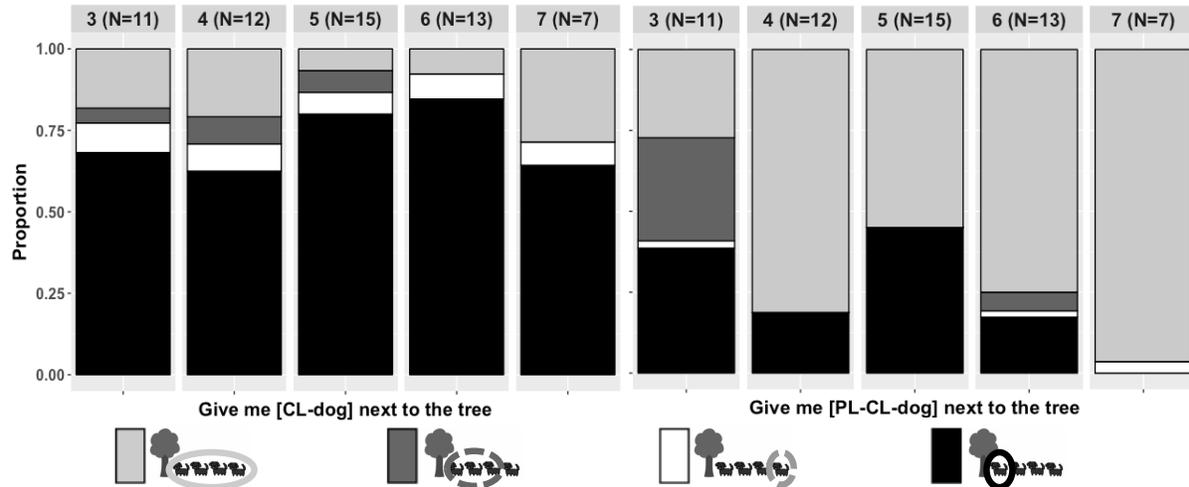
#### 7.4.2. Experimental sentences

Table 3 presents the percentage and frequency of children’s responses in the three experimental conditions. Like adults, children treated *các*-CL-N and *những*-CL-N sequences similarly, with no significant differences in the distribution of responses ( $\chi$ -squared = 0.39, df = 3,  $p$  = 0.94). Thus, we collapse across these conditions in all subsequent analyses.

**Table 3.** Percentage (frequency) of children’s response types in experimental conditions

Response type		Condition:	Sg. Def.	Pl. Def.	Pl. Def (?)
			CL-N	<i>các</i> -CL-N	<i>những</i> -CL-N
Singular	Closest		73.27% (85)	26.72% (31)	26.72% (31)
Singular	Non-closest		7.76% (9)	0.86% (1)	1.72% (2)
Plural	Maximal		14.66% (17)	64.66% (75)	64.66% (75)
Plural	Non-maximal		4.31% (3)	7.76% (9)	6.90% (8)

Figure 3 divides children’s responses by age group, collapsing across the two plural conditions. In the singular definite condition (ex. *Give me CL dog next to the tree*), children of all age groups produced a majority of target responses, choosing the closest dog to the tree. This is not surprising, given their preference for this response in the singular control trials. Interestingly, however, the second most common response—and therefore their most common error—was the maximal set of dogs, which is the expected answer if they are looking for the maximal set possible, independent of number.



**Figure 3.** Proportion of children's response types, by age group, in singular (left) and plural experimental conditions (right). Response types include singleton closest to the landmark (black), singleton but not the closest one (white), plural proper subset (medium gray) and plural maximal set (light grey).

Turning to the plural conditions (ex. *Give me các/những* CL dog next to the tree), we find a similar pattern. Children 4 and older produced a majority of target answers, i.e., they chose the maximal set of dogs. And just as in the singular condition, their most common error was a definite response of the wrong number, namely the closest single dog. Children age 3 were evenly split between number and definiteness errors.<sup>7</sup>

We now turn to children's comprehension of number and then of definiteness, examined independently.

