

A Contrastive Study of Thai and Vietnamese Collocations

Kanchanita Khunthamsakul¹, Achara Wongsothorn² and Suwaree Yordchim³

¹⁻³ Linguistics Program, Suan Sunandha Rajabhat University

Abstract

This study implemented a contrastive analysis to find the similarities and the differences between collocations in Thai and Vietnamese that convey the five sensory perceptions of: sight, taste, smell, hearing, and touch. In total 548 collocations were studied. Data analysis was divided into 3 parts: (1) the components of collocations used in Thai and those used in Vietnamese; (2) the 8 categories of the collocation combination structures: a noun with a noun, a noun with an adjective, a noun with a verb, a verb with a verb, a verb with an adverb, a verb with a noun, an adjective with an adjective, and an adjective with a noun; and (3) the comparison between collocations used in Thai and those used in Vietnamese, which were categorized into three types. The results showed that 1) there were 404 similar collocations or 73.72%, and 144 different collocations or 26.28%, 2) concerning the collocation combination structures, there were 422 similar collocation structures or 77.01% while there were 126 different collocation structures or 22.99%, and 3) there were more similarities than differences. The study revealed that there were 372 collocations of similar classifications or 67.88% and 176 collocations of different classifications or 32.12%. The collocations are arbitrary, thus, this study of the similarities and the differences of collocations in the languages can be a guideline for improving and increasing efficiency in both learning and teaching Vietnamese in terms of implementing correct, appropriate, and natural choices of collocations in classroom interaction.

Keywords: Base word, Collocator, Collocations, Thai, Vietnamese

การศึกษาการเปรียบเทียบคำปรากฏร่วมวลีในภาษาไทยและภาษาเวียดนาม

กานต์ชนิตา คุณธรรมสกุล, อัจฉรา วงศ์โสธร และสุวรรีย์ ยอดฉิม

บทคัดย่อ

การศึกษานี้ มีวัตถุประสงค์เพื่อวิเคราะห์เปรียบเทียบความเหมือนและความแตกต่างของคำปรากฏร่วมวลีในภาษาไทยและภาษาเวียดนามที่เกี่ยวกับประสาทสัมผัสทั้ง 5 คือ รูป รส กลิ่น เสียง สัมผัส โดยการนับจำนวนและคิดเป็นร้อยละ ข้อมูลมีจำนวนทั้งสิ้น 548 คำปรากฏร่วมวลี โดยแบ่งการวิเคราะห์ออกเป็น 3 ส่วน คือ ส่วนที่ 1 ศึกษาส่วนประกอบคำปรากฏร่วมวลีในภาษาไทยและภาษาเวียดนาม ส่วนที่ 2 ศึกษาโครงสร้างคำปรากฏร่วมวลีในภาษาไทยและภาษาเวียดนาม ซึ่งมีทั้งหมด 8 กลุ่ม คือ คำนามกับคำนาม คำนามกับคำคุณศัพท์ คำนามกับคำกริยา คำกริยากับคำกริยา คำกริยากับคำกริยาวิเศษณ์ คำกริยากับคำนาม คำคุณศัพท์กับคำคุณศัพท์ และคำคุณศัพท์กับคำนาม ส่วนที่ 3 ศึกษาเปรียบเทียบประเภทของคำปรากฏร่วมวลีในภาษาไทยกับภาษาเวียดนามมีทั้งหมด 3 ประเภท การศึกษาส่วนที่หนึ่งพบว่า มีความเหมือนมากกว่าความแตกต่าง คือมีส่วนประกอบคำปรากฏร่วมวลีที่ความเหมือนกัน 404 คำปรากฏร่วมวลี คิดเป็นร้อยละ 73.72 และมีความแตกต่างกัน 144 คำปรากฏร่วมวลี คิดเป็นร้อยละ 26.28 ส่วนที่สองพบว่า โครงสร้างคำปรากฏร่วมวลีมีความเหมือนมากกว่าความแตกต่าง คือ มีความเหมือนกัน 422 โครงสร้างคำปรากฏร่วมวลีคิดเป็นร้อยละ 77.01 และมีโครงสร้างคำปรากฏร่วมวลีที่มีความแตกต่างกันคิดเป็นร้อยละ 22.99 ส่วนที่

