

A CORPUS-BASED STUDY OF CLASSIFIERS IN VIETNAMESE*

Thanh Hai Tran
University of Manitoba

Vietnamese has a rich inventory of classifiers. In this research, 150 classifiers have been found in three semantic types: human, animate, and inanimate. This research investigates the natural data of Vietnamese classifiers in terms of pragmatics and discourse structure on a corpus-study basis. I will present the distribution and frequency of the classifiers identified in the research and discuss the co-occurrence of two classifiers in Vietnamese, in which each classifier usually carries different properties. One of them may be omitted without changing the meaning of the noun, but probably losing some properties depending on the classifier type. I will also discuss the similarities and differences in the use of the two classifiers *cái* (inanimate) and *chiếc* (individual) in the paper.

1. Introduction and background

Vietnamese is a Mon-Khmer language of the Austroasiatic language family. It is an isolating and non-inflectional language with three regional dialects: Northern (Hanoi), Central (ThuaThien-Hue Province), and Southern (Ho Chi Minh City). A prominent feature of Vietnamese is that it has a very complex classifier system, which attracts attention from researchers within the country and worldwide. A classifier in Vietnamese is a word that categorizes the noun by grouping the thing denoted by the noun it precedes into a generalized classification (Diep 2005).

The complex classifier system in Vietnamese has led to controversial arguments among researchers about the number of classifiers, functions, and structures. As one of the isolating languages, Vietnamese tends to have a large number of numeral classifiers (Aikhenvald 2000), but researchers claim different numbers of classifiers in Vietnamese (Emeneau 1951). Many researchers claim about 140 or 150 classifiers (Adams 1989; Thompson 1965). Cao (1998) argued that there are three classifiers in Vietnamese only: *người* (human), *con* (animate), and *cái* (inanimate) as in (1-3) respectively while it is reported to have 195 classifiers by Nguyen (2002).

- (1) người mẹ¹
CL(human) mother
'the mother' (2.120)

*This research was presented at the 2019 CLA conference at the University of British Columbia in Vancouver, BC. I am very grateful to Dr. Terry Janzen, University of Manitoba, for all his valuable supervision, suggestions and encouragement. I would like to send my sincere thanks to Dr. Nicole Rosen, University of Manitoba, for her helpful comments. My special thanks also go to the CLA 2019 audience for their feedback. All errors are my own.

¹All the examples in the paper are from the Vietnamese Narrative Corpus (Tran 2018), unless otherwise indicated.

- (2) một *con* mèo
 one CL(animate) cat
 ‘a/one cat’ (1.138)
- (3) cái bếp
 CL(inanimate) kitchen
 ‘the kitchen’ (1.39)

There is a large discrepancy between the two opinions since these three general classifiers are the most common and widely recognized while other linguists see classifiers in a broader sense as the word preceding the noun to indicate the type that noun belongs to (Diep 2005). According to many researchers, in addition to the three general classifiers, there are many other specific classifiers in Vietnamese including *cây* (tree, long), *quả* (fruit, round), and *hòn* (round) as in (4-6) respectively (Emeneau 1951; Nguyen, D. H. 1957; Thompson 1965; Nguyen, P. P. 2002). Löbel (1996:172) claimed the ten ‘core’ classifiers of Vietnamese which are *cái* (inanimate), *cây* (tree, long), *chiếc* (individual), *con* (animate), *hòn* (stone, round), *quả* (fruit, round), *quyển* (volume), *sợi* (hair, thread, cord), *tấm* (a flat piece of material) and *tờ* (a sheet of paper, document).