### 7.4.3. Results: Number

To address the question of whether children distinguish singular from plural definite noun phrases, we compared the proportion of plural responses provided in singular (CL-N) test conditions relative to plural (*các/những*-CL-N) test conditions, collapsing across definite and indefinite responses. Two-tailed t-tests revealed significantly more plural responses in plural relative to singular conditions, for every age group (all  $p < 0.05$ ).

Despite distinguishing between singular and plural, however, 3- and 5-year-olds produce a surprisingly low rate of plural responses in the plural conditions (59% and 55%, respectively), at least compared with what has been reported on children's number comprehension in other languages and their own behavior in the control condition. Indeed, while Munn et al. (2006) report 80-95% target number responses in this task among English- and Spanish-acquiring children ages 5 and younger, the Vietnamese-acquiring children did not reach that same level of

<sup>7</sup> The split mostly occurred between subjects, not within. Four children were responsible for most of the plural indefinite responses, while four different children were responsible for most of the singular indefinite responses.

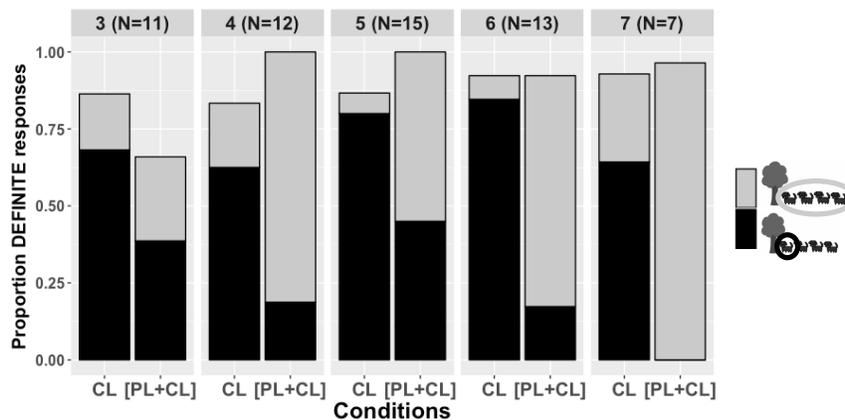
number accuracy until age 6 (see Table 4). This is in line with what has been found for pluralizer comprehension in Mandarin (Munn et al., 2009), Japanese, and Korean (Nakano et al., 2009).

**Table 4.** Proportion target number, collapsing across definite and indefinite responses. Experimental conditions of English and Spanish versions (Munn et al., 2006) and our Vietnamese-language Experiment 1.

Language and age group	SG-Def	PL-Def
English (N=15) 3;0-5;5	83.8	81.3
Spanish (N=20) 3;2-4;11	90	95
Vietnamese (N=11) 3;2-3;11	77	59
Vietnamese (N=12) 4;0-4;10	71	81
Vietnamese (N=15) 5;0-5;11	87	55
Vietnamese (N=13) 6;0-6;11	92	81
Vietnamese (N=7) 7;0-7;7	71	96

#### 7.4.4. Results: Definiteness

Turning to the second research question, we assessed Vietnamese children’s comprehension of definiteness by examining the proportion of definite responses they produced in the test conditions, whether plural (i.e., the maximal set) or singular (i.e., the closest animal to the landmark). For convenience, we refer to these as plural definite and singular definite responses, although we are aware that strictly speaking such terms are an interpretation of children’s behavior. Figure 4 presents the proportion of plural definite responses in grey and singular definite responses in black. Children showed a strong tendency toward definite responses—even if these responses did not always match the number of the noun phrase in question. Even the 3-year-olds produced a majority of definite responses (86.3% in the singular condition and 65.9% in plural conditions).



