

Child Comprehension of Number and Definiteness: A View from Classifiers and Pluralizers in Vietnamese

Ni-La Le

Michigan State University

lenila@msu.edu

1 Introduction

Acquisition studies of lexical pluralizers in Japanese, Korean and Mandarin have shown that the acquisition of plural morphemes in these languages is quite protracted in comparison to English (Zhang, 2006; Park, 2010; Nakano et al., 2009; Kim et al., 2012). English-speaking children or other languages that grammaticalize number produce the plural marker on nouns in 90% of obligatory contexts by 24 months and are sensitive to its semantic consequences even in novel words by 36 months of age (Brown, 1973; Mervis & Johnson, 1991; Kouider et al., 2006). On the other hand, children speaking classifier languages acquire these plural morphemes much later: the adult-like interpretations are not mastered until around the age of 6 or 7 (Kim, 2008; Munn et al., 2009; Nakano et al., 2009). This fact may be explained by the fact that pluralizers in these languages are not only optional but also usually portmanteau morphemes which encode additional information beside number, such as classifier (e.g., animacy), definiteness, or uniform versus associative interpretation (Zhang, 2006; Munn et al., 2009; Li et al., 2009; Nakano et al., 2009; Kim, 2011; Kim et al., 2012). As is well-known, portmanteau morphemes are more difficult for children (Karmiloff-Smith, 1979; cf. Peters, 1997).

Vietnamese is a classifier language that also has pluralizers. Noun phrases with just a classifier must be interpreted as definite and singular, and a plural interpretation can be obtained by the addition of one of the pluralizers *các* or *những*. The pluralizer *những* requires that the noun phrase be modified. The pluralizer *các*, on the other hand, does not need modification and is consistently interpreted as definite. Meanwhile, the interpretation of *những* is not as clear. While all studies are unanimous in saying that *các* is definite (Nguyen, T. C., 1975; Diep & Hoang, 1998; Nguyen, H. T., 2004 among others), the literature on Vietnamese is divided regarding *những*. Most authors claim that *những* indicates only a subset of the whole given set (Thompson, 1965; Nguyen, T. C., 1975) and thus it is an indefinite (Nguyen, H. T., 2004), but some other linguists argued the interpretation of *những* vary on a continuum between indefinite and definite (Cao, 1998; Bui, 2000a).

This paper reports a comprehension study of Vietnamese noun phrases that have classifiers (CL): [CL-N(oun)] phrases, which are interpreted as singular and definite, and [CL-N] phrases preceded by two pluralizers *những* and *các* which may have different interpretation in terms of definiteness. The results of our study are going to strengthen and expand current findings on singular/plural distinction as well as the interaction between number and definite features in classifier languages. It also could offer some empirical observations that can contribute to the controversial theoretical issue in the literature of Vietnamese linguistics with respect to the status of *những*. In particular, we show that Vietnamese-speaking children until the age of 5 do not know the plurality and definite properties of these pluralizers. On the other hand, they seem to master

the singularity of bare CL quite early and also successfully pick out the target unique object in the definite singular condition.

2 Linguistics background

2.1 Vietnamese noun phrases and the properties of pluralizers in Vietnamese

The phrase in (1) shows the linear order of a Vietnamese noun phrase with full compositional units, which is All/most-PL/Num/Q-CL-N-Adj-Dem-PP-RC.

- (1) *tất cả những/hai cái chiếc chén sứ Trung Quốc màu nâu hình vuông cỡ
all PL /two Foc CL bowl ceramic China color brown shape square size
trung đó của anh ấy mà tôi mới thấy hôm qua
medium that of elder brother that that I just see yesterday*
'all those (two) brown medium-sized square Chinese ceramic bowl sets of his that I just saw
yesterday'

Most nouns in Vietnamese can combine with Qs, numerals, attributive adjectives (Adj), demonstratives (Dem), relative clauses (RC) or prepositional phrases (PP) but only through the addition of an obligatory CL. The lack of CLs yields unacceptable NPs, as illustrated in (2).

- (2) a. *mỗi /hai *(con) chó*
each/two CL dog
'each dog/ two dogs'
- b. **(cuốn) sách *{hay /này /mà tôi mới mua /của cô ấy}*
CL book good /this /that I just buy /of aunt that
'the good book/ this book/ the book I just bought/ her book'

Much like bare nouns in many other classifier languages such as Thai or Mandarin, bare nouns in Vietnamese can be compatible with a very wide range of interpretations: depending on the predicates and the contexts, a bare noun can have generic, existential, indefinite or definite reading. Bare nouns are also neutral in number, i.e., can be understood as singular or plural, as seen in (3).

- (3) She gave me a pen/pens yesterday
- a. *Bút bị gãy.*
Pen PAR-negative mood break
'The pen/s was/were broken.'
- b. *Tôi làm gãy bút.*
I make break pen
'I broke the pen(s).'

Despite their wide range of interpretations as such, bare nouns in Vietnamese seem to be most fit to generic reading only. In all other contexts, there is always another structure that is preferred, probably because it is more informative. For example, in a singular definite inducing context like (4), CL-phrases are preferred compared to bare nouns like those in (3). Thus in a context in which someone gave me one pen yesterday, it is appropriate to say (4a) or (4b).

- (4) a. [**Cây bút**] bị gãy.
 CL pen PAR-neg mood break
 ‘The pen was broken.’
 b. Tôi làm gãy [**cây bút**].
 I make break CL pen
 ‘I broke the pen.’

This is not uncommon: bare classifier phrases are cross-linguistically preferred over bare nouns in most singular and definite contexts (Simpson et al., 2011: 184-90).

Although *một* ‘a/one’ is also associated with singular interpretation like [CL-N], [*một* ‘a/one’-CL-N] is obligatorily indefinite. The examples in (5-6) illustrate the contrast between two constructions in terms of definiteness: [CL-N] is banned in existential constructions (5b) where English indefinites are often used, while [*một*-CL-N] cannot be anaphoric (6a).

- (5) a. Có [**một con chó**] ngoài sân.
 Have one CL dog outside yard
 ‘There is a black dog in the yard.’
 b. #Có [**con chó**] ngoài sân.
 Have CL dog outside yard
 ‘There is a black dog in the yard.’

- (6) She has a dog and a cat.
 a. #Tôi thích [**một con chó**].
 I like one CL dog
 ‘I like the dog.’
 b. Tôi thích [**con chó**].
 I like CL dog
 ‘I like the dog.’

Other phrases with numerals, except for *một* ‘a/one,’ refer to an indefinite set (7a) or an anaphoric plural set (7b).¹

- (7) a. Cô ấy có [**ba con chó**] và [**hai con mèo**]...
 Aunt that have three CL dog and two CL cat
 ‘She has three dogs and two cats....’
 b. ...Tôi hay đùa với [**ba con chó**].²
 I often joke with three CL dog
 ‘...I often play with the three dogs.’

Although Vietnamese does not have any definite or indefinite determiners, different types of NPs show a clear division of labor among the overt morphology in terms of definite/indefinite and generic interpretations, as summarized in Table 1 below.

¹ Q-CL-N basically can have all interpretations that Num-CL-N can, depending on the particular Q and contexts.

² Num-CL-N can be interpreted as definite (i.e., ‘I often play with her three dogs’) or indefinite (i.e., ‘I often play with three dogs which are not her dogs’). However, if the speaker aims to a definite expression, this construction is less preferred than *các* (plural) or the Q *mấy* (wh-word).

Table 1: Interpretations of some common Vietnamese noun phrase types

Noun phrase type	Generic	Indefinite		Definite	
		Singular	Plural	Singular	Plural
Bare N	yes	yes	yes	yes	yes
CL-N	no	no	no	yes	no
<i>một</i> ‘one’-*(CL)-N	no	yes	no	no	no
Num-*(CL)-N	no	no	yes	no	yes (?)
<i>các</i> -*(CL)-N	no	no	no	no	yes
<i>những</i> -*(CL)-N-*(MOD)	yes	no	yes	no	yes (?)

As seen in Table 1, *những* and *các* combine with CLs to derive plural noun phrases. Both *những* and *các* require a CL in the noun phrase (8), cannot co-occur with numerals and most Qs (9a), except for *tất cả* ‘all’ and *hầu hết* ‘most’, for which either of the pluralizers has to be present (9b), and are definite-like, similar to pluralizers in other classifier languages (10).

(8) *những/các* *(*con*) *chó mực*

PL CL dog ink

‘(the) black dogs’

(9) a. **mấy/ba những/các chiếc thuyền gỗ*

some/three PL CL boat wood

‘some/three wooden boats’

b. *tất cả/hầu hết những/các chiếc thuyền gỗ*

all /most PL CL boat wood

‘all/most the wooden boats’

(10) a. Situation 1: X asks for blue pens in a store.

Answer: **Tiệm có /không bán các/những cây bút xanh.*

Store yes /no sell PL CL pen blue

‘The store does/doesn’t carry blue pens.’

b. Situation 2: She just brought five dogs home (some are black, some are white).

She says: ✓*Các/những con chó mực đang ngủ.*

PL CL dog ink Prog sleep

‘The black dogs are sleeping.’

In terms of differences, *những* has a more complicated nature than *các*. First of all, *những* has to be licensed by some sort of restriction on the noun phrase (11a), while *các* does not need one (11b).

(11) a. *những cuốn sách *{hay /này /mà tôi mới mua /của cô ấy}*

PL CL book good /this /that I just buy /of aunt that

‘the good books/ these books/ the books I just bought/ her books’

b. *các con chó*

PL CL dog

‘the dogs’

Second, the behavior of *các* is consistent with it being definite whereas the interpretations of *những* can vary. The morpheme *các* is always compatible with a definite interpretation of the

noun phrase, which, according to Heim (1991), satisfies familiarity and maximality. As it denotes familiarity, *các* cannot appear in existential sentences like (10b) and requires discourse anaphora, i.e., it is discourse dependent (e.g., (12a) cannot be said out of blue).

(12) *Ngày xưa ngày xưa có *các chàng hoàng tử rất thích ngựa.*

Once upon a time have PL CL prince very like horse

‘Once upon a time, there were princes that really liked horses.’