สามพบว่า ประเภทของคำปรากฏร่วมวลีมีความเหมือนมากกว่าความแตกต่าง คือมีความเหมือนกัน 372 ประเภท คิดเป็นร้อยละ 67.88 และมีความแตกต่างกัน 176 ประเภท คิดเป็นร้อยละ 32.12 คำปรากฏร่วมวลีไม่มีกฎเกณฑ์ตายตัว ดังนั้น การศึกษาความเหมือนและความแตกต่างของภาษาจะเป็นแนวทางในการปรับปรุงและช่วยเพิ่มประสิทธิภาพในการเรียนการสอนภาษาเวียดนามและเป็นแนวทาง ในการเลือกใช้คำที่เกิดร่วมกันอย่างถูกต้อง เหมาะสม และเป็นธรรมชาติ

คำสำคัญ คำศัพท์ฐาน คำศัพท์ร่วม คำปรากฏร่วมวลี ภาษาไทย ภาษาเวียดนาม

1. Introduction: Statement and Significance of the Research Problems

How can we learn a foreign language efficiently and effectively? What does one have to do? Despite adequate knowledge of both vocabulary and grammar, one may still not be capable to write or articulate a correct sentence. This might be caused such problems as incorrect word choice, because of similar meanings, or incorrect syntactic combinations. Sometimes one can make up a sentence grammatically correct but it is not natural for a native speaker. These challenges cannot be solved by just opening up a dictionary.

We have chosen to study the topic of collocation which is an essential factor in the teaching and learning a foreign language, but the problem of collocation is often ignored and treated with a grain of salt. This is because the use of collocation is arbitrary. Thus, the emphasis on rote learning of individual words is not sufficient to make one good at learning a foreign language, (Wong Wattanaphichet, 1999, p. iv). “Collocations” are groups of words used together, this linguistic feature of “collocations” has numerous aliases that are worth mentioning: “words in juxtaposition”, “word combinations” or “intertextuality”. The royal Institute uses the word “Collocators” (คำปรากฏร่วมวลี) /kham^M bpra^M goht^L ruaam^F champho/ or (สหพบ) /sa^L ha^L boht^L/. However, for this study, the researcher uses “collocation” (คำปรากฏร่วมวลี) /kham^M bpra^M goht^L ruaam^F wali/ which is derived from the Latin meaning “combining/collocating”. This word has been used in English long before J.R. Firth the “father” of collocation, first brought it into use as a terminology in linguistics in 1951.

Hill (1993:1) noted that concerning collocation competence native speakers of a language would instinctively know whether the spoken collocations were correct, naturally combined, or irregularly used. Lewis and Hill (1998) additionally noted that “Collocations” can be expected because if one head word appears, the other word will follow.

The Oxford Advanced Learner’s Dictionary of Current English (1974) indicates that collocations are both arbitrary and unpredictable. At first it seemed that these two opinions were incompatible, but if considered more broadly, these two opinions on collocations have compatible points of views as follows:

- 1) ฝนตกหนัก /fohn^R dtohk^L nak^L/ (The rain falls heavily.)
- 2) แดดออกจ้า /daaet^L aawk^L jaa^F/ (The sun shines brightly.)
- 3) ลมพัดแรง /lohm^M phat^H raaeng^M/ (The wind blows strongly.)

All three sentences have the same word order. Each sentence has the same number of words. They all convey about the three phenomena. The rain, the sun and the wind have different verbs to co-occur like the rain co-occurs with fall, the sun co-occurs with shine, and the wind co-occurs with blow. When they need adverbs to show high levels of quantity, three different adverbs are used.

Therefore, the adverbs and verbs used in these sentence samples are collocations for each kind of phenomena. If the collocations are irregularly used; for example, *"the rain shines; the rain falls brightly; the rain blows.", or *"the wind falls; the wind shines brightly; *the wind blows brightly." The above sentences would be irregular.