**Figure 4.** Proportion plural maximal responses (grey) and singleton responses closest to the landmark item (black) in experimental conditions: CL-N (singular, definite; left) and PL-CL-N (plural, right).

To check for developmental trends in the rate of what we refer to as definite responses, we ran a two-way ANOVA with condition (singular, plural) as a within-subjects factor and age group (3, 4, 5, 6, 7) as a between-subjects factor. There was a significant effect of age group ( $F = 8.049$ ,  $p < 0.001$ ), reflecting an overall increase in definite responses over time, as well as a significant interaction between condition and age group ( $F = 4.219$ ,  $p < 0.01$ ). Sub-t-tests between age groups, using Bonferroni-adjusted alpha levels of 0.005 per test (alpha 0.05 divided by 10 comparisons) revealed that the rate of definite responses produced in singular conditions remained equally high across age groups (all  $t > -0.952$ , all  $p > 0.346$ ), while in plural conditions 3-year-olds produced fewer definite responses relative to other age groups (all  $t < -3.245$ , all  $p < 0.002$ ). In sum, the overall rate of definite responses remains high over the course of development, with only 3-year-olds producing fewer definite responses than the rest—and then only in the plural conditions.

What does change over the course of development instead appears to be the ratio of *plural* definite responses (in grey) relative to *singular* definite responses (in black). In the plural conditions, the ratio of plural definite responses increases from 41.38% among 3-year-olds to 100% among 7-year-olds, as children gradually learn that *các*-CL-N and *những*-CL-N sequences encode plurality. In the singular conditions, the ratio of plural maximal responses remains low throughout, between 7.69% (5-year-olds) and 30.77% (7-year-olds), as even the youngest children seem to assume that CL-N sequences are to be interpreted as definite and singular.

## 7.5. Discussion

The major difference between our results and those of Munn et al. (2006) is that Vietnamese children commit number errors rather than definiteness errors. Spanish- and English-acquiring children from Munn et al. (2006) produced number-target responses across singular and plural conditions and failed to produce definite responses in the singular definite condition. In contrast, children in our study committed more number errors but have a preference for answers that are compatible with a definite interpretation in both the singular and the plural condition.

However, one might argue that the reason that the Vietnamese children in our study produced target responses in the singular definite condition is being over interpreted since that is their default preference also for the indefinite singular control. Recall that in the singular *indefinite* control conditions of our study, both adults and children showed a distinct preference for the animal closest to the relevant landmark—even though an equally acceptable response would have been to choose any other single animal on that side of the display. Carried over to the definite singular experimental condition, this default preference may have resulted in children choosing the right response for the wrong reasons.

But why would Vietnamese children have a default preference for the animal closest to the landmark, while the English- and Spanish-acquiring children showed a preference for the animal closest to themselves? Perhaps our decision to use 4 animals per landmark rather than 3 made a difference. A restriction had to be imposed for the singular and the item closest to the

child was too far to also satisfy close to the landmark. After all, the animal closest to the participant is the animal farthest from the landmark, and if 4 animals are used instead of 3, this may put that animal into a grey area no longer considered close enough to count as “next to” the landmark. To see if this manipulation would make a difference, we repeated the study once more with 3 animals per side.

## 8. Experiment 2

### 8.1. Subjects

56 children were recruited from kindergartens in Ho Chi Minh City, Vietnam, with 2 exclusions for refusal to participate and 12 exclusions for failure to provide at least 2 correct responses for the 6 control trials. The remaining 42 children (mean: 4;5, range: 2;7-5;6, 22 males) comprised 12 3-year-olds (M=3;3, range: 2;7-3;11), 15 4-year-olds (M=4;5, range: 4;0-4;10), and 15 5-year-olds (M=5;4, range: 5;0-5;7). Adults include 20 participants ages 17 to 42 (4 males) from Ho Chi Minh City.

### 8.2. Methods

Materials and design were identical to Experiment 1, except that 3 animals instead of 4 were presented next to each landmark, and the training phase was omitted.