(13) *Cô ấy nuôi năm con chó và một con mèo.*

Aunt that raise five CL dog and one CL cat

‘She has five dogs and one cat.’

a. *Các con chó rất trung thành.*

PL CL dog very loyal

‘The dogs are very loyal.’

b. *#Các con chó màu đen.*

PL CL dog color black

Intended: ‘Some (two, three, or four) of the dogs are black.’

OK: ‘The dogs are black.’

An indefinite reading is impossible when *các* precedes [CL-N]: it cannot introduce a set of entities in the discourse, which makes (12) ungrammatical. *Các* also imposes a maximality restriction. It cannot refer to a subset of the entities previously mentioned, as shown in (13b), but must pick the whole set in the discourse (13a) and therefore does not behave as a demonstrative, explaining why (14a) is unacceptable.

(14) a. **Các con chó đang ngủ, các con chó đang giỡn.*

PL CL dog Prog sleep PL CL dog Prog play

‘*The dogs are sleeping, the dogs are playing.’

b. *Các con chó đó đang ngủ, các con chó đó đang giỡn.*

PL CL dog that Prog sleep PL CL dog that Prog play

‘Those dogs are sleeping, those dogs are playing.’

In addition to being discourse-anaphoric, [*các*-CL-N] can also be used in all other primary situations licensing the use of definite determiners in English, French and many other languages (Simpson et al., 2011:172): it can refer to entities which have not been mentioned explicitly but are visible and uniquely identifiable for both speaker and hearer (15a) or to culturally unique, familiar entities (15b), and *các* noun phrases are felicitous in bridging contexts like (15c, d).

(15) a. *Đưa tôi các cây búa.*

Give I PL CL hammer

‘Pass me the hammers.’

b. *Chiến tranh giữa các vì sao*

War between PL CL star

‘the war between the stars (Star wars)’

c. ‘She just bought a new tea set yesterday...’

...*Các cái tách hơi bé.*

PL CL cup rather small

‘The teacups are pretty small.’

- d. *Các quyển sách toán ở chỗ nào?*
 PL CL book math at place which
 ‘Where are the math books?’ (asking a librarian)

As for *những*, the data is more complicated. As mentioned above, most authors in the literature agree that *những* indicates only a subset of the whole given set (Thompson, 1965; Nguyen, T. C., 1975; Nguyen, H. T., 2004), which seems to suggest *những* is somewhat similar to English Q *some (of)*. However, as shown in (10b) above, *những* is compatible with anaphoric readings. The examples in (16) show that *những* can also be definite in the sense that it can refer to the whole set of contextually unique items.

- (16) a. *Những/các bức tranh trên tường vừa được tháo xuống.*
 PL CL painting on wall just PASS-positive remove down
 ‘The paintings on the wall were taken down.’
 b. *Cô ấy tháo những/các bức tranh trên tường xuống.*
 Aunt that remove PL CL painting on wall down
 ‘She took the paintings on the wall down.’

Like *các*, what *những* picks out in sentences such as (16) has to be the entire set, i.e., all the pictures on the wall, not some of them. Another evidence for definite-like nature of *những* is that *những* can be interchangeable with *các* in all contexts in (15), as long as some kind of modification is provided, as illustrated in (17) below.

- (17) a. *Đưa tôi những cây búa màu đỏ.*
 Give I PL CL hammer color red
 ‘Pass me the red hammers.’
 b. *chiến tranh giữa những vì sao đêm*
 war between PL CL star night
 ‘the war between the night stars’
 c. ‘She’s just bought two tea sets. One is made in Japan, the other one is made in Vietnam...’
...Những cái tách Nhật hơi bé.
 PL CL cup Japan rather small
 ‘The teacups in the Japanese set are pretty small.’
 d. *Những quyển sách toán ở chỗ nào?*
 PL Cl book math at place which
 ‘Where are the math books?’ (asking a librarian)

However, unlike *các* which can combine with kinship terms used as singular personal pronouns to create Plural pronouns, *những* cannot, as shown in (18).

- (18) a. *Cô có thích chó mực không?*
 Aunt yes like dog ink no
 ‘Do you (SG) like black dogs?’
 b. *[Các/*những cô] có thích chó mực không?*
 PL aunt yes like dog ink no
 ‘Do you (PL) like black dogs?’

Meanwhile, *những* is compatible with many typical indefinite constructions. It can appear with *wh*-phrases while *các* cannot:

- (19) a. *Cô ấy nói những/*các gì?*
 Aunt that say PL what
 ‘What did she say?’
 b. *Cô ấy đi những/*các đâu?*
 Aunt that go PL where
 ‘Where did she go?’

Những can appear in ‘there are...’ expressions but *các* would yield ungrammatical sentences:

- (20) *Ngày xưa ngày xưa có những/*các chàng hoàng tử rất thích ngựa.*
 Once upon a time have PL CL prince very like horse
 ‘Once upon a time, there were princes that really liked horses.’

The pluralizer *những* can also allow generic readings with individual-level predicates, like those in (21), while [*các*-CL-N] does not have this interpretation (22), unless there is a modifier of some sort on the noun phrase, as seen in (23).

- (21) a. *Những con chó mực thường trung thành.*
 PL CL dog ink usually loyal
 ‘Black dogs are usually loyal.’
 b. *Tôi thích những con chó mực.*
 I like PL CL dog ink
 ‘I like black dogs.’
 (22) a. *#Các con chó thường trung thành.*
 PL CL dog usually loyal
 ‘Dogs are usually loyal.’
 b. *#Tôi thích các con chó.*
 I like PL CL dog
 ‘I like dogs.’
 (23) a. *Các con chó mực thường trung thành.*
 PL CL dog ink usually loyal
 ‘(The) black dogs are usually loyal.’
 b. *Tôi thích các con chó mực.*
 I like PL CL dog ink
 ‘I like (the) black dogs.’

These examples in (23) show that restriction on the noun phrase not only licenses the presence of *những* in a noun phrase, but also allows the generic interpretation of *các*. Therefore, if there is restriction on the noun phrase, *những* and *các* allow both generic and non-generic definite readings, especially when they combine with ‘most’ (24) or intentional verbs (25).

- (24) a. *Hầu hết những/các tòa nhà cao tầng có thang máy*
 Most PL CL house tall floor have ladder machine
 ‘Most **những/các** tall buildings have elevators.’
 b. *Reading 1: Most of the tall buildings (e.g., on campus) have elevators.*
Reading 2: As for tall buildings, most have elevators.
- (25) a. *Tôi muốn mua những/các con chó mực.*
 I want buy PL CL dog ink
 ‘I want to buy **những/các** black dogs.’
 b. *Reading 1: There are specific black dogs I want to buy.*
Reading 2: I want to buy whatever dogs that are black.

The summary of the syntactic and semantic properties of *các* and *những* is in Table 2.

Table 2: *The properties of những and các*

	NHỮNG	CÁC
Co-occur with ‘all/most’	yes	yes
Co-occur with other quantifiers	no	no
Co-occur with numerals	no	no
Require classifiers	yes	yes
Require restriction on the NP	yes	no
Require D-linked	yes	yes
Co-occur with <i>wh</i> -elements	yes	no
Appear in ‘ <i>there are...</i> ’ construction	yes	no
Allow generic reading	yes	no*
Co-occur with kinship terms	no	yes

* *in general, except in the presence of modification.*

These PLs are obligatory not only for ‘all’ and ‘most’ but also for the plural interpretation of CL-phrases. In Vietnamese, [CL-N] and [CL-N-Dem] are consistently interpreted as singular and can be Pluralized by adding *những/các*, as shown in (26) and (27).

- (26) a. *con chó (mực)* b. *các con chó* c. *những con chó mực*
 CL dog ink PL CL dog PL CL dog ink
 ‘the (black) dog’ ‘the dogs’ ‘the black dogs’
- (27) a. *con chó này* b. *các con chó này* c. *những con chó này*
 CL dog this PL CL dog this PL CL dog this
 ‘this dog’ ‘these dogs’ ‘these dogs’

As we have seen, everytime one of these pluralizers is present we have a plural interpretation. But what is the interpretation of the plural? Nomoto (2013) claims that pluralizers in classifier languages are a ‘genuine plural,’ i.e. always have more-than-one interpretation (p.102), unlike English plurals, which can allow one or more-than-one interpretation, in some contexts (Sauerland et al., 2005). In particular, English plural noun phrases can appear in yes/no questions and existential sentences with a one or more-than-one interpretation, as illustrated in (28) and (29), respectively.

- (28) Q: ‘Do you have daughters?’
 A: ‘Yes, I have one daughter.’

(29) *I think there were dogs in our yard last night. I heard some barking.*

The answer to the question in (28) shows that the interpretation is not obligatorily a more-than-one interpretation. Similarly, one can say (29) even when there is only one dog in the yard. Additionally, English plurals can behave as dependent plurals, as illustrated (30). Dependent Plurals are interpreted as distributed singular indefinites.

(30) a. *Unicycles have wheels.*

b. *Unicycles have a wheel.* (Minor, 2014: Ex. 1 & 2)

Another ability of English plural as a dependent plural is that it allows a long distance dependency. For example, if Bob wants Bill to marry Ann, who happens to be a famous linguist, and Kate wants Bill to marry Jane, who is also a well-known linguist, a plural form of *linguist* can be used in English (31a) but prohibited in Vietnamese (31b).

(31) a. *Bob and Kate want Bill to marry famous linguists.*

b. #*Bob và Kate muốn Bill cưới những nhà ngôn ngữ học nổi tiếng.*

Bob and Kate want Bill marry PL CL language study famous

In Vietnamese, the pluralized noun phrases are banned in questions, *there* constructions with a one and more-than-one interpretation, and dependent plurals. It seems to be hard for *những/các* to have a weak interpretation: they are discourse linked and are interpreted as definite. To express a one or more-than-one interpretation, a bare noun is preferred, as seen in (32-33).

(32) a. *Cô có con gái không?*

Aunt have child female no

‘Do you have daughters?’

b. **Cô có những/các đứa con gái không?*

Aunt have PL CL child female no

‘Do you have daughters?’