- (4) một *cây* tre trăm đốt dài lắm
 one CL(tree, long) bamboo hundred knots long very
 ‘a very long bamboo tree of hundred knots’ (1.25)
- (5) hai *quả* bầu khô
 two CL(fruit, round) gourd dry
 ‘two dried gourds’ (2.69)
- (6) một *hòn* đá rất lớn
 one CL(stone, round) stone very big
 ‘a very big stone’ (2.276)

Typologically, many nouns may occur with more than one classifier (Dixon 1986). In Vietnamese, one noun can combine with more than one independent classifier, depending on a particular, shaped-related property of the referent in focus (Aikhenvald 2000; Nguyen 1957). The two functions that Vietnamese classifiers perform are to individuate the object denoted by the noun, and to classify, characterize or describe objects through definite features (Ly 1998). Vietnamese has numeral classifier system with two primary functions: classification and individualization (Bisang 1993, 1999). The most typical structure of Vietnamese classifier constructions has been claimed to be Numeral - Classifier - Noun (Aikhenvald 2000; Nguyen 1957; Thompson 1965).

The corpus for this research comes from children’s literature, folktales. The data consists of one hundred and seventeen Vietnamese folktales randomly selected from two books published in 2013 and 2016 in Vietnam. They cover a wide variety of topics including animals, country, family, talented people, and festivals. Each story is from about three to fifteen pages long. The folktales which are supposed to have originated many years ago, orally transmitted from generation to generation, were then collected, written and

Table 2: Frequency of human classifiers

<i>Human classifier</i>	<i>No. of occurrences</i>	<i>Frequency (No. of tokens per 10,000 words)</i>
người (human)	321	28.28
ông (human, male, old)	119	10.48
cô (human, female, young)	104	9.16
đứa (human, young)	93	8.19
thằng (human, male, low social status)	77	6.78
nhà (human)	38	3.35
bà (human, female, old)	35	3.08
cậu (human, male)	27	2.38
thầy (teacher, master)	25	2.20
chàng (human, male, young)	25	2.20
'others'	148	13.04
Overall	1012	89.16

The next frequent classifiers are *đứa* (human, young) and *thằng* (human, young, low social status) at the rates of 8 and 6 per 10,000 words respectively as in (9-10). In fact, in (10), two classifiers co-occur. This phenomenon will be discussed in section 3.2.

- (9) một *đứa* con trai khôi ngô
 one CL(human, young) child male smart
 'a smart son' (2.100)
- (10) hai *thằng* kẻ trộm
 two CL(human, l.s.s) CL(human) steal
 'two (male) thieves' (1.155).

2.2 Animate (non-human) classifiers

Twelve animate (non-human) classifier types have been found with 374 tokens in the corpus. Table 3 shows the frequency of frequent animate (non-human) classifiers in the data. The most frequent classifier is *con* (animate), which occurs with the majority of nouns denoting animals as in (2) at the rate of about 27 per 10,000 words.

Other animate classifier types are not used often. They occur at the rate of one or less per 10,000 words, such as *đàn* (herd/group) and *bầy* (herd/group) as in (11-12). These two classifiers in Vietnamese carry the same meaning as they indicate a group/herd of animals in general.

- (11) một *đàn* kiến
 one CL(herd/group) ant
 'an army of ants' (1.76)

- (12) một *bầy* cạp
 one CL(herd/group) tiger
 ‘a group/an ambush of tigers’ (2.101)

Table 3: Frequency of animate (non-human) classifiers

<i>Animate classifiers</i>	<i>No. of occurrences</i>	<i>Frequency (No. of tokens per 10,000 words)</i>
con (animate)	316	27.84
đàn (herd, group)	21	1.85
bầy (herd, group)	9	0.79
chú (human, male, young)	8	0.71
‘others’	20	1.76
Overall	374	32.95

It is surprising that eight human classifiers have been found to go with animate (non-human) nouns anthropomorphically, in which personification is likely to be used. One of the two kinship terms, *chú* (human, male, young), is used with nouns denoting animals as in (13). Other animate non-human classifiers are rarely used in the data.

- (13) một *chú* gấu
 one CL(human, male, young) bear
 ‘a bear’ (2.102)

2.3 Inanimate classifiers

A total of 113 inanimate classifier types have been identified among 930 tokens in the corpus. Table 4 shows the frequency of five most frequent inanimate classifiers. The other classifiers are less frequent, so I put them into the ‘others’ category. This collective group includes 108 different infrequent classifiers in the data.