Perhaps native speakers might be able to guess and understand them. However, they are not in natural language. For this point of view; collocations can be expected or guessed. In Thai when we talk about the rain it can be expected to be used with the verb "fall" (ตก) /dtohk^L/ for the collocators and adverb "heavily" (หนัก) /nak^L/. They will not use "shine" for the verb and "brightly" (จ้า) /jaa^F/ for the adverb. The idea that collocations are arbitrary and unpredictable can be explained further that there are not any rules saying that when the head word is "the rain" (ฝน) /fohn^R/ in Thai, the verb that follows has to be "fall" (ตก) /dtohk^L/. The words that follow might be "an adjective- persistent" (คร่ำม) /khreum^H/, "another verb – come soon" (มาเร็ว) /ma meuu^F/ or "a prepositional phrase – from the North" (จากทางเหนือ) /jaak^L thaang^M neuua^R/. Likewise, the verb "fall" (ตก) /dtohk^L/ might follow other headwords, not only "the rain" (ฝน) /fohn^R/; for example, สีตก /see^R dtohk^L/ (The color is washed out.), or หุ้นตก /hoon^F dtohk^L/ (The stock falls.), etc.

Collocations are the word combinations of two or more words used together. That is when one head word comes, the other follows. It is a group of words that needs to be together sentences which are commonly used among native speakers, and which are correctly and naturally used. Besides, some collocations can be neither replaced by others nor have their word order switched even though substituted words have the same or similar meaning.

Collocations are words which are combined together and sound right to native English speakers. Any other combination may sound wrong and unnatural, such as "quick food". "Collocations are not words which we 'put together'. They co-occur naturally. Collocation is the way in which words co-occur in natural text in statistically significant ways" (Lewis, 2000, p. 132). Collocation is a "continuum that moves from compound words (second-hand), through

lexical chunks (bits and pieces), including idioms (out of the blue) and phrasal verbs (do up), to collocations of more or less fixedness (set a new world record)" (Thornbury, 2002, p.116).

Collocation is a problem is a big hurdle for foreign language learners and also an obstruction to the correct and efficient use of the target language. There might be reasons that each word can have many functions, or each word can appear in many collocation combinations.

Due to the fact that the majority of the problems often occur in grammar and vocabulary systems, we have chosen to compare collocations in Thai with that of Vietnamese for the coming of ASEAN Community (AC). Steps of comparison are applied in different aspects looking for the similarities and differences in both languages, and to point out errors in language learning in order to apply the results for the improvement and increase the efficiency of teaching and learning foreign languages.

Being native Vietnamese and Thai, we have observed plenty of collocations in both Vietnamese and Thai. Even though a number of Vietnamese collocations have completely different meanings from those of Thai, there are still a number of Thai collocations which have similar meaning with those of Vietnamese.

Vietnamese and Thai are within the same family of isolating language, and they share similar language structure. For instance, the collocations in Thai and Vietnamese can be translated word for word: "strong tea" *chè đăc²/cɛ² dak⁵* is an equivalent of "ชาเข้ม" /*chaa^M khem^F*. The structural order is the same: *chè² /jɛ²* is a noun meaning "ชา" /*chaa^M*, and *đăc² /dak⁵* is an adjective meaning "เข้ม" /*khem^F*. Thus, the structure of head noun and modifier [noun + an adjective] indicates that Thai and Vietnamese share this structure.

In addition, some Thai collocations convey meanings of shading such as colour terms "สีดำ" /*see^R dam^M* (the black color", which has many shades of meaning such as "ดำขลับ" /*dam^M khlap^L* / "ดำสวย" /*dam^M sa^L luay*, "ดำขำ" /*dam^M kham^R*, "ดำคล้ำ" /*dam^M khlaam^H*, "ดำปิ๊ด" /*dam^M bpit^H bpee^R*, "ดำปื้อ" /*dam^M bpee^R*, "ดำทะมึน" /*dam^M tha^H meun^F*, etc. The use of each depends on appropriate, correct, and natural word combinations such as "glossy black hair", "beautiful black hair", "black skin complexion", "darkened face", "pitch black water", "charcoal black water", "gloomy black sky", etc.