(33) a. *Có chó (mực) ngoài sân.*

Have dog ink outside yard

‘There is/are (a) black dog(s) in the yard.’

b. #*Có những/các con chó mực ngoài sân.*

Have PL CL dog ink outside yard

‘There is a black dog in the yard.’

(34) a. *Nhân mã có một cái sừng.*

Human horse have one CL horn

‘Unicorns have a horn.’

b. **Nhân mã có những/các cái sừng to*

Human horse have PL CL horn big

‘Unicorns have big horns.’

Although the data is inconclusive with respect to the interpretation of the plural morpheme itself, it is impossible to obtain a one or more-than-one interpretation of the pluralizer and therefore we argue that it has a more-than-one interpretation.

So far, our observations have shown *những* and *các*, on one hand, has inherent plural interpretations, and on the other hand, do not behave exactly as plural markers in traditional sense. This may come from the fact that they, like the plural morphemes in other classifier languages, are also portmanteau morphemes³ (morphemes that encode more than one piece of information, cf. Peters 1997). In particular, besides number features, they contain in/definite information.

2.2 The acquisition of definites

Definites, in general, have a uniqueness presupposition: a definite noun phrase refers to a unique object (singular) or a maximal unique set of entities (plural) in the discourse (Kadmon, 1990; Heim, 1991; Roberts, 2003 among others). From the acquisition point of view then, the correct interpretation of a definite derives from two factors: (i) knowing that definites require a relevant maximal set (which is a single unique item in a singular condition) and (ii) being able to pick out the right domain in which maximality is defined.

Many studies have shown evidence that children do not always use definites like adults. They often incorrectly use the definite to refer to only one of identical objects in the context, e.g., *Give me the ball!* instead of *Give me a ball!* (Maratsos, 1976; Karmiloff-Smith, 1979). There are three different hypotheses that have used to explain the misuse of definites of children as such. Both Maratsos (1976) and Karmiloff-Smith (1979) suggest a child uses the definite for any referent that is under his own focus of attention ('egocentrism,' Maratsos, 1976:63; Karmiloff-Smith, 1979:72).⁴ Alternatively, observing that children still overuse the definite even there is no element in focus, Wexler (2011) argues that their errors are due to the lack of the Maximality presupposition of the definite ('no Maximality Hypothesis,' p. 25). Other work such as Drozd (2001), Miller & Schmitt (2004), Munn et al. (2006) has proposed another possibility, i.e., children have difficulties in finding a right domain for determiners ('Domain Restrictions'). In particular, the fact that English children in the experiment by Munn et al. (2006) failed in singular conditions (e.g., *Give me the frog next to the pond*) but not in plural conditions (e.g., *Give me the frogs next to the pond*) is compatible with the idea that (i) children know the definite refer to a maximal set (at least in plural conditions) and (ii) children have problems with domain restrictions on the definite (at least with the implicit ones) (p. 385-386).

2.3 The acquisition of plurality

The distinction between singular and plural sets can be expressed in different ways in different languages, such as lexical quantifiers or morphology on the nouns, verbs, adjectives, or determiners. In terms of interpretation, plural markers can be interpreted as 'one and more-than-one' in some languages like English (e.g., *I didn't eat cookies*) or always as 'more-than-one' in some other languages such as Korean or Japanese (Kim, 2008; Nomoto, 2013; Liter et al., 2013).

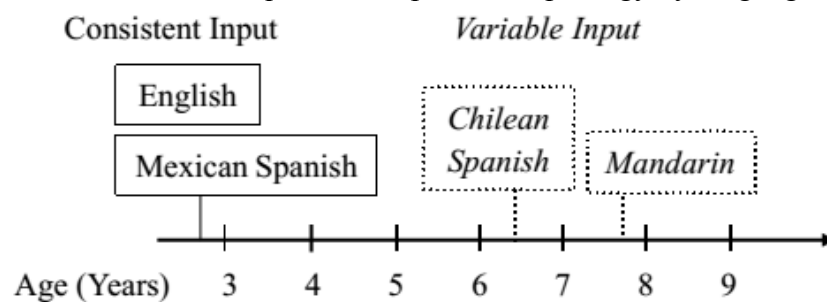
³ While Korean *-tul*, like *những* and *các*, does not have any restrictions regarding the degree of animacy of the noun (Kim, 2008:163), Mandarin *-men* and Japanese *-tachi* can be used mostly with human nouns (Fehri & Vinet, 2004:12; Li, 2009:94; Nomoto, 2013:103-105). When used with non-human nouns, *-men* and *-tachi* either personify the entities or express the speaker's affection toward them (Nomoto, 2013:103-105). All of them also act as specific or definite.

⁴ However, they assume different reasons for this 'egocentric' interpretation in children. For Maratsos, although definites need the specificity of reference to both speakers and hearers, children do not have the latter. According to Karmiloff-Smith, however, children seem to have a more deictic version of the definite than adults.

In terms of morphology, there is a tendency that Plural is the element morphologically marked in the plural/singular pair. If a language chooses to morphologically mark only one side of this distinction, plural will be the one bearing the overt morphology (Corbett, 2000).

Many studies of English plurality have suggested that (i) English-speaking children master the conceptual distinction between *one* and *more-than-one* between 20 months and 24 months of age (Fenson et al., 1994; Barner et al., 2007), (ii) children between 24 and 36 months old can produce the plural marker in correct contexts and use it to learn novel words (Brown, 1973; Mervis and Johnson, 1991; Kouider et al., 2006), and (iii) they are more sensitive with plurality marked lexically (e.g., ‘*Look, there ARE SOME blickets*’) than plurality only marked by bound morphemes (e.g., *Look at the blicketS*) (Kouider et al., 2006; Wood et al., 2009). Meanwhile, children speaking classifier languages such as Mandarin, Japanese, or Korean do not have the adult-like interpretations of the plural markings until much later, i.e., between 6 and 8 years old (Kim, 2008; Nakano et al., 2009; Munn et al., 2009; Li et al., 2009). As we mentioned above, this may be due to the fact that there is more information than just number itself encoded in plural morphemes in these languages (cf. footnote 3, p. 14) but it could also be that plural in these languages is not as frequent and is optional (Miller, 2007; Miller & Schmitt, 2009). For example, as shown by the work by Miller (2007) and Miller & Schmitt (2009), children speaking Mexican Spanish which marks plural morphology systematically and obligatorily are able to use plural markers in comprehension tasks by age 3. Children speaking Chilean Spanish whose plural morpheme is sometimes not overtly present in the input under the lenition process with all syllable-final [-s], on the other hand, take longer to master plural morphology. Nakano et al. (2009) summarizes the timeline of acquisition of plural morphology in languages with consistent and variable input in Figure 1, which shows that children have difficulties to learn variable plural morphology and it is even harder for them to acquire pluralizers that are not only optional but have extra constraints/information besides number.

Figure 1: Timeline of the acquisition of plural morphology by language



(Nakano et al., 2009:5)

Similarly, in Vietnamese, although the pluralizers are unambiguously plural, they are not obligatorily marked in every plural noun phrase because bare nouns can be interpreted as plural (or singular). This leads to two difficulties for children from the acquisition point of view. First, the input in classifier languages is likely to provide the child with information for lack of plural morphology, while evidence for a pluralizer appears in a much smaller proportion of the input. Second, many studies have shown that variability in the input that causes ambiguity delays

acquisition (Yang, 2002; Miller, 2007; among others). In acquiring the plurality of *những/các*, Vietnamese-speaking children seem to deal with three types of variability at once: (i) both a bare form and a pluralized form are associated to a plural interpretation, as mentioned above, (ii) there are two different lexical items with the same plural feature but different restrictions (with or without modification), and (iii) the interpretation of *những* can vary since it can appear in both definite contexts and indefinite expressions like ‘*there are*’ or *wh*-phrases.

In summary, we may predict that Vietnamese children should master these morphemes later than their English peers due to their optionality and the input variability. However it may be that the particularities of Vietnamese may lead Vietnamese speaking children to learn a bit faster than Mandarin, Japanese, or Korean children.⁵ In terms of the interaction between number and definiteness, Vietnamese-speaking children may have a command of the plurality of *những* and *các* quite early (as they are lexical pluralizers) and will acquire number before in/definite features, since definiteness is also protracted in all other languages. In addition, we should expect some differences in the acquisition path of *những* compared to that of *các* for two reasons. As *những* has a higher frequency in the input than *các* (Bui, 2000a: 11), children might start learning *những* earlier. On the other hand, the interpretations of *những* which are more complicated and less consistent than *các* may delay the adult-like version of this pluralizer for children.

3 Research questions

The experiment tested child comprehension of number and definiteness of Vietnamese noun phrases with or without pluralizers. In particular, I want to investigate the path Vietnamese-speaking children acquire the three noun phrase types of interest, i.e., how they develop to the point they can associate [CL-N] with a singular, definite reading, *các*-CL-N with a plural, definite reading, and interpret plural noun phrases with the pluralizer *những* like adults. In order to find out the answer to that general question, the experiment tries to address the three questions in (35).

(35) *Research questions:*

- a. **Question 1:** At which age do children know the number information (singular vs. plural) encoded in these constructions?
- b. **Question 2:** How do children interpret these noun phrase types in terms of definiteness?
- c. **Question 3:** Do *các* and *những* behave differently in terms of definiteness, as claimed by the literature?

In this experiment, I replicate Munn et al. (2006) in testing children’s knowledge of number and definiteness in a context that allows the same scenario. However, their experiments investigated definites in languages that definite and number features are encoded somewhat separately: English definite determiner *the* does not contain any number information and in the Spanish experiment, the singular/plural distinction is encoded in both the definite determiners and

⁵ Vietnamese pluralizers do have less discourse restrictions than other languages: they do not allow associative readings (as in Japanese and Korean) and it is not a portmanteau morpheme with a CL, i.e., do not have any requirement with respect to animacy (as in Mandarin, Japanese and Korean). They also have less optionality: they are required to pluralize a noun phrase if there is a CL (such as [CL-N] or [CL-N-Dem]). They are also obligatory in the presence of ‘all’ or ‘most.’

the nouns. On the contrary, these two pieces of information are tangled together in the same form in Vietnamese, so they are not as independent in terms of acquisition. The empirical findings will not only show the developmental course of mastery of these kinds of information encoded in Vietnamese pluralizers and bare CL-phrases, but also may be able to shed some lights on traditional description of the morpheme of *những*.