As Table 4 shows, five most frequent inanimate classifier types in the data include *cái* (inanimate) as in (3) at the rate of about 20 per 10,000 words, *cây* (tree, long) 12 per 10,000 words, *quả* (fruit, round), *hòn* (round), and *chiếc* (inanimate) as in (4-6) and (14) at the rates of 4 to 2 per 10,000 words. These five classifiers are among the ten core classifiers of Vietnamese that have been claimed by Lobel (1996:172).

- (14) *chiếc* thuyền
 CL(inanimate) boat
 ‘a boat’ (1.86)

It is interesting to see that both *cái* (inanimate) and *chiếc* (individual) are used with non-living things and indicate a non-specific unit of objects as in (3) and (14) (Emeneau 1951; Nguyen 1957; Löbel 1996, 2000). In the data *cái* (inanimate) is the most frequent while *chiếc* (individual) is the fourth most frequent. However, *cái* (inanimate) is widely recognized as the general inanimate classifier and receives special attention from prior

researchers while *chiếc* (individual) does not. Whether they are both general classifiers and they are the same or different will be discussed in more detail in 3.3.

Table 4: Frequency of inanimate classifiers

<i>Inanimate Classifiers</i>	<i>No. of occurrences</i>	<i>Frequency (No. of tokens per 10,000 words)</i>
cái (inanimate)	235	20.70
cây (tree, long)	140	12.33
quả (fruit, round)	52	4.58
chiếc (inanimate)	37	3.26
hòn (round)	29	2.55
'others'	437	38.50
Overall	930	81.94

3. Discussion

3.1 Overview of Vietnamese classifiers in the corpus

Overall, 150 classifiers have been identified in the corpus with the total number of 2316 tokens including 34 human, 12 animate (non-human), and 113 inanimate classifiers, with eight human classifiers used with animate (non-human) nouns and *con* (animate) used with both human and inanimate nouns. The results of this study are more in line with the view that Vietnamese has about 200 hundred classifiers (Nguyen 2002) although the findings in this research are limited to only one genre due to the limited amount of data in a genre.

Frequency is not balanced among classifier types in the data, and the frequency of tokens in three semantic classifier types (inanimate, human, and animate) is not balanced, either (see Table 1). The inanimate classifier has a high type frequency with the greatest variety of inanimate classifiers because nouns denoting non-living things may belong to many different classifications and have many different properties. In addition, in Vietnamese several different classifiers can be used with a noun (Nguyen 1957). On the contrary, the human classifier type has the highest token frequency, so in terms of lexical strength it would be the strongest. However, “lexical connections among low-frequency items are stronger than those among high-frequency items” as Bybee (1985:133) argues. That is, high-frequency classifiers tend to be stored as whole autonomous units, while low-frequency classifiers are stored with strong connections to others (Bybee 1985, 2006). Thus, in this case, speakers tend to associate inanimate classifiers more strongly with one another than they do to human classifiers, and this leads to a greater lexical strength for the inanimate classifier type.

Frequency is also not balanced among classifier construction categories in the corpus. Among nine classifier constructions found in the data, classifier construction 1, (Numeral) - Classifier - Noun - (Adjective Phrase/Noun Phrase/Verb Phrase/Prepositional Phrase) - (Demonstrative/Ordinal Number/Wh-word) - (Possessive Pronouns), is the most frequent,

with the widest distribution of classifier types. Thirty-two out of 34 human classifiers, 11 out of 12 animate (non-human) classifiers, and 105 out of 113 inanimate classifiers have been found in this construction. Overall, this construction accounts for 85% of all the classifier tokens in the data.