Similarly, there are words with similar meanings in the Vietnamese collocation, of which it is difficult to decide which word should be used such as *đen* /*dɛn⁰* meaning "the black color." There are six words which mean "the black color" such as *huyền* /*hwian²* (glossy black), *mun* /*mun⁰* (Ebony black), *mực* /*muɯk⁵* (the black ink), *ô* /*ʔo⁰* (shiny black), *hóng* /*hɔŋ¹* (smoky black), and *thâm* /*tham⁴* (gloomy black sky). These words are used differently, depending on the word used to combine, to make them correct, appropriate, and natural collocation forms of

use such as: *tóc huyền* /tɔk¹ hwian²/, *mắt huyền* /mat¹ hwian²/, *mèo mun* /mew² mun⁰/, *chó mực* /cɔ¹ muək⁵/, *ngựa ô* /ŋua⁵ ɔ⁰/, *bỏ hồng* /bɔ² hɔŋ¹/, *áo thâm* /ʔa:w¹ thəm⁴/, etc. It can be observed that *huyền* /hwian²/ is applied to the color of hair, or eyes, *mun* /mun⁰/ to a cat's hair, *mực* /muək⁵/ to a dog's hair, *ô* /ɔ⁰/ to a horse's hair, *hồng* /hɔŋ¹/ to the colors of smoke, ashes or coal, and *thâm* /thəm⁴/ to the color of a mouth and eyes, and clothes such as *trời thâm* /tɿj² thəm⁴/ (gloomy black sky), *áo* or *vải thâm* /ʔa:w¹ vaj¹ thəm⁴/ (blackish shirt /cloth), *môi thâm* /moi¹ thəm⁴/ (color of lips) and *mắt thâm* /mat¹ thəm⁴/ (color of eyes) etc.

The purpose of the study on collocations in Thai and Vietnamese is to establish a data set that acts as a guideline to those interested in learning Thai and Vietnamese so that they might be able to identify the correct, appropriate and natural words to suit the context. The data collection was conducted through translated books from Vietnamese into Thai from three different periods providing a variety of collocation samples. They are¹

- 1) "The collection of Nam Cao's short stories" *Nam Cao* (รวมเรื่องสั้นของนามกา) /ruam^M reuang^F san^F khaawng^R naam^M gao^M/; is a great and famous literary piece of Vietnam, written by Trần Hữu Tri, with Nam Cao as a pen name, and translated from Vietnamese into Thai by Viriyanun, in the Translation Project of the Neighboring Countries' Works of Sathiamkoses-Nakaprateep Foundation, first printed in 1986. There are 9 short stories in all. Nowadays, it is still used as a supplementary external reading book for classes in Vietnam.
- 2) "Memoire of Nhật Ký Đặng Thùy Trâm" *Nhật Ký Đặng Thùy Trâm* (บันทึกของดั่งถุ่ยเจิม) /ban^M theuk^H khaawng^R dang^L thuy^L jeerm^L/; first printed in 2005, is a book with total sales of more than 400,000 copies in Vietnam, written by Đặng Thùy Trâm, translated into Thai by Montira Rato, an expert in Vietnamese language at the Faculty of Arts, Chulalongkorn University.
- 3) "Let me have a return ticket to childhood one more time" *Cho tôi xin một vé đi tuổi thơ* (ขอตั๋วหนึ่งใบกลับไปสู่วัยเด็ก) /khaaw^R dtuaa^R neung^L bai^M glap^L bpai^M suu^L wai^M dek^L/; is a novel, winning the 2010 Asian Award for the best creative literary piece from Vietnam, written by Nguyễn Nhật Ánh, translated by Montira Rato and is also listed as one of the best sellers in Vietnam.