4 Hypotheses and predictions

In general, we hypothesize that Vietnamese children acquire both plurality and definiteness later than their peers of other inflectional languages in which these two features are separately realized as different morphemes and these morphemes appear systematically and consistently, unlike Vietnamese where number is not grammaticalized and can be expressed via different means (e.g., bare nouns, other Qs); more importantly, it is encoded together with definiteness in the same form. This means it should be difficult for children to master the pluralizers *những/các*.

In terms of acquisition order, we also hypothesize that Vietnamese-speaking children may master the number feature earlier (as *những* and *các* are lexical pluralizers), but do not have adult-like interpretations of bare CL-phrases and the two pluralizers in terms of definiteness until much later, because definiteness is proven to be difficult for children in all languages even when it manifests independently from number like in English. Therefore, we predict children to be able to distinguish singular from plural in this experiment (one vs. more-than-one) even if they give us the wrong target set (not the closest one/s).

As mentioned above, the non adult-like interpretations of definites can be accounted for by different hypotheses, including the Maximality Hypothesis (36a) and the Domain Restriction Hypothesis (36b).

(36) *Alternative Hypotheses:*

- a. H_{1a}: Children do not have the Maximality presupposition (Wexler 2011)
- b. H_{1b}: Children have difficulties with domain restrictions (Drozd, 2001; Miller & Schmitt, 2004; Munn et al., 2006; among others)

If children do not have the Maximality presupposition, they should fail to give maximal answers for both singular and plural definite noun phrases. In contrast, if children have difficulties with domain restrictions, then we expect to see them fail only in singular definite condition while never give a non-maximal plural response in plural definite condition⁶.

Children's performance in the experiment can then be searched for (i) the acquisition order between number and definiteness and for (ii) these predictions to see which hypothesis is borne out. In addition, we should be able to observe whether *những* tends to pick out a whole set or a subset.

5 Methods

This is an act-out task, replicating Munn et al.'s (2006) design in testing English and Mexican children.

⁶ They may give singular answers if the plurality is not acquired yet.

5.1 Experimental set up

The experiment used 4 types of animal⁷ toys, which were dogs, cats, roosters, and penguins (6 identical toys per type). The scene included a toy house and a tree apart from each other. Six animals of the same type were divided in two groups: three of them lined up next to the tree while the other three lined up next to the house (Figure 2). Participants' task was to pick out a set of animal toys next to either the tree or the house corresponding to what they heard. In that scenario, a definite singular expression (i.e., bare CL-phrases) refers to the one closest to the tree/house (e.g., cat number 6 in Figure 2) while a definite plural noun phrase (i.e., *các*-CL-N) picks out all three toys close to the tree/house (e.g., cat number 4, 5, and 6). If *những* actually indicates a subset of entities like how it has been described in many theoretical work, participants can pick two out of three animal toys (e.g., cat number 5 and 6).

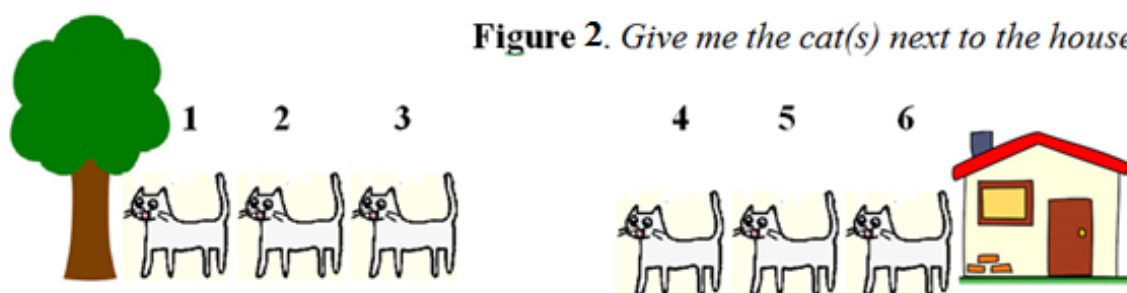


Figure 2. Give me the cat(s) next to the house

5.2 Subjects

We tested 56 Vietnamese-speaking children ages 2;7 to 5;7 from two kindergartens in Ho Chi Minh City (Vietnam)⁸ and 22 adults in the same city during the summer of 2015. Child subjects received a small gift (a candy, a pencil, or a small animal toy) for participating; adult subjects did not receive any compensation. Data from two children who refused to complete the task and other 12 children who were not correct for at least one item of each control construction is eliminated. Data from two adults who did not get all control items correct is also not considered. The results presented in this paper are from 42 native Vietnamese children (M: 4;5, range: 2;7-5;6, 22 males) and 20 adults (M: 23;11, range: 16;8-41;6; 4 males).

5.3 Materials

This experiment used six forms of Vietnamese noun phrases, listed in Table 3. Our inventory consisted of 48 sentences (6 NP types⁹ x 4 animal types x 2 displays (tree or house)). From there, 4 versions with 12 sentences for each were designed. In each version, there were three experimental conditions, as illustrated in (37): definite singular (2 items), definite plural (2 items), and in/definite plural (2 items).

- (37) a. *Đưa cho cô con chó đứng gần cái nhà* CL (SG, Def)
Give for aunt CL dog stand near CL house
'Give me the dog next to the house.'

⁷ Animal nouns require the same CL *con*, which is a general CL for animate entities. We want to make sure children know the CL and thus it is not a confounding factor influencing their performance.

⁸ One autistic child at the age of 6;10 who was also tested for courtesy is not included.

⁹ All can combine with either *những* or *các*.

- b. *Đưa cho cô các con chó đứng gần cái nhà* *các* (PL, Def)
 Give for aunt PL CL dog stand near CL house
 ‘Give me the dogs next to the house.’
- c. *Đưa cho cô những con chó đứng gần cái nhà* *những* (PL, In/Def)
 Give for aunt PL CL dog stand near CL house
 ‘Give me (some of) the dogs next to the house.’

Table 3: *Noun phrase types used in the experiment*

Condition	Noun phrase type	Target referent
singular (SG), definite (Def)	CL-N	SG-closest
plural (PL), Def	<i>các</i> -CL-N	PL-exhaustive
PL, in/definite (In/Def)	<i>những</i> -CL-N-modifier	all three items ¹⁰
SG, indefinite (Indef)	<i>một</i> ‘one’-CL-N	any one item among the three
All (1)	<i>tất cả</i> - <i>các</i> -CL-N	all three items
All (2)	<i>tất cả</i> - <i>những</i> -CL-N	all three items

Sentences with [*một* ‘one’-CL-N], [*tất cả* ‘all’-*những*-CL-N], and [*tất cả* ‘all’-*các*-CL-N], as illustrated in (38), were used as control conditions (2 sentences each, so 6 in total). The control items were always ordered after all test items to avoid any potential contrast strategies, especially between pluralizers and *all*.

- (38) a. *Đưa cho cô một con chó đứng gần cái nhà* **one** (SG, Indef)
 Give for aunt one CL dog stand near CL house
 ‘Give me a dog next to the house.’
- b. *Đưa cho cô tất cả các con chó đứng gần cái nhà* All *các*
 Give for aunt all PL CL dog stand near CL house
 ‘Give me all the dogs next to the house.’
- c. *Đưa cho cô tất cả những con chó đứng gần cái nhà* All *những*
 Give for aunt all PL CL dog stand near CL house
 ‘Give me all the dogs next to the house.’

Những and *các* were tested in separate blocks in 4 versions and the item order within each block was randomized. Some children received *những* block before *các* (2 versions) while some were tested on *các* before *những* (2 versions). We also included two versions in which there was a mixture between *những* and *các* items. One filler item followed after every four tested items. These filler items were yes-no questions taken from an experiment on Vietnamese pronouns by Forsythe et al. (2015).

5.4 Procedure

The experiment had two sessions: Pretest and Test and the entire procedure, including the Pretest (familiarization phase), lasted approximately 30 minutes in total. Children were tested individually and videotaped¹¹ (as parents consented). During the Pretest, a child was asked to

¹⁰ If *những* is indefinite, subjects should pick out 2 items. However, as we discussed in the background in section 2.1, despite what the literature has claimed before, it seems to appear mostly in definite contexts (p. 8-9). Therefore, definite responses are chosen as target/correct responses.

¹¹ Videotapes were sometimes used to verify coding of experimental data.

name the toys, the landmarks, and answer the question ‘*What is next to what?*’ in order to ensure that s/he recognized the displays and understood what *next to* means. During the Test, s/he was asked to give the experimenter or the puppet what she wanted, e.g., either ‘the cat next to the house/tree’ or ‘the cats next to the house/tree’ (Figure 2 above).

Adults participated in two experiments in a single session¹² without the training phase. They were tested in a large group with a paper and pencil version of the task. The adults’ results would provide an estimate goal of child language development.

5.5 Coding of responses

In the response sheet, the items in the displayed were numbered from 1 to 6 from left to right. The experimenter recoded child subjects’ responses on the sheet by noting down the item number(s) they gave her. The responses were then transferred to a spreadsheet for coding. Each response was coded into one of the five categories: ‘SG closest’ (item 1 or 6), ‘SG non-closest’ (item 2, 3, 4, or 5), ‘PL exhaustive’ (all three items as 1+2+3 or 4+5+6), ‘PL non-exhaustive’ (two items 1+2, 2+3, 1+3, 4+5, 5+6, or 4+6), and ‘others’ (e.g., 3+4, 1+2+3+4+5+6). All responses were also marked as target in three different dimensions: number, definite, and both, listed in Table 4 below.