### 8.3. Results

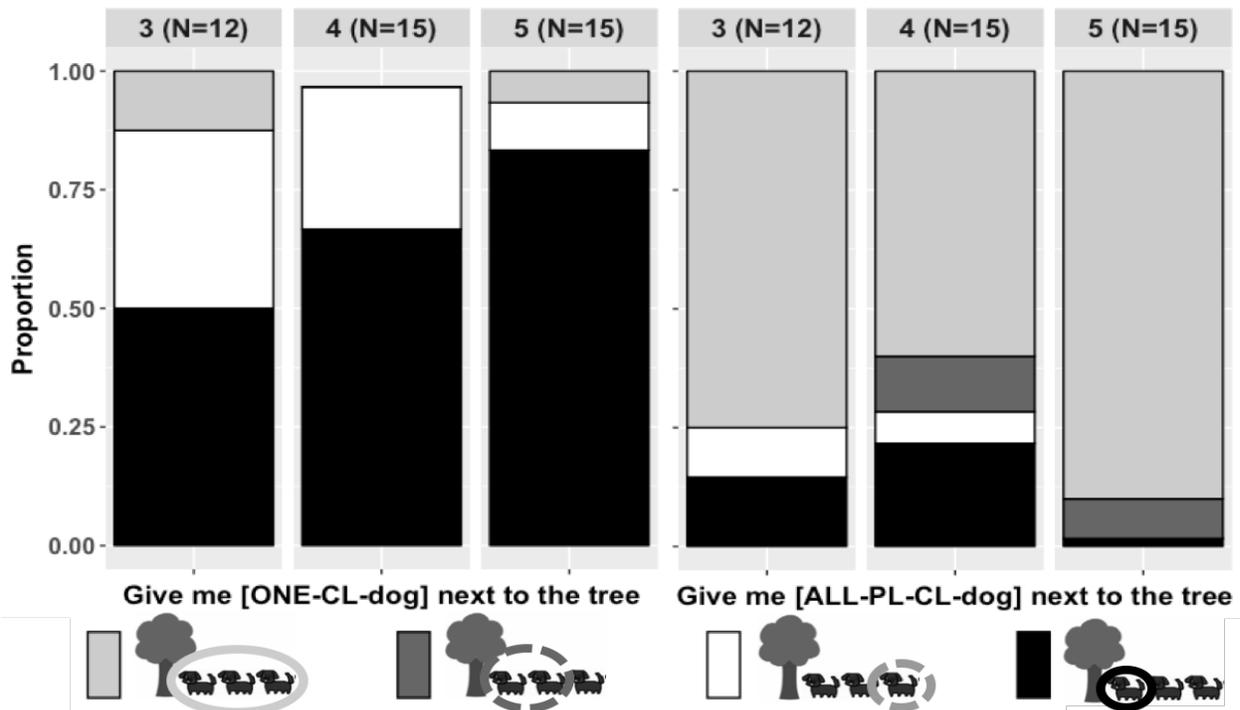
Adults produced 100% target responses in all control conditions. In the singular control condition (ex. *Give me **one CL dog** next to the tree*) where both definite and indefinite singular responses were acceptable, adults once again showed a preference for what we are calling the definite response (closest single dog), choosing it 87.5% of the time. In the experimental conditions, adults produced slightly fewer target responses as compared to Experiment 1. Nevertheless accuracy was still very high, with 82.5% correct in the singular definite condition, 92.5% correct in the plural definite condition with pluralizer *các*, and 90.0% correct in the plural definite condition with pluralizer *những*.