It should be noted that the Numeral in the classifier construction Numeral - Classifier - Noun in this corpus is optional. To be specific, only 23% of all the tokens in the corpus follow this pattern. However, 60.3% of the tokens in the data have the construction Classifier - Noun. The results suggest that the pattern Classifier - Noun is a better candidate to be the prototypical classifier construction in Vietnamese, not the pattern Numeral - Classifier - Noun as claimed by Aikhenvald (2000) and other prior researchers. This finding also supports Daley (1998)'s discussion. In other words, if the pattern Numeral - Classifier - Noun is considered to be the prototypical classifier construction in Vietnamese, the Numeral is thus optional.

The three general classifiers *người* (human), *con* (animate) and *cái* (inanimate) are the most frequent in the three semantic classifier types. The frequency rates of these classifiers greatly differ. The classifier *người* (human) accounts for 32%, with 321 occurrences out of 1012 tokens. It seems to be the most semantic neutral classifier (Löbel 1996:174). The choice of human classifiers probably depends on the speaker's attitudes, perceptions, and/or social status, but this is a topic for further research. Thus, special classifiers in the human type can be said to be complex, especially the ones homonymous with kinship terms. For the animate classifier type, *con* (animate) accounts for 84%, with 316 occurrences out of 374 tokens. Meanwhile, *cái* (inanimate) accounts for 27%, with 235 occurrences out of 930 inanimate classifier tokens. The frequencies of the other inanimate classifiers vary widely, ranging from 14% such as *cây* (tree, long), *quả* (fruit, round) to less than 1% such as *con* (tornado) or *viên* (stone).

Interestingly, the classifier *con* (animate) is also used with some nouns indicating humans and inanimate things. This makes the Vietnamese classifier system more complex.

3.2 The co-occurrence of two classifiers in Vietnamese

It is interesting to find that two classifiers in Vietnamese co-occur in the data. This evidence supports Aikhenvald's claim that "different types of numeral classifiers may co-occur and display different properties" (Aikhenvald 2000:112). The co-occurrence of two classifiers has been found in two constructions in the corpus. However, this phenomenon is infrequent in the data.

Classifier construction (Numeral) - Classifier - Classifier - Noun - (Adjective Phrase) - (Demonstrative) occurs 64 times in the human classifier type, once in the animate (non-human), and once in the inanimate classifier type. Three pairs of co-occurred classifiers in the human classifier type are '*người con*' (human - animate), '*đứa con*' (human, young - animate), and '*cô con*' (human, female, young - animate). The first classifier in these pairs are all human classifiers in which *người* (human) is the general classifier; *đứa* (human, young) is used to refer to young persons of unspecified sex; and *cô* (human, female (f), young (y)) is the one indicating young females. The second classifier in these pairs is the general classifier *con* (animate). All these double classifiers, which are made of three

different human classifiers and the general classifier *con* (animate), go with nouns denoting boys or girls as in (15a-17a). This construction Classifier - Classifier - Noun can optionally have a numeral as a premodifier and/or a prepositional phrase, adjective phrase, or demonstrative as post-modifiers as in (15a-17a).

- (15) a. *một đứa con gái trong làng*
 one CL(human, young) CL(animate) girl in village
 ‘a girl in the village’ (1.29).
 b. *một đứa gái trong làng*
 one CL(human, young) girl in village
 ‘a girl in the village’
- (16) a. *người con trai*
 CL(human) CL(animate) boy/man
 ‘the boy’ (2.228).
 b. *người trai*
 CL(human) boy/man
 ‘the boy/man’
- (17) a. *cô con gái đẹp*
 CL(human, f, y) CL(animate) girl beautiful
 ‘the beautiful girl’ (1.141).
 b. *cô gái đẹp*
 CL(human, f, y) girl beautiful
 ‘the beautiful girl’²

The second classifier *con* (animate) in (15a-17a) can be omitted without changing the meaning of the noun phrases as in (15b-17b). This means that the second classifier *con* (animate) in this case may not perform an important function in the noun phrases.

The double classifier occurs only once in the animate (non-human) classifier type in this construction is *anh chàng* as in (18a).