Table 4: Definition of target responses

Condition (2 items/condition/subject)	Target referent		
	Number	Definite	Both
CL-N	SG	1/6	SG closest
<i>các</i>	PL	1+2+3/4+5+6	PL exhaustive
<i>những</i>	PL	1+2+3/4+5+6	PL exhaustive
one	SG	1/6	SG closest/SG non-closest
all <i>các</i>	PL	1+2+3/4+5+6	PL exhaustive
all <i>những</i>	PL	1+2+3/4+5+6	PL exhaustive

6 Results

Responses in each condition are shown in Table 5 for adults and Table 6 for children, illustrated by the graph in Figure 3. Overall, adults were doing very well, as expected: 100% correct rate for control items and 88.33% on experimental items. It is clear from adults’ performance that *các* and *những* were treated as definite plurals: adults chose a maximal plural set for 92.5% of *các* items and 90% of *những* items. The tendency of adults to choose the closest item was shown not only in the CL condition (singular, definite) (82.5%) but also in the ‘one’ condition (singular, indefinite) (87.5%), and seemed to be preserved in children as well (70.24% and 67.86% for CL and ‘one’, respectively). Children were also adult-like in control conditions (80.95% correct): the percentage correct responses were 92.86% for ‘one’, 76.2% for ‘all *các*’, and 73.81% for ‘all *những*’. Among three experimental conditions, children understood the singular definite items best (CL-N) with the rate of 70.24% of responses. However, children did not do well with both pluralizers. In particular, they overwhelmingly favored a singular interpretation for all

¹² The other experiment also tested definite properties of *những* and *các* using a Truth Value Judgement Task design. Its results will not be presented in this paper.

experimental items, including noun phrases with pluralizers (63.1% and 71.44% for *các* and *những*, respectively). More specifically, there was a bias towards singular **definite** responses across all types of tested noun phrases, except for *all* whose maximality seemed to be already acquired by child subjects. In general, adults' correct responses rate is twice as high as that of children on experimental items (88.33% compared to 42.06%).

Table 5: Adult responses (N = 20): proportion and number of responses (target responses in shaded cells)

Condition (40 responses /condition)		Experimental items			Control items		
		CL	các	những	one	all các	all những
Singular	Closest	82.50% (33)	5% (2)	5% (2)	87.50% (35)	0% (0)	0% (0)
	Non-closest	10% (4)	0% (0)	0% (0)	12.50% (5)	0% (0)	0% (0)
Plural	Exhaustive	7.50% (3)	92.50% (37)	90% (36)	0% (0)	100% (40)	100% (40)
	Non-exhaustive	0% (0)	2.50% (1)	5% (2)	0% (0)	0% (0)	0% (0)
Others		0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)
Total correct responses		94.17%	88.33%		100%		

Table 6: Child responses (N = 42): proportion and number of responses

Condition (84 responses /condition)		Experimental items			Control items		
		CL	các	những	one	all các	all những
Singular	Closest	70.24 (59)	47.62 (40)	57.15 (48)	67.86 (57)	11.91 (10)	13.1 (11)
	Non-closest	17.86 (15)	15.48 (13)	14.29 (12)	25.00 (21)	5.96 (5)	4.77 (4)
Plural	Exhaustive	10.72 (9)	32.15 (27)	23.81 (20)	5.96 (5)	76.20 (64)	73.81 (62)
	Non-exhaustive	1.20 (1)	4.77 (4)	2.39 (2)	0 (0)	5.96 (5)	8.34 (7)
Others		0 (0)	0 (0)	2.39 (2)	1.20 (1)	0 (0)	0 (0)
Total correct responses		61.51%	42.06%		80.95%		

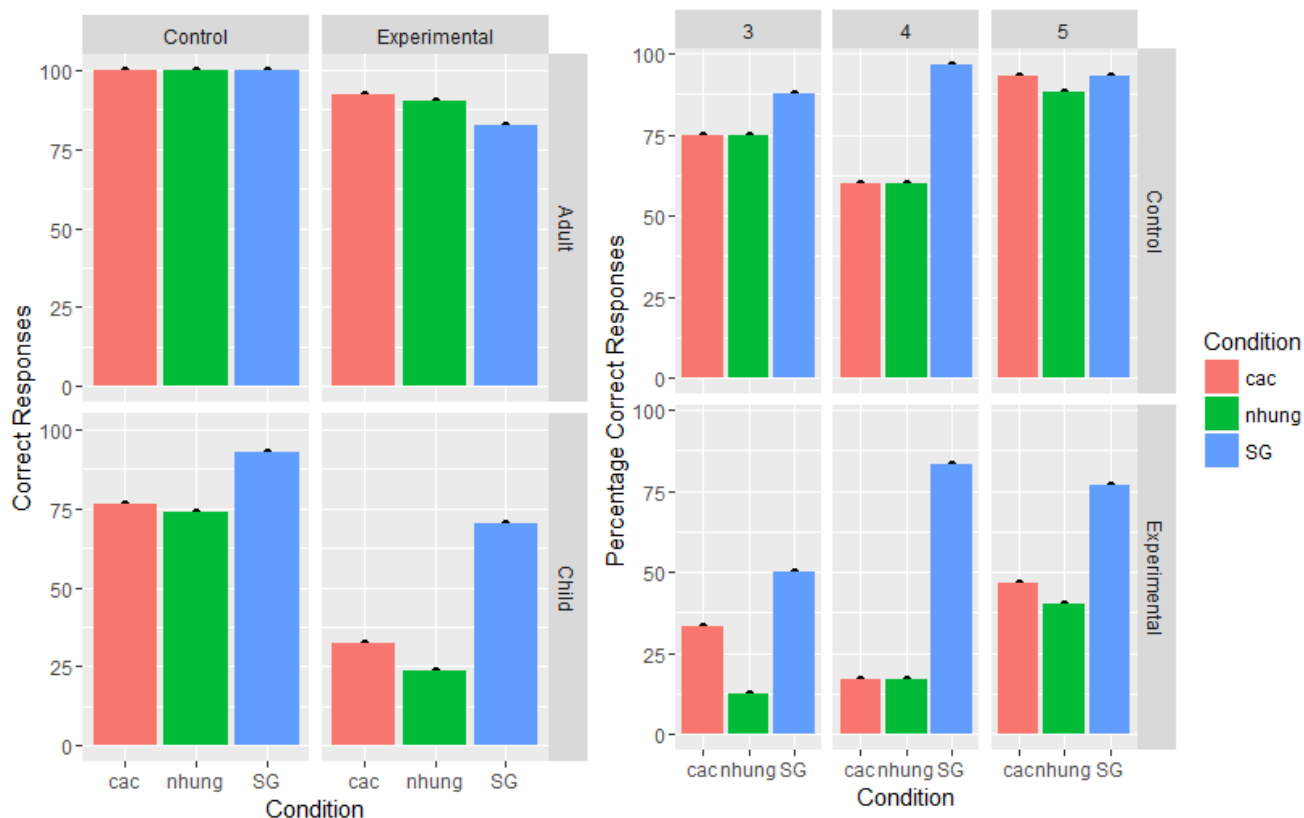


Figure 3: Percentage of correct responses (child & adult)

Figure 4: Percentage of correct responses across age groups (child)

The performance of child subjects in each age group (3 years old, $N = 12$; 4 years old, $N = 15$; and 5 years old, $N = 15$) were also examined (see Table 7). The percentage correct responses (i.e., target on both number and definiteness) in each condition per age group are summarized in Table 8 and illustrated in Figure 4 (above). As for control items, children of age 5 is higher than other groups (91.11% vs. 79.17/ 72.22%). With respect to experimental items, children of all ages do much better in singular condition than in conditions with pluralizers. Among all groups, 3 year-olds perform worst in the singular condition while 4-year-olds are worst with respect to *các/những*. A three-way ANOVA was conducted on the percentage correct responses with test type (Control vs. Test), condition (singular, plural *các*, plural *những*), and age group (3, 4, and 5 years old) as within-subjects factors. A significant effect of age group $F = 8.294$, $**p = .004 < .05$, condition $F = 16.574$, $***p = 1.81e-07 < .05$, and test type $F = 58.774$, $***p = 4.38e-13 < .05$ was found. There was also a significant interaction between condition and test type $F = 3.065$, $*p = .048 < .05$. Three sub-t-tests were run to decide exactly which age group leads to the differences in children's level of accuracy. The results suggested a significant difference between 5-year-olds and 3-year-olds ($t = -2.158$, $p = .034 < 0.05$) while 4-year-olds and 3-year-olds do not behave significantly different from each other.

Table 8: Percentage of correct responses across age groups (child)

Condition	Experimental items				Control items				Total
	CL	các	những	total	one	all các	all những	total	
3 years old (N=12)	45.83	33.33	12.5	30.56	87.5	75	75	79.17	54.86
4 years old (N=15)	83.33	16.67	16.67	38.89	96.67	60	60	72.22	55.56
5 years old (N=15)	76.67	46.67	40.00	54.44	93.33	93.33	86.67	91.11	72.78

Next, child responses were analyzed in three steps testing the research questions which were stated in (35) and are reordered here in (39).

(39) *Analysis steps:*

- Do children treat *những* and *các* differently (35c), given that adults unanimously comprehend both of them as definite pluralizers, as seen in Table 5?
- Do they know CI-N is associated with a singular reading and the pluralizers are added to derive a plural interpretation (35a)? In order to answer this question, I will isolate the singular/plural distinction.
- How do they interpret *những* and *các* in terms of definiteness (35b)? This question can be addressed by looking at maximal (exhaustive/closest) responses.

6.1 Results of experimental items

Recall the general question of this study is child acquisition of the three constructions of interest: classifier phrases with or without pluralizers. As Figure 3 already tells us the children tested do not have the adult-like interpretation of the pluralizers in Vietnamese, this section examines whether *các/những* are treated differently by children.

6.1.1 *Những* vs. *các*

As seen in Table 6 and 7, although children have a much lower target rate of *những/các* compared to adults, the exhaustive interpretation outweighs the non-exhaustive reading as clearly and as strongly (32.15% vs. 4.77% for *các* and 23.81% vs. 2.39% for *những*), differing from the traditional claim in the literature that *những* refers to a subset. Table 9 shows the percentage child correct responses to *những* and *các* in Experimental sentences, in terms of number (plural), definiteness (maximal), and both (target).