Children’s responses in the control conditions are reported in Table 5. Within the singular control conditions, an ANOVA with age group (3, 4, 5) and experiment (Expt.1, Expt.2) as between-subjects factors reveals no significant difference in the rate of target responses. Nevertheless, the reduction of animals in the display does seem to have diminished children’s default preference for the closest dog; children chose this response 67.86% of the time, down from 85.53% by 3-5-year olds in Experiment 1. This is confirmed with an ANOVA comparing the rate of definite responses (i.e., single closest dog or maximal set of dogs) in the singular control condition provided by 3-, 4- and 5-years olds across the two experiments, which revealed a significant difference between the two experiments ( $F = 3.948, p < 0.05$ ) and a main effect of age group ( $F = 5.806, p < 0.01$ ).

Within the plural control conditions, an ANOVA with age group (3, 4,5) and experiment (Expt.1, Expt.2) as between-subject factors reveals a main effect of age ( $F = 6.677, p < 0.01$ ), but crucially no main effect of experiment on the rate of target responses. There was a significant interaction between age group and experiment ( $F = 7.072, p < 0.001$ ), such that 4-year-olds produced more target responses in Experiment 1 relative to Experiment 2 ( $t(104.98) = -3.10, p < 0.01$ ), but this was the only age group within which there was a significant difference. Figure 5 divides children's responses by age group, collapsing across the two plural control conditions (all-*các*-CL-N and all-*những*-CL-N sequences).

**Table 5.** Percentage (frequency) of children's response types in control conditions

Response type		Condition: <i>one</i> CL N	all <i>các</i> CL N	all <i>những</i> CL N
Singular	Closest	67.86% (57)	11.91% (10)	13.1% (11)
Singular	Non-closest	25.00% (21)	5.96% (5)	4.77% (4)
Plural	Maximal	5.96% (5)	76.20% (64)	73.81% (62)
Plural	Non-maximal	0% (0)	5.96% (5)	8.34% (7)
Plural	Other	1.20% (1)	0% (0)	0% (0)



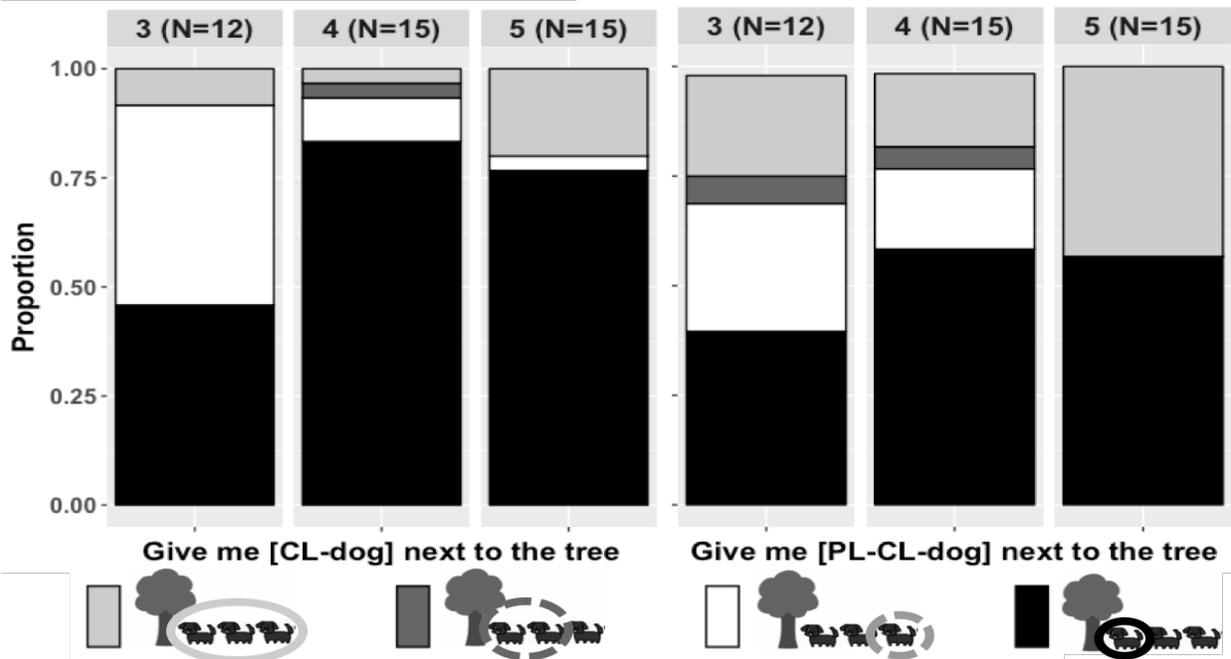
**Figure 5.** Proportion of children's response types, by age group, in singular control condition (left) and plural control conditions (right). Response types include singleton closest to the landmark (black), singleton but not the closest one (white), plural proper subset (medium gray) and plural maximal set (light grey).