- (18) a. *anh chàng hổ ngu ngốc*
 CL(human, male, young) CL(human, male, young) tiger stupid
 ‘the stupid tiger’ (2.67).
 b. *con hổ ngu ngốc*
 CL(animate) tiger stupid
 ‘the stupid tiger’

In fact, both *anh* (human, male, young) and *chàng* (human, male, young) in (18a) are human classifiers to indicate young male. They are thus combined to emphasize the characteristic of being young and male of the tiger. In this case, personification is used for the animate non-human noun ‘tiger’, so the double classifier is used to emphasize that the tiger is definitely a young male one. That means, either of the two classifiers in (18a) can

² The examples in (b) and (c) in (15-21) are not from the corpus as they are given for illustration only.

frequency with 37 occurrences (4%). *Cái* (inanimate) appears in six different classifier constructions while *chiếc* (individual) is used in only one construction in the data.

Cái (inanimate) and *chiếc* (individual) are two of the ten ‘core’ classifiers claimed by Löbel (1996:172), and both indicate unspecified things. However, *cái* (inanimate) can be used for all kinds of non-living things, even for some small living things such as *cái kiến* (CL ant) as claimed by Nguyen (1957:127). Meanwhile, it is claimed that *chiếc* (individual) is the “classifier for nouns denoting artificial, individual items” (Löbel 2000:298). It is argued that *chiếc* is “one of a pair”, such as *chiếc đũa* ‘a chopstick’, *chiếc giày* ‘a shoe’, and *chiếc tất* ‘a sock, stocking’ (Emeneau 1951:106). It is also discussed that two nouns which do not “denote members of pairs but are classified with *chiếc* are *chiếu* (mat) and *nhẫn* (ring)” (Emeneau 1951:106). In addition, *chiếc* (individual) classifies nouns denoting boats and ships, and also vehicles (alternatively classified with *cái* (inanimate)) such as *chiếc tàu* ‘ship’, *chiếc thuyền* ‘boat’, and *chiếc xe* ‘vehicle’ (Emeneau 1951:106).

The results of this study show that *chiếc* (individual) is not only used with nouns denoting “one of a pair” or vehicles but also with many other individual items. To be specific, *chiếc* (individual) and *cái* (inanimate) can be used interchangeably for certain nouns to indicate non-living individual items from small to bigger ones as in (22-27).

- (22) a. *cái* chày
CL(inani) pestle
‘the pestle’ (1.102)
- b. *chiếc* chày
CL(inani) pestle
‘the pestle’ (2.112, 2.113)
- (23) a. một *cái* gùi lớn
one CL(inani) papoose big
‘a big papoose’ (2.147)
- b. *chiếc* gùi mới
CL(inani) papoose new
‘the new papoose’ (2.147)
- (24) a. bốn *cái* bánh
four CL(inani) cake
‘four cakes’ (1.43)
- b. bốn *chiếc* bánh
four CL(inani) cake
‘four cakes’ (1.43)
- (25) a. *cái* giường
CL(inani) bed
‘the bed’ (1.39)
- b. *chiếc* giường
CL(inani) bed
‘the bed’ (2.315)

- (26) a. *cái* lưới
CL(inani) net
'the net' (2.291)
- b. *chiếc* lưới
CL(inani) net
'the net' (2.311)
- (27) a. một *cái* thuyền lớn
one CL(inani) boat big
'a big boat' (2.180)
- b. một *chiếc* thuyền lạ
one CL(inani) boat strange
'a strange boat' (2.312)

However, the data shows that *chiếc* (individual) is used for a limited number of nouns, including small individual items such as axe, sword, jar, or bigger items such as net, blanket, bed; vehicles such as boat; “one in a pair” such as a chopstick, a shoe. In contrast, the result shows that only *cái* (inanimate) can be used for abstract nouns and nouns denoting things that can not be moved such as *nhà* ‘house, building’ as in (28a-29a), while *chiếc* (individual) can not occur with these types of nouns as in (28b-29b). In any linguistic environment, the classifier *chiếc* (individual) can never go with abstract nouns such as ‘trick’ or nouns indicating unmovable things such as ‘house, building’.