Table 9: Percentage and number correct responses (out of total responses) for pluralizers

Group	CÁC			NHỮNG		
	Plural	Maximal	Target	Plural	Maximal	Target
3 years old	41.67 (10/24)	33.33 (8/24)	33.33 (8/24)	20.83 (5/24)	12.5 (3/24)	12.5 (3/24)
4 years old	23.33 (7/30)	16.67 (5/30)	16.67 (5/30)	23.33 (7/30)	16.67 (5/30)	16.67 (5/30)
5 years old	46.67 (14/30)	46.67 (14/30)	46.67 (14/30)	40.00 (12/30)	40.00 (12/30)	40.00 (12/30)

This section is interested in target responses (i.e., those responses that are correct in **both** number and definiteness) while percentage plural and maximal responses will be discussed in Section 6.2 and 6.3 below. As expected from the ANOVA test run earlier, there is a difference between age groups: 5-year-old children do better than younger ones for both *các* and *những*. The older children who got the number feature of *những/các* correctly also were accurate in terms of their maximality. This was not the case for younger group: some of them who were right in terms of number were unsuccessful to pick out a maximal set. The percentages from Table 8 shows 4-year-olds treat *những/các* exactly the same while t-tests indicate that there is no significant difference between *những* and *các* for other age groups: $t = 1.239$, $p = .231$ (3 year olds), and $t = -0.35675$, $p = .724$ (5 year olds). Therefore, two conditions *những* and *các* are going to be collapsed as a single condition ‘pluralizers’ (PL) henceforth. T-tests between age groups also show only significant difference between 4-year-olds and 5-year-olds: $t = 2.368$, $*p = .022 < .05$.

6.1.2 [CL-N] vs. the pluralizers

Now let consider whether children differ in their comprehension between [CL-N] (which is singular, definite) and noun phrases with *những/các* (which, based on adults’ performance, are plural, definite). The percentage correct responses for each noun phrase type is presented in Table 10.

Table 10: Percentage correct responses to noun phrases with and without pluralizers (child)

Condition	[CL-N]	Pluralizers
3 years old	45.83	22.92
4 years old	83.33	16.67
5 years old	76.67	43.33
All children	70.24	27.98

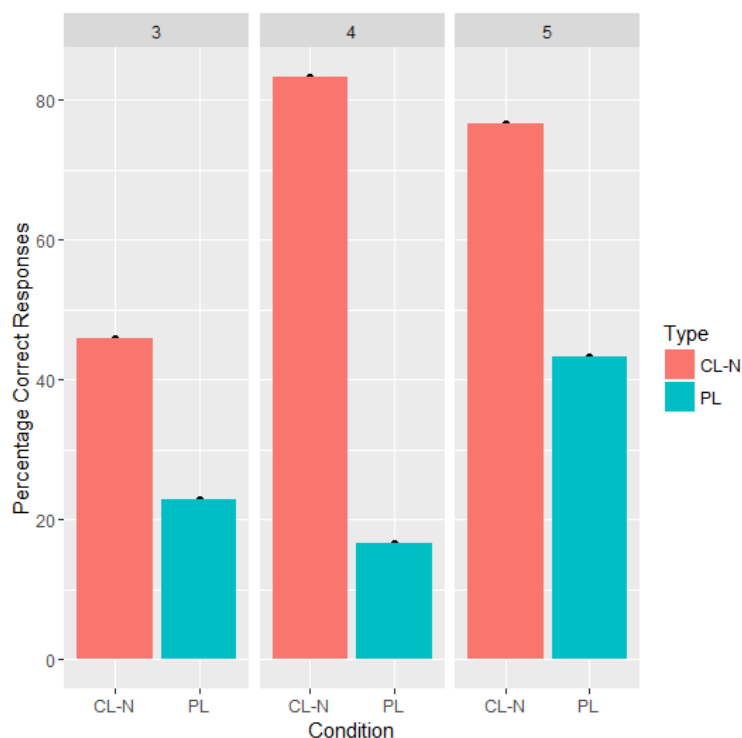


Figure 5: Percentage of correct responses for [CL-N] and [những/các-CL-N] per age group (child)

T-tests are run to see whether the difference between the two conditions is significant, and if yes, whether there is an age effect (Is it across the board? Which age group motivates that difference?). The result of each comparison is summarized in Table 11.

Table 11: T-tests comparison between conditions across and within age groups (child)

Comparison	Age group	Results	Conclusion
[CL-N] vs. pluralizers	all children	$t = -5.166, ***p = 1.577e-06$	Significant difference
	within 3 year-olds	$t = -1.372, p = .186$	Not significant difference
	within 4 year-olds	$t = -5.859, ***p = 2.796e-06$	Significant difference
	within 5 year-olds	$t = -2.507, *p = .017$	Significant difference
Within [CL-N]	3 y.os vs. 4 y.os	$t = -2.187, *p = .041$	Significant difference
	3 y.os vs. 5 y.os	$t = -1.784, p = .089$	Marginal difference
	4 y.os vs. 5 y.os	$t = -.498, p = .622$	Not significant difference
Within pluralized NPs	3 y.os vs. 4 y.os	$t = .584, p = .562$	Not significant difference
	3 y.os vs. 5 y.os	$t = -1.629, p = .109$	Not significant difference
	4 y.os vs. 5 y.os	$t = 2.368, *p = .022$	Significant difference

Table 11 and these tests suggest some following points. First, overall children are significantly better at understanding classifier phrases than they do noun phrases with pluralizers, especially children at the age of 4 and 5. Second, 4- and 5-year-olds got more correct responses for [CL-N] than 3-year-olds, even significantly more in the case of 4-year-old children. Third, 5-year-olds got significantly more correct responses for pluralizers than younger children. The age effects are illustrated in Figure 5 above.

6.2 Number

The study was initially interested in these expressions (bare classifier phrases and pluralizers) because these items encode more than one piece of information: number and definiteness. Thus, in order to fully understand how they are acquired by children, we also have to look at the acquisition of each feature. We are going to deal with number first and then maximality, since the latter must be defined within the former.

The percentage target responses in terms of number (i.e., singular for [CL-N] and plural for *những/các*) regardless of their accuracy in terms of definite is given in Table 12, followed by the summary of t-tests run to statistically examine children's comprehension of number indicated in these conditions.

Table 12: Percentage and number correct responses with respect to number (child)

Condition	[CL-N]	PL
3 years old	91.67 (22/24)	31.25 (15/48)
4 years old	93.33 (28/30)	23.33 (14/60)
5 years old	80 (24/30)	43.33 (26/60)
All children	88.09 (74/84)	32.74 (41/168)

Table 13: T-tests comparison between conditions across and within age groups with respect to NUMBER (child)

Comparison	Age group	Results	Conclusion
[CL-N] vs. pluralizers	all children	$t = -4.509, ***p = 2.005e-05$	Significant difference
	within 3 year-olds	$t = -4.802, ***p = 3.521e-05$	Significant difference
	within 4 year-olds	$t = -6.406, ***p = 1.479e-07$	Significant difference
	within 5 year-olds	$t = -2.77, **p = .009$	Significant difference
Within [CL-N]	3 y.os vs. 4 y.os	$t = 0.151, p = .881$	Not significant difference
	3 y.os vs. 5 y.os	$t = 0.923, p = .365$	Not significant difference
	4 y.os vs. 5 y.os	$t = -0.837, p = .41$	Not significant difference

Within pluralized NPs	3 y.os vs. 4 y.os	$t = 0.658, p = .514$	Not significant difference
	3 y.os vs. 5 y.os	$t = -0.917, p = .363$	Not significant difference
	4 y.os vs. 5 y.os	$t = 1.687, p = .097$	Marginal difference

In our experiment, as these statistics show, children in general were significantly more accurate in interpreting [CL-N] as singular than they were in associating pluralizers with plurality and this difference is strong across all age groups. In particular, they treat *những/các* as indicating ‘more-than-one’ not even close to chance (32.74 < 50%). There is no age effect either in singular or plural condition. Figure 6 demonstrates the figures in Table 12 and visualizes the significant difference between two noun phrase types present within all age groups.

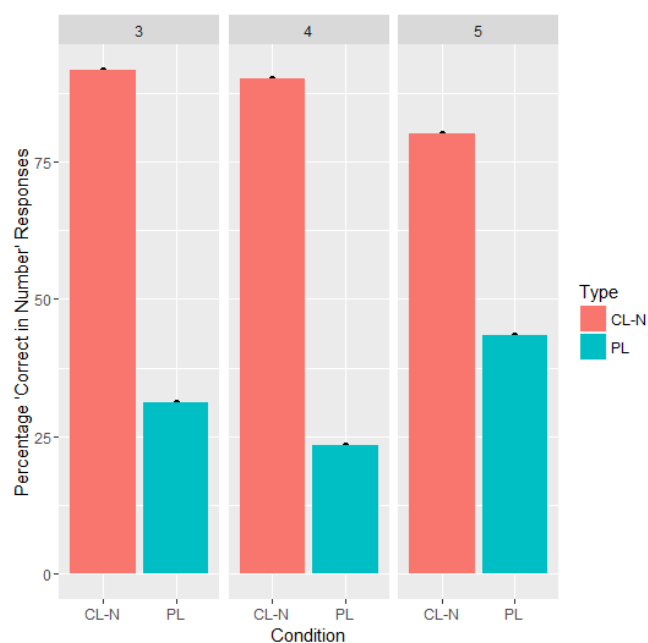


Figure 6: Percentage of correct responses for [CL-N] and [*những/các*-CL-N] in terms of number (child)

6.3 Maximality

The experiment is set up in the way that if children do know these constructions are definite, they would pick out the closest animal to the given landmark in the singular condition and the maximal set (3 animal toys close to the given landmark) in the plural condition. To be concise, this paper calls both choices as maximal responses. In consistent with the coding in Table 4, the percentage correct of maximal responses is coded in two ways: the number of maximal responses out of total responses (i.e., the target responses), which is used to compare between experimental and control items, and the number of maximal responses out of only plural choices for *các*, *những*, and *tất cả các/những* ‘all,’ and out of only singular choices for [CL-N] and *một* ‘one’ (i.e., within only ‘correct in number’ responses), which is used to tease apart between ‘No Maximality Hypothesis’ and ‘Domain Restriction Hypothesis’. Table 14 and 15 summarize the number and percentage maximal responses in these two different calculations.