Children’s responses in the experimental conditions are shown in Table 6 and the division by age group is reported in Figure 6. In the singular definite condition, children once again produced mostly target responses (70.24% target, compared to 71.05% by 3-5-year olds in Experiment 1). That is, despite the change in children’s default preference in the singular *indefinite* condition, their performance in the singular *definite* condition remains unchanged when the number of animals in the display is reduced. A two-way ANOVA with age group (3, 4, 5) and experiment (Expt.1, Expt.2) as between-subjects factors revealed no difference between the two experiments in the rate of closest to the landmark responses provided in the singular definite condition, although there was a main effect of age ( $F = 3.351, p < 0.05$ ).

In the plural definite conditions with pluralizers *các* and *những*, children produced far fewer target responses compared to children from Experiment 1 (23.81-32.15% target responses in the two plural conditions, as compared 55.26% by 3-5-year olds in Experiment 1). However, the pattern of their responses was qualitatively the same, in that their most common error was to produce a maximal response of the wrong number, rather than a plural, non-maximal response. The *overall* rate of definite responses (i.e., singular closest and plural maximal responses) remained fairly high, especially considering that children in Experiment 2 were younger (between 79.79-80.95% of children’s responses in the plural conditions qualified as either definite singular or definite plural, compared to 90.13% by 3-5-year olds in Experiment 1).

**Table 6.** Percentage (frequency) of children’s response types in experimental conditions

Response type		Condition:	SG-def CL N	PL-def <i>các</i> CL N	PL-def <i>những</i> CL N
Singular	Closest		70.24% (59)	47.62% (40)	57.15% (48)
Singular	Non-closest		17.86% (15)	15.48% (13)	14.29% (12)
Plural	Maximal		10.72% (9)	32.15% (27)	23.81% (20)
Plural	Non-maximal		1.20% (1)	4.77% (4)	2.39% (2)
Plural	Other		0% (0)	0% (0)	2.39% (2)



**Figure 6.** Proportion of children's response types, by age group, in singular (left) and plural experimental conditions (right). Response types include singleton closest to the landmark (black), singleton but not the closest one (white), plural proper subset (medium gray) and plural maximal set (light grey).

As with Experiment 1, we used a 2x3 ANOVA with condition (singular, plural) as a within-subjects factor and age group (3, 4, 5) as a between-subjects factor to test for developmental trends in the overall rate of definite responses, regardless of number. Similarly to the previous experiment, there was a main effect of age ( $F = 23.092, p < 0.001$ ) and no effect of condition ( $F = 0.015, p = 0.90$ ), but this time there was no interaction ( $F = 1.509, p = 0.223$ ). That is, children produced more maximal responses as they grew older, in both conditions, not just the plural condition.

#### 8.4. Discussion

In sum, children in Experiment 2 again appear to interpret singular and plural noun phrases as definite—even when the number of animals is reduced and their default preference patterns change. In the singular control condition, which allows multiple responses, children no longer showed as strong a preference for the animal closest to the tree; yet in the singular definite experimental condition, they continued to choose the singular definite response. This indicates that Vietnamese children seem to be able to restrict the reference of the noun phrase in order to satisfy the uniqueness presupposition of the singular definite noun phrase earlier than their Spanish- and English-acquiring counterparts.