- (28) a. *cái* mẹo
CL(inani.) trick
'the trick' (1.135)
- b. **chiếc* mẹo
CL(inani.) trick
'the trick'
- (29) a. *cái* nhà
CL(inani.) house
'the house' (1.38, 1.39)
- b. **chiếc* nhà
CL(inani) house
'the house'³

In sum, the classifier *chiếc* (individual) can be used for certain nouns while the general classifier *cái* (inanimate) can be used with almost all nouns denoting inanimate things. I assume that only *cái* (inanimate) is the general classifier while *chiếc* (individual) is of limited use in comparison with *cái* (inanimate). To be specific, *chiếc* (individual) may not be used for abstract nouns and unmoveable things although it can go with a variety of nouns indicating a small object to a big one including ‘cake, bed, and boat’.

³ The examples in (28b-29b) are not found in the corpus, but are given for illustration.

4. Conclusion

In conclusion, the study has brought an overall picture of classifier use in Vietnamese and provided evidence to support the claims of prior research that Vietnamese classifiers can be words used to individuate nouns and categorize nouns into different classification with reference to their properties such as age, shape, gender or social status (Nguyen, D. H. 1957; Nguyen, P. P. 2002; and Nguyen, H. T. 2013). The results of the study show that 150 classifiers are used in three semantic types in this corpus. This means that the number of Vietnamese classifiers is not confined to three general classifiers as Cao (1998) argues. However, the actual number of Vietnamese classifiers may be higher than the number of classifiers found in this study as the corpus for this research is of one genre only. There might be other classifiers in Vietnamese spoken and written discourse of other genres, which requires further research. The findings of this research, therefore, support the claims of previous researchers that Vietnamese has about 200 classifiers (Emeneau 1951; Nguyen, D. H. 1957; Nguyen, P. P. 2002).

In addition, the data shows that two classifiers co-occur in Vietnamese although this phenomenon is restricted to a limited number of classifiers in the data. It is observed that in Vietnamese double classifiers, one may be a general classifier while the other can be a specific classifier. The general classifier is usually in the first position, but may be in the second position in some other cases. One of the classifiers in the double classifier construction in Vietnamese usually can be omitted. The primary meaning of the noun phrase does not change when the general one is left out. In contrast, the noun phrase may lose some properties when the specific classifier is omitted. However, in some cases, both classifiers are specific, so the noun phrase just loses some properties or pragmatic meaning when one of the classifiers is removed. I argue that the use of classifiers in Vietnamese is very complex, especially double classifiers, and this requires further research.

In sum, the study has found out interesting findings which support previous research and set the foundation for my future research on Vietnamese classifiers as there are still many other issues about classifiers in this language that need to be investigated. Vietnamese classifiers are obligatory in many cases, but not in some other cases (Simpson and Ngo 2018). Nevertheless, this issue is not discussed in the paper as it requires another study. This research investigates the use of classifiers in a single genre, folktale narrative, the findings of the study are thus limited to that genre. The use of classifiers, especially double classifiers, in conversational discourse and written discourse of various genres is an issue for my future study.

References

- Adams, Karen Lee (1989). *Systems of numeral classification in the Mon-Khmer*, Nicobarese and Asian subfamilies of Austroasiatic. Research School of Pacific Studies, The Australian National University Canberra, Australia.
- Aikhenvald, A. Y. (2000). *Classifiers: A Typology of Noun Categorization Devices*. Oxford University Press.
- Bisang, Walter (1993). *Classifiers, Quantifiers and Class Nouns in Hmong*. Studies in Language, pp.1-51.
- Bisang, Walter (1999). Classifiers in East and Southeast Asian languages: counting and beyond. In