Table 14: Percentage and number of maximal responses (out of correct responses) (child)

Group	Singular condition		Plural condition	
	[CL-N]	One	Pluralizers	All
3 years old	45.83 (11/24)	50 (12/24)	22.91 (11/48)	75 (36/48)
4 years old	83.33 (25/30)	66.67 (20/30)	16.67 (10/60)	60 (36/60)
5 years old	76.67 (23/30)	83.33 (25/30)	43.33 (26/60)	90 (54/60)
All children	70.24	67.86	27.97	75

Table 15: Percentage and number of maximal responses (out of ‘correct number’ responses) (child)

Group	Singular condition		Plural condition	
	[CL-N]	One	Pluralizers	All
3 years old	50 (11/22)	57.14 (12/21)	73.33 (11/15)	100 (36/36)
4 years old	89.29 (25/28)	68.97 (20/29)	71.43 (10/14)	83.72 (36/43)
5 years old	95.83 (23/24)	89.29 (25/28)	100 (26/26)	91.53 (54/59)
All children	79.73%	73.08%	85.45%	91.3%

6.3.1 Definiteness in singular condition

The two constructions [CL-N] and [*một* ‘one’-CL-N] are similar in number (singular) but different in terms of definiteness: the former is definite while the latter is indefinite. However, as seen in Table 14, children gave maximal answers not only in the definite context but also in the indefinite condition. A two-way ANOVA on proportion of target responses with condition ([CL-N] and ‘one’) and age group (3, 4, and 5) as within-subject factors was conducted and showed there was no significant difference between two conditions ($F=.035$, $p=.869$) or age groups ($F=6.335$, $p=.128$) and there is no interaction between age groups and conditions ($F=.010$, $p=.931$). This suggests children equally associate both expressions with a definite interpretation.

6.3.2 Definiteness in plural condition

Initially, the experiment aimed to test the difference between *các* which is a definite pluralizer and *những* which was commonly assumed as an indefinite plural. However, as seen in the results presented in 6.1.1, they both seem to be associated with a definite plural interpretation. Therefore, we turn into the comparison between the two pluralizers on one hand and ‘all’ on the other hand, which also requires the presence of these pluralizers. A two-way ANOVA on proportion of maximal responses (out of all responses) with condition (pluralizers and ‘all’) and age group (3, 4, and 5) as within-subject factors was conducted and showed there was a significant difference between two conditions ($F=27.095$, $*p=.035 < 0.05$) but there is no age affect ($F=2.204$, $p=.276$) and there is also no interaction between age groups and conditions ($F=.052$, $p=.841$). However, when we ran another similar two-way ANOVA on the proportion of maximal responses out of only ‘correct in number’ responses, there was no longer a significant difference between two conditions ($F=2.435$, $p=.259$). This may suggest that if a child knows *những/các* are plural, s/he also knows they are definite.

6.3.3 ‘No Maximality’ Hypothesis vs. ‘Domain Restriction’ Hypothesis

As discussed in Section 4, the ‘No Maximality’ Hypothesis predicts children, when dealing with definite items, to fail on both singular and plural conditions while the Domain Restriction’

Hypothesis predicts them not to fail in the plural condition but may fail in the singular condition. We are going to compare the maximal responses between these two conditions in our experiment to test these hypotheses out. The percentage of maximal responses in Table 15 will be used here because those responses that were wrong with respect to number could not participate in the maximality. Figure 7 illustrates the percentage maximal responses in the singular definite and the plural definite conditions.

Again, another two-way ANOVA on the percentage of maximal responses out of those correct in number, with conditions ([CL-N] vs. pluralizers) and age groups (3, 4, and 5) as within-subject factors, was conducted. The results show that there is significant different among age groups ($F=14.208$, $***p=.00036 <.05$) but no significant difference between two conditions ($F=0.007$, $p=.932$) and there is also no interaction between age groups and conditions ($F=0.172$, $p=.679$). Table 16 summarizes the results of the sub t-tests, in which 3-years old group differ significantly from the rests.

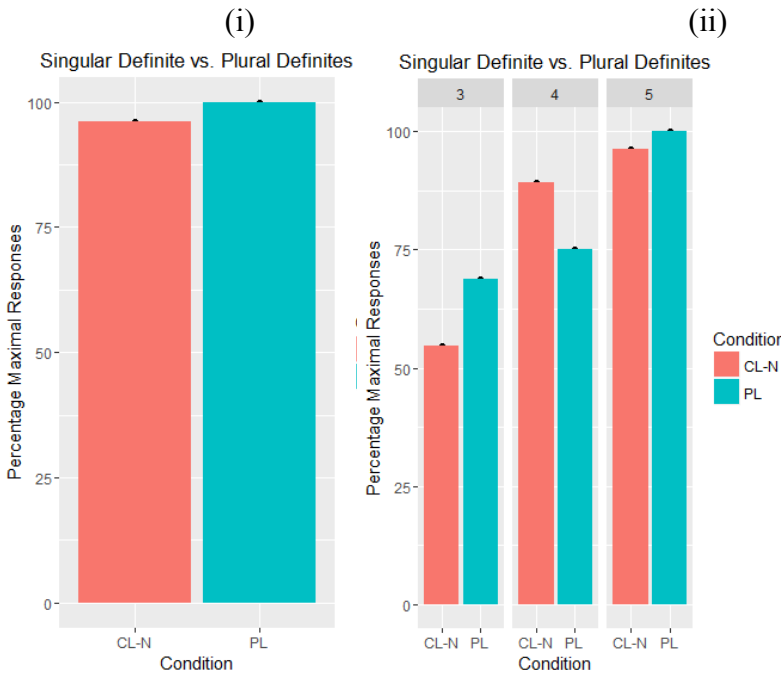


Figure 7: Percentage of maximal responses for singular definite and Plural definite expressions: (i) all child subjects; (ii) per age group

Table 16: T-tests comparison between age groups wrt maximal responses for [CL-N] and [những/các-CL-N]

Comparison	Age group	Results	Conclusion
Within [CL-N]	3 y.os vs. 4 y.os	$t(1, 15) = -2.14, *p = .048$	Significant difference
	3 y.os vs. 5 y.os	$t(1, 11) = -2.82, *p = .016$	Significant difference
	4 y.os vs. 5 y.os	$t(1, 18) = -0.79, p = .437$	Not significant difference
Within pluralized NPs	3 y.os vs. 4 y.os	$t(1, 13) = -0.27, p = .79$	Not significant difference
	3 y.os vs. 5 y.os	$t(1, 7) = -1.93, p = .095$	Not significant difference
	4 y.os vs. 5 y.os	$t(1, 7) = -1.53, p = .171$	Not significant difference

This means Vietnamese children tended to pick out the maximal set equally for both singular and plural definites. In other words, the maximality errors were not significantly different between the singular and the plural condition. While this result poses an opposition to the ‘No Maximality’ Hypothesis (children seldom associated definite pluralizers with non-maximal sets), it does not support the ‘Domain Restriction’ Hypothesis either: Vietnamese-speaking children were able to figure out the unique animal that a singular definite expression like [CL-N] refers to.

7 General discussion

Overall, Vietnamese-speaking children acquire the interpretation of bare classifier phrases (singular, definite) quite early, around 4 years old (4-year-olds in this experiment comprehend [CL-N] correctly 83.33% of the time). On the contrary, pluralizers in Vietnamese seem to be difficult for children of all ages tested to interpret: in this experiment children succeeded only 27.98% of the time with *những/các* noun phrases, and until 5 years old they still can be correct only at the rate of 43.33%. As we have seen the protraction in child acquisition of pluralizers in other classifier languages in previous studies, this is not surprising. However, the unexpected findings from this experiment are (i) opposite to our predictions and to the obligatoriness of *những/các* in many plural noun phrase types in Vietnamese, children seem to struggle with the plural feature of these pluralizers (even at the age of 5): they interpret them as singular for 67.26% of the times, and (ii) number feature and definite feature encoded in *những/các* seem to be a feature bundle and thus may be learned at the same time: Children made much less correct responses in terms of number than in terms of maximality (32.74% vs. 85.45%). The percentage correct responses for plural conditions were boosted significantly with the presence of ‘all’ (in control sentences). This may be because children know the lexical meaning of ‘all,’ in the same way they know ‘one.’

In general, children participating in this experiment displayed strong bias towards singularity and uniqueness. Singular choices were found not only in the singular condition but also in the plural condition at very high rates, suggesting children do not know the plurality of these pluralizers until much later. Meanwhile, within the singular condition, the closest animal toy was given not only for the definite expression (CL) but also in the indefinite item (‘one’). This unfortunately prevents us from claiming for sure children know what definite is for the singular condition.

With respect to the status of *những*, it patterns with *các* not only in the adult’s results in the sense that both indicate a maximal plural set, but also in children’s performance in every aspect such as percentage correct responses regarding number and maximality or types of errors. While the nature of *những* is still a controversial theoretical issue in the literature of Vietnamese linguistics, our results offer some empirical observations that can be used to justify one way or the other: in this case, *những* is in favor of a definite interpretation.

Our results also suggest development path in acquiring [CL-N] and the pluralizers. On one hand, [CL-N] seems to emerge first before children can acquire pluralizers. By age 3, children already interpret [CL-N] quite well and adult-like around 4. As we seen in the analysis of the results, child comprehension of these items is, in general, better over time. The difference between 3-year-olds and 5-year-olds is significant (cf. p. 26-27).

8 Conclusion

There is very little research on the acquisition of the Vietnamese language and even less on the acquisition of noun phrases in Vietnamese¹³. Our experiment was one of the first attempts to examine how children and adults interpret the number and definiteness of CL-phrases with and without the two pluralizers *những* and *các* and investigate in which order Vietnamese-speaking children acquire them. Based on the results, some preliminary conclusions can be drawn as follows. First, children from 3 to 5 years old know the interpretation of [CL-N] better than the interpretation of the pluralizers *những/các*. In fact, by age 5, Vietnamese speaking children still do not master fully the plural feature of these two pluralizers, although this develops over time. Second, children by age 4 can consistently associate [CL-N] with a unique reference. Children who already acquire the plural feature of *các/những* also interpret them as indicating a maximal set, i.e., definite, most of the time. Third, *những* can refer to an entire set, i.e., it has maximality presupposition like *các*.