¹ Thai-language.com (TLC), P.8

In order to get the right, appropriate and sufficient sampling groups, the researchers searched for additional data from movies, dramas, music videos, news and website; examining grammar, many phrases, and common terms. Moreover, we conducted inquiries with people using Vietnamese language. Thus, in this study 548 Vietnamese collocations were found with comparisons to similar collocations in Thai.

According to, Art-ong Jumsai Na Ayudhyam (1997) the best technique for strengthening values is to control the five senses. He said that "We do not always see and hear reality, everything is in our mind, if our mind is dirty, everything we see and hear will be dirty, we interpret everything and give it meaning,". The key to good development is giving love and understanding and learning to think positively.

2. Objectives of the Research

The purpose of this study is to analyze and compare collocations in Thai and Vietnamese concerning sight, taste, smell, hearing, and touch; as follows:

- 1) to analyze and compare the components, meaning, and use of collocations
- 2) to analyze and compare the word formation of collocation
- 3) to analyze and compare types of collocation
- 4) to study the ratio of similarity and difference of collocation

3. Scope of the Research

The research is a comparative analysis for similarity and difference of the collocations in Thai and Vietnamese with the following scope:

The sampling of this study has made use of short stories translated from Vietnamese into Thai. In addition, additional data were collected from TV, radio, printed matters, together with interviews with Vietnamese-born Thais who are still using Vietnamese language in their daily life.

Procedures for our contrastive analysis follow Whitman Randal's approach (1970, p.191–197). They include four steps: 1) Description—to describe the characteristics of both languages, 2) Selection—to choose the format for comparison of similarities and differences, 3) Contrast—to contrast the chosen format, 4) Predict—to think of the possible problems that can happen in language learning from the foundation of those comparisons.

4. Basic Assumptions

4.1 The analysis and comparison of collocations in Thai and Vietnamese are done through the application of the study guidelines of collocations from the following:

1) Hausmann (1984): the components consist of the base and the collocator;

2) Benson, Benson and Ilson (1986): a study of grammatical collocation and lexical collocation, of which I have chosen to study the lexical collocation or word combinations;

3) Lewis (2000): classifications of collocations into three types, i.e., (1) weak collocations—words which can co-occur with other words; such as “พ่อครัว /phaaw^F khruaa^M/, พ่อบ้าน /phaaw^F baan^F/, พ่อพระ /phaaw^F phra^H/, พ่อบุญธรรม /phaaw^F boon^M tham^M/, etc.” (2) medium-strength collocations—different words with the same meaning can co-occur with other words; such as “กินข้าว /gin^M khaao^F/, ทานข้าว /thaan^M khaao^F/, รับประทานข้าว /rap^H bpra^L thaan^M khaao^F/, ปรอดหัว /bpuaat^L huaa^R/, ปรอดศีรษะ /bpuaat^L see^R sa^L/, etc. and (3) strong collocations—limited and specific words which cannot be changed such as “ขี้มูก” /khee^F muuk/ not “อุจจาระมูก” /oot^L jaa^M ra^H muuk/ and “แกงขี้เหล็ก” /gaaeng^M khee^F lek^L/ not “แกงอุจจาระเหล็ก” /gaaeng^M oot^L jaa^M ra^H lek^L/ etc.

4.2 The transcription system represents Thai language pronunciations and seeks only to represent the sounds of the original language. So, the pronunciation guide systems described here are all primarily transcription systems (although in a few cases, they do indicate some additional details about the spelling of the original word) from thai-language.com (TLC) such as คุณเก็บเสื้อไว้ไหน /^{M L F H R} which tone indication, there are 5 tones: 1) mid tone; x^M (เสียงสามัญ) /siiang^R saa^R man^M/, 2) low tone; x^L (เสียงเอก) /siiang^R aehk^L/, 3) falling tone; x^F (เสียงโท) /siiang^R tho:h^M/, 4) high tone; x^H (เสียงตรี) /siiang^R dtree^M/, 5) rising tone; x^R (เสียงจัตวา) /siiang^R jat^L dta^L waa^M/). The Thai-language.com enhanced phonemic transcription system is recommended for beginning students and is the default setting for transcriptions throughout that site. However, one can select an alternate Romanization scheme in the site control panel, including IPA, Paiboon, AUA, the Royal Thai General System, ISO 11940 transliteration, and others.