- Gvozdanovic, J. (Ed.) *Numeral Types and Changes Worldwide*, pp. 113-185. Mouton de Gruyter, Berlin.
- Bybee, Joan (2006). *Frequency of Use and the Organization of Language*. Oxford University Press.
- Bybee, Joan (1985). *Morphology: A study of the relation between meaning and form*. John Benjamins Publishing Company.
- Cao, Xuan Hao (1998). *Tiếng Việt: Mấy vấn đề về Ngữ Âm, Ngữ Pháp, Ngữ Nghĩa (Vietnamese: Issues in Phonetics, Syntax, and Semantics)*. Ho Chi Minh City: Nhà Xuất bản Giáo dục (Education Publisher).
- Daley, Karen Ann. (1998). Vietnamese classifiers in narrative texts. A Publication of the Summer Institute of Linguistics and the University of Texas at Arlington.
- Diep, Quang Ban. (2005). *Ngữ pháp tiếng Việt (Vietnamese Grammar)*. Thua Thien - Hue: Nhà Xuất bản Giáo dục.
- Dixon, R.M.W. (1986). Noun Classes and Noun Classification in typological perspective. In Craig, Colette (Ed.), *Noun classes and categorization*, (pp.105-112). John Benjamins Publishing Company. Amsterdam/Philadelphia.
- Emeneau, M. B. (1951). *Studies in Vietnamese (Annamese) Grammar*. University of California Press. Berkeley and Los Angeles.
- Löbel, Elisabeth (1996). Classifier systems versus noun class systems: A case study in Vietnamese. In Barbara Unterbeck et. al. (Ed.) *Approaches to gender*. Berlin: Mouton de Gruyter.
- Löbel, Elisabeth (2000). Classifiers versus genders and noun classes: A case study in Vietnamese. In Barbara Unterbeck et. al. (Ed.) *Gender in Grammar and Cognition*, pp. 259-319. Berlin: Mouton de Gruyter.
- Ly, Toan Thang (1998). *Representation of space in Vietnamese classifiers*. Mon-Khmer Studies 29:71-80.
- Nguyen, Dinh Hoa (1957). *Classifiers in Vietnamese*. Word, 13:1, 124-152, DOI: 10.1080/00437956.1957.11659631.
- Nguyen, Hung Tuong (2013). The Vietnamese noun phrases. In Hole, Daniel; Lobel, Elisabeth (Eds.), *Linguistics of Vietnamese – An International Survey*, (pp. 57-87). De Gruyter Mouton.
- Nguyen, Phu Phong (2002). *Những vấn đề ngữ pháp tiếng Việt (Vietnamese grammar issues)*. Nhà Xuất bản Đại học Quốc gia Hà Nội (National University Publisher – Hanoi).
- Pham, Giang and Kohnert, Kathryn (2008). *A corpus-based analysis of Vietnamese con and cái*. Mon-Khmer Studies 38: 161-171.
- Simpson, Andrew and Ngo, Binh (2018). Classifier syntax in Vietnamese. *J East Asian Linguist* 27:211-246. Springer Science + Business Media B.V., part of Springer Nature. <https://doi.org/10.1007/s10831-018-9181-5>
- Thompson, L. C. (1965). *A Vietnamese Grammar*. Seattle: University of Washington Press.
- Tran, T. Thanh Hai (2018). *A corpus-based study of the functions and structure of classifier constructions in Vietnamese*. Manuscript. Generals paper. University of Manitoba.
- Microsoft Excel 2016.
- Corpus:
- Ôn Như Nguyễn, Văn Ngọc (2016). *Truyện cổ nước Nam (Vietnamese folktales)*. Nhà Xuất bản Kim Đồng (Kim Dong Publisher). Part: People, Book 1.
- Lữ, Huy Nguyên and Đặng, Văn Lung (Ed.) (2013). *100 truyện cổ tích Việt Nam (100 Vietnamese folktales)*. Nhà Xuất bản văn học Đông Á (East Asia Literature Publisher).