Some further research can be made based on the results and design of this current experiment. The fact that children gave equal maximality errors in our experiment, though can be another piece of evidence against ‘No Maximality’ Hypothesis, has told us nothing in terms of ‘Domain Restriction’ Hypothesis. In addition, the fact children often gave the same responses (the unique item) in both definite and indefinite singular conditions, complicates the meaningfulness of their well performance for singular definite items. Furthermore, one can argue although *những* in fact could indicate a subset as well as a whole set, it might be weird to use *những* to target just 2 out of 3 items and that was why it behaved similar to *các* in our experiment. One way to get around these problems is to increase the animal toys used in the experiment (currently 3 per side), which makes ‘some’/indefinite meaning more felicitous and hopes to weaken the bias towards singularity/closest. Also, as children by age 5 still do not fully master the pluralizers, we might want to increase the age of subjects to 7 or 8, the age around which children of classifier languages master their pluralizers.

¹³ Tran’s doctoral dissertation, “The Acquisition of Vietnamese Classifiers” (2011) is the first and the only study on the acquisition of the Vietnamese language until now. However, her conclusions are drawn from a very small number of subjects (four children who were aged from 1;11 to 2;5). In addition, the study is mainly interested in syntactic errors and order of acquisition of different types of Cl-phrases and, therefore, all of the participants are not in the age range within which, as studies of other classifier languages have shown, the semantic properties of Cls are critically developed (between 3 and 6 years old) (Matsumoto, 1985a&b; Erbaugh, 1986; Li et al., 2008; Cheung et al., 2010).

References

- Barner D., Thalwitz D., Wood J., Yang S., & Carey S. (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.
- Borer, H. (2005). Structuring Sense, Vol I: In Name Only. Oxford: Oxford University Press.
- Brown, R. (1973). A First Language: The Early Stages. London: Allen & Unwin.
- Bui, M. H. (2000a). Về một số đặc trưng ngữ nghĩa - ngữ pháp của “những” và “các” [About Semantic and Grammatical Properties of *Những* and *Các*]. *Tạp chí Ngôn ngữ* [Linguistic Journal] 3: 16-26.
- Bui, M. H. (2000b). Về vấn đề quán từ và nhận diện quán từ trong tiếng Việt [About Articles and Identifying Articles in Vietnamese]. *Tạp chí Ngôn ngữ* 12: 1-16.
- Chierchia, G. (1998). Plurality of Mass Nouns and the Notion of “Semantic Parameter.” In: *Events and Grammar*. Rothstein, S. (ed.). Dordrecht (Netherlands): Kluwer, 53-103.
- Cheung, P., Barner, D., & Li, P. (2010). Syntactic Cues to Individuation in Mandarin Chinese. *Journal of Cognitive Science*, 10: 135-47.
- Corbett, G. G. (2000). Number. Cambridge: Cambridge University Press.
- Diep, B., & Hoang, T. (1998). Ngữ pháp Tiếng Việt [Vietnamese Grammar]. Hanoi: Education Publisher.
- Droz, K. F. (2001). Children’s Weak Interpretations of Universally Quantified Questions. In: *Language Acquisition and Conceptual Development*. Bowerman, M. & Levinson, S. (eds.). Cambridge University Press, 340-76.
- Erbaugh, M. S. (1986). Taking Stock: The Development of Chinese Noun Classifiers Historically and in Young Children. In: *Noun Classes and Categorization*. C. Craig (ed.). Amsterdam: John Benjamins, 399-436.
- Fehri, A. F. & Vinet, M. (2004). Number and Classifier Distributions in Arabic and Chinese and their complementarity. *Linguistic Research* 9: 9-52.
- Fenson, L. et al. (1994). Variability in Early Communicative Development. *Monographs of the Society for Research in Child Development* 59 (5): i-185.
- Forsythe et al. (2015). Experiments on Pronouns in Spanish and Vietnamese.
- Kadmon, N. (1990). Uniqueness. *Linguistics and Philosophy* 13: 273-324.
- Kang, B. (1994). Plurality and Other Semantic Aspects of Common Nouns in Korean. *Journal of East Asian Linguistics* 3 (1): 1-24.
- Karmiloff-Smith, A. (1979). A Functional Approach to Child Language. Cambridge University Press.
- Kim, J. (2008). The Parameterization of Plural Markings in Classifier Languages. *Kansas Working Papers in Linguistics*, 30: 159-73.
- Kim, M. (2011). The Acquisition of Korean Pluralizer *-tul*. Michigan State University: MA Thesis. ProQuest Dissertations Publishing.
- Kim, C., O’Grady, W., & Deen, K. (2012). The Extrinsic Plural Marker in Korean: Five Studies. Manuscript.

- Kouider, S., Halberda, J., Wood, J., & Carey, S. (2006). Acquisition of English Number Marking: The Singular-Plural Distinction. *Language Learning and Development* 2: 1-25.
- Heim (1991). Artikel und Definitheit (Article and Definiteness). In: *Semantics. An International Handbook of Contemporary Research*. Stechow, A. v. & Wunderlich, D. (eds.). Berlin: de Gruyter.
- Li, P., Barner, D., & Huang, B. (2008). Classifiers as Count Syntax: Individuation and Measurement in the Acquisition of Mandarin Chinese. *Language Learning and Development* 4(4): 249-90.
- Li, P., Barner, D., Oruga, T., Yang, S., & Carey, S. (2009). Does the Conceptual Distinction Between Singular and Plural Sets Depend on Language?. *Developmental Psychology*, 45(6): 1644-53.
- Li, Y. A. (2009). Plurality in a Classifier Language. *Journal of East Asian Linguistics*, 8(1): 75-99.
- Liter, A., Heffner, C., & Schmitt, C. (2013). Inclusive Plural Interpretations Depend on Grammaticalized Number: an Argument from Artificial Language Learning. Paper presented at the *Workshop on the Acquisition of Quantification*. Amherst, MA: University of Massachusetts.
- Lyons, C. (1999). *Definiteness*. Cambridge University Press.
- Maratsos, M. (1972). *The Use of Definite and Indefinite Reference in Young Children*. Cambridge University Press.
- Matsumoto, Yo (1985a). Japanese Numeral Classifiers: Their Structure and Acquisition. Tokyo (Japan): Sophia University. Master's Thesis.
- Matsumoto, Yo (1985b). Acquisition of Some Japanese Numeral Classifiers: The Search for Convention. *Stanford University Papers and Reports in Child Language Development* 24: 89-96.
- Mervis, C., & Johnson, K. E. (1991). Acquisition of the Plural Morpheme: A Case Study. *Developmental Psychology* 27(2): 222-35.
- Miller, K. and Schmitt, C. (2004). Wide-scope Indefinites in English Child Language. In: *Proceedings of Generative Approaches to Language Acquisition 2003 (GALA)*. Jacqueline, v. K. & Baauw, S. (eds.). Utrecht: LOT, 317-28.
- Miller, K. (2007). Variable Input and the Acquisition of Plurality in Two Varieties of Spanish. Ph.D. Dissertation. Michigan State University.
- Miller, K. & Schmitt, C. (2009). Variable vs. consistent input: comprehension of plural morphology and verbal agreement in children. In: *Merging Features: Computation, Interpretation, and Acquisition*. Brucart, J., Gavarró, A., & Solà, J. (eds.). New York, NY: Oxford University Press, 123-38.
- Munn, A., Miller, K. & Schmitt, C. (2006). Maximality and Plurality in Children's Interpretation of Definites. In: *Proceedings of BULCD 30*. Bamman, D., Magnitskaia, T., & Zaller, C. (eds.). Cascadilla Press, 377-87.
- Munn, A. & Schmitt, C. (2005). Number and Indefinites. *Lingua* 115: 821-55.

- Munn, A., Zhang, X., & Schmitt, C. (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.
- Nakano, N., Park, S., & Schmitt, C. (2009). A Comparative Study of the Acquisition of Plural Morphemes in Japanese and Korean. In: *Proceedings of the 11th Tokyo Conference on Psycholinguistics*.
- Nguyen, T. C. (1975). Từ loại Danh từ trong Tiếng Việt Hiện đại [The Word Class of Nouns in Modern Vietnamese]. Hanoi: Publisher of Social Sciences.
- Nguyen, T. H. (2004). The Structures of the Vietnamese Noun Phrase. Boston University: PhD Dissertation.
- Nomoto, H. (2013). Number in Classifier Languages. UMI Dissertations Publishing.
- Park, S. (2010). The Acquisition of Plural Morphemes in Korean. Michigan State University: MA Thesis. ProQuest Dissertations Publishing.
- Peters, A. M. (1997). Language Typology, Prosody and the Acquisition of Grammatical Morphemes. In *The Crosslinguistic Study of Language Acquisition*, Vol 5. Slobin, D.I. (ed.). Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Roberts, C. (2003). Uniqueness in definite noun phrases. *Linguistics and Philosophy* 26: 287-350.
- Sauerland, U., Yatsushiro, K., & Anderssen, J. (2005). The Plural is Semantically Unmarked. In: *Linguistic Evidence: Empirical, Theoretical, and Computational Perspectives*. Kepser, S. & Reis, M. (eds.). Mouton de Gruyter.
- Simpson, A., Soh, H. L., & Nomoto, H. (2011). Bare Classifiers and Definiteness: A Cross-linguistic Investigation. *Studies in Language* 35(1): 168-93.
- Thompson, L. (1965). A Vietnamese Grammar. Seattle: University of Washington Press.
- Wexler, K. (2011). Cues Don't Explain Learning: Maximal Trouble in the Determiner System. In: *The Processing and Acquisition of Reference*. Gibson, E. & N. J. Pearlmuter (eds.). Cambridge, Mass: The MIT Press, 15-42.
- Wood, J., Kouider, S., & Carey, S. (2009). Acquisition of Singular-Plural Morphology. *Developmental Psychology*, 45 (1): 202-6.
- Yang, C. (2002). Knowledge and Variation in Natural Language. Oxford: Oxford University Press.
- Zhang, X. (2006). Chinese *-men* and Associative Plural. Paper presented at *the International Conference on East Asian Linguistics*, Toronto University.