For plural noun phrases like *các con chó đứng kế cái cây* and *những con chó đứng kế cái cây* ('the dogs next to the tree'), results are qualitatively similar to the previous experiment.

Vietnamese children are unlike their cross-linguistic peers in that they frequently commit number errors, despite producing an overwhelming proportion of maximal responses.

## 9. General Discussion

The acquisition of definite noun phrases appears to take a different route in Vietnamese than it does in English and Spanish. With respect to number comprehension, Vietnamese children may distinguish between singular and plural noun phrases, but their ability to associate plural-marked phrases (*các/những*-CL-N) to plural referents lags behind their ability to associate singular phrases (CL-N) to singular sets. And significantly, their ability to associate noun phrases to sets of the correct cardinality lags behind that of English- and Spanish-acquiring children completing the same task.

With respect to definiteness, on the other hand, Vietnamese-acquiring children appear to surpass their American and Mexican peers, if we consider the distinction between the singular indefinite and the singular indefinite responses, as showing that this is not a simple preference. All ages tested showed a strong tendency to associate the definite singular noun phrase with the closest animal (unique singleton set) and the plural noun phrases with the maximal set of animals (unique plural set)—even if that set did not always satisfy the number features of the noun phrase in question.

In sum, Vietnamese children's simply did not produce many definiteness errors. This raises the question of what explains the difference in behavior across languages. That is, why do Vietnamese children succeed with definiteness and commit number errors, while Spanish- and English-acquiring children succeed with number and have problems in the singular condition?

One possibility is that the semantics of the noun phrase locative modifier might differ in Vietnamese. Maybe the phrase *đứng kế cái cây*, which we have translated as *next to the tree*, really means something more like *adjacent to the tree*. However, we reject this explanation based on semantic truth value judgments from native speakers, who answer yes to sentences like (7), even in scenarios where the dog with the bow is not the one adjacent to the tree.

- (7) *Con chó đeo nơ có đứng kế cái cây không?*  
CL dog wear bow yes stand next CL tree no  
'Is the dog with the bow next to the tree?'



Instead, we would like to suggest that what causes the discrepancy in performance between speakers of the different languages is the information that children choose to prioritize. Specifically, we propose that when the task becomes demanding, all groups of children have difficulty simultaneously coordinating information from number and from definiteness, but,

while Vietnamese-acquiring children prioritize definiteness, resulting in number errors, Spanish- and English-acquiring children prioritize number, resulting in definiteness errors. And we suggest that this difference in behavior has to do with the fact that number is obligatorily marked on Spanish and English nouns, while in Vietnamese, overt number marking only exists in the context of definite noun phrases.

If we assume that Spanish- and English-speaking children prioritize number over definiteness it could be that their adult-like behavior in the plural conditions may simply amount to a default interpretation of plural as maximal, since the maximal set is the easiest plural set to access. More studies will determine whether this reinterpretation of the plural definite results in English and Spanish has some merit. In either case, this study highlights the importance of cross-linguistic research in contributing to our understanding of how children acquire semantic concepts. Specifically, what our results suggest is that when languages encode the same semantic primitives in different ways, this has consequences for how children interpret those semantic primitives.