4.3 The International Phonetic Alphabet (IPA) is used for Vietnamese language pronunciations. There are two major standards, Hanoi and Saigon; represented here are the central dialects. There are 6 tones: 1) the *ngang* tone is level at around the mid level; x⁰ *ba* (three), 2) the *huyền* tone starts low falling (breathy); x¹ *ba* (lady) 3) the *sắc* tone starts as mid rising, tense; x² *ba* (governor) 4) the *hỏi* tone starts a mid falling (-rising),harsh; x³ *ba* (poison) 5) the *ngã* tone is mid rising glottalized; x⁴ *bã* (residue) 6) the *ngặng* tone starts midfalling, glottalized, short; x⁵ *ba* (at random)

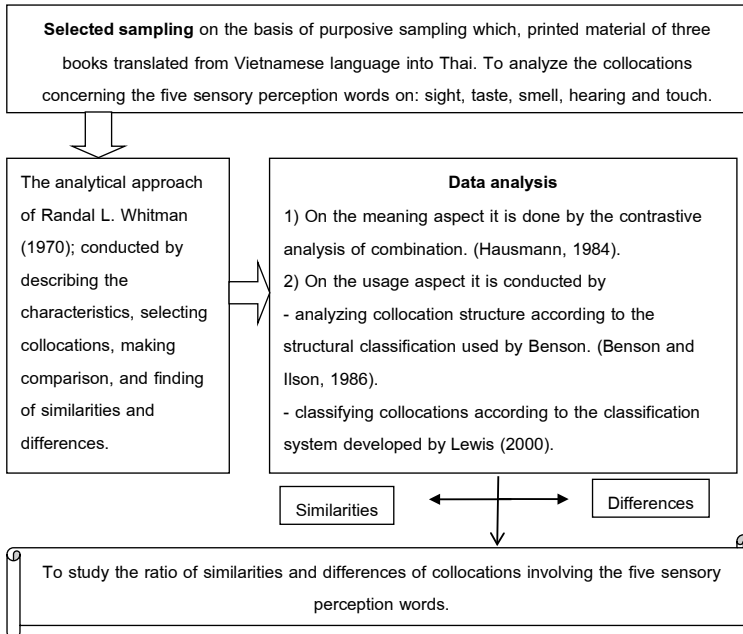
5. Significance of the Study

An analysis and a comparison of similarities and differences between the collocations of both languages will enable Thai and Vietnamese language learners to use the words correctly, appropriately, and naturally like the native speakers.

Moreover, the format for analysis and comparison of similarities and differences can be applied to the use of other languages.

6. Conceptual Framework

Figure1: Conceptual Framework



7. Research Methodology

The contrastive analysis of Thai and Vietnamese collocations was implemented in 3 steps as follows:

7.1 Sampling selection

We chose to study collocations of the five sensory perception words which are considered to be essential for the communication. If one has a good communication channel, communication capacity should be efficient and effective.

The five sensory perception words comprise sight, taste, smell, hearing, and touch. The knowledge of perception is composed of two forms: One is the abstract form which is the perception through the mind; and the other form, a concrete form. The five sensory perception words are considered important things that enable people to stimulate their good development and make good adaptations of themselves to the environments. These will sharpen their skills (<http://www.laorutis.dusit.ac.th/research10.php>).

In the Buddhist teaching, “ผัสสะ” /phat^l sa^l/ means contact, percussion and touch which cause feeling. “ผัสสะ” /phat^l sa^l/ is the interaction of three things which are internal senses (eyes, tongue, nose, ears, body and mind), external senses (sight, taste, smell, hearing, โภภุญัพพะ /pho:ht^R thap^H pha^H/, ธรรมารมณ /tham^M maa^M rohm^M/) and spirit (<http://th.wikipedia.org/wiki/>).

In marketing, Martin Lindstorm, a new generation marketer, who is one of the world's 100 influential persons, so much so that Time magazine praised him for what he has done and established through his “thought” into “a book” called *Brand sense: build powerful brand through touch, taste, smell, sight and sound* or in Thai version: “กลยุทธ์สร้างแบรนด์ทรงพลังด้วย รูป รส กลิ่น เสียงและสัมผัส” /gohn^M la^H yoot^H saang^F braaen^M sohng^M pha^H lang^M duay^F ruup^F roht^H glin^L siiang^R lae^H sam^R phat^L/ (<http://www.martinlindstrom.com/>).

Therefore, We selected the sampling “Base” (คำศัพท์ฐาน) /kham^M sap^L thaan^R/ involving the five sensory perception words which frequently occur in the three books by applying Hausmann's (1984) analytical stance in selecting the sampling as follows:

Table1: Base words concerning the five sensory perception words

Channel	Sense	Base
1. Sight	Eyes	thin, tall, fat, color and face
2. Taste	Tongue	mouth, speak, taste
3. Smell	Nose	smell
4. Hearing	Ears	sound
5. Touch	Body/Mind	heart, hard, rough, soft, lose, smile, cry, laugh

The words “กลิ่น” /glin^L/ (smell), “ใจ” /jai^M/ (heart), “ปาก” /bpaak^L/ (mouth), “พูด” /phuut^F/ (speak), “ผอม” /phaawm^R/ (thin), “สูง” /suung^R/ (tall), “อ้วน” /uaan^F/ (fat), “รส” /roht^H/ (taste), “สี” /see^R/ (colour), “แข็ง” /khaeng^R/ (hard), “หยาบ” /yaap^L/ (rough), “อ่อน” /aawn^L/ (soft), “เสีย” /siai^R/ (lose), “เสียง” /siiang^R/ (sound), “หน้า” /naa^F/ (face), “ยิ้ม” /yim^H/ (smile), “ร้องไห้” /raawng^H hai^F/ (cry), “หัวเราะ” /huuaa^R raw^H/ (laugh) are base which co-occur with collocator; for example,

- 1) The word group that co-occurs with “กลิ่น” /glin^L/ (smell) entails: กลิ่นหอม /glin^L haawm^R/, กลิ่นเหม็นเขียว/glin^L men^R khiaao^R/, กลิ่นสาบ /glin^L saap^L/, etc.
- 2) The word group that co-occurs with “ใจ” /jai^M/ (heart) entails: ใจกว้าง /jai^M gwaang^F/, ใจดำ /jai^M dam^M/, ใจแคบ /jai^M khaaep^F/, etc.
- 3) The word group that co-occurs with “ปาก” /bpaak^L/ (mouth) entails: ปากกล้ำ /bpaak^L glaa^F/, ปากแข็ง /bpaak^L khaeng^R/ ปากดี /bpaak^L dee^M/, etc.
- 4) The word group that co-occur with “พูด” /phuut^F/ (speak) entails: พูดหวาน /phuut^F waan^R/, พูดแข็ง /phuut^F khaeng^R/, พูดมัน /phuut^F man^M/, etc.
- 5) The word group that co-occurs with “ผอม” /phaawm^R/ (thin), “สูง” /suung^R/ (tall), “อ้วน” /uaan^F/ (fat) entails: ผอมแห้ง /phaawm^R haeng^F/, ผอมเกร็ง /phaawm^R greng^M/, ผอมบาง /phaawm^R baang^M/, สูงชัน /suung^R chan^M/, สูงโปร่ง /suung^R bpro:hng^L/, สูงระฟ้า /suung^R ra^H faa^H/, อ้วนท้วน /uaan^F thuaan^H/, อ้วนพี /uaan^F phee^M/, etc.
- 6) The word group that co-occurs with “รส” /roht^H/ (taste) entails: รสกลมกล่อม /roht^H glohm^M glaawm^L/, รสดี /roht^H dee^M/, รสฝาด /roht^H faat^L/, etc.
- 7) Word group that co-occurs with “สี” /see^R/ (colour) includes: ขาวนวล /khaao^R nuaan