## References

- Barner, David, Dora Thalwitz, Justin Wood, Shu-Ju Yang, and Susan Carey. 2007. On the relation between the acquisition of singular-plural morpho-syntax and the conceptual distinction between one and more than one. *Developmental Science* 10(3). 365-73.
- Brown, Roger. 1973. *A first language: The early stages*. London: Allen & Unwin.
- Bui, Manh Hung. 2000. Về một số đặc trưng ngữ nghĩa - ngữ pháp của “những” và “các” [On the semantic and grammatical properties of *những* and *các*]. *Tạp chí Ngôn ngữ* [Linguistic Journal] 3. 16-26.
- Cao, Xuan Hao. 1998. *Mấy vấn đề ngữ âm-ngữ nghĩa-ngữ pháp trong tiếng Việt* [Some issues on phonetics, semantics, and grammar in Vietnamese]. Hanoi: Education Publisher.
- de Cat, Cécile. 2011. Explaining children’s over-use of definites in partitive contexts. *First Language* 32. 137-50.
- Drozd, Kenneth. 2001. Children’s weak interpretation of universally quantified sentences. *Conceptual development and language acquisition*, ed. by Melissa Bowerman and Stephen Levinson, 340-76. Cambridge: Cambridge University Press.
- Fenson, Larry, Philip S. Dale, Steven J. Reznick, Elizabeth Bates, Donna J. Thal, and Stephen J. Pethick. 1994. Variability in early communicative development. *Monographs of the Society for Research in Child Development* 59(5). 1-185.
- Karmiloff-Smith, Annette. 1979. *A functional approach to child language*. Cambridge: Cambridge University Press.
- Kouider, Sid, Justin Halberda, Justin Wood, and Susan Carey. 2006. Acquisition of English number marking: The singular-plural distinction. *Language Learning and Development* 2. 1-25.
- Lê, Ni-La, and Cristina Schmitt. 2016. Plurality in classifier languages: A view from Vietnamese pluralizers. *Proceedings of the 42nd Annual Meeting of the Berkeley Linguistics Society*, 153-74. Berkeley, CA: University of California.
- Maratsos, Michael. 1976. *The use of definite and indefinite reference in young children*. Cambridge: Cambridge University Press.

- Mervis, Carolyn, and Kathy Johnson. 1991. Acquisition of the plural morpheme: A case study. *Developmental Psychology* 27(2). 222-35.
- Modyanova, Nadya, and Ken Wexler. 2007. Semantic and pragmatic language development: Children know 'that' better. *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)*, ed. by Alyona Belikova, Luisa Meroni, and Mari Umeda, 297-308. Somerville, MA: Cascadilla Press.
- Munn, Alan, Karen Miller, and Cristina Schmitt. 2006. Maximality and plurality in children's interpretation of definites. *Proceedings of BULCD 30*, ed. by David Bamman, Tatiana Magnitskaia, and Colleen Zaller, 377-87. Somerville, MA: Cascadilla Press.
- Munn, Alan, Xiaofei Zhang, and Cristina Schmitt. 2009. Acquisition of Plurality in a Language without Plurality. In: *Merging Features: Computation, Interpretation, and Acquisition*. Brucart, J., Gavarró, A., & Solà, J. (eds.). Oxford: Oxford University Press, 310-27.
- Nguyen, Tai Can. 1975. *Từ loại danh từ trong tiếng Việt hiện đại* [The word class of nouns in modern Vietnamese]. Hanoi: Publisher of Social Sciences.
- Nguyen, Hung Tuong. 2004. *The structures of the Vietnamese noun phrase*. Boston University: PhD Dissertation.
- Nakano, Nao, Hye Sun Park, and Cristina Schmitt. 2010. A comparative study of the acquisition of plural morphemes in Japanese and Korean. Paper presented at *the 11th Tokyo Conference on Psycholinguistics*. Tokyo.
- Schaeffer, Jeannette, and Lisa Matthewson. 2005. Grammar and pragmatics in the acquisition of article systems. *Natural Language & Linguistic Theory* 23(1). 53-101.
- Thompson, Laurence. 1965. *A Vietnamese grammar*. Seattle: University of Washington Press.
- Wexler, Ken. 2011. Cues don't explain learning: maximal trouble in the determiner system. *The Processing and Acquisition of Reference*, ed. by Edward Gibson and Neal J. Pearlmutter, 15-42. Cambridge, MA: The MIT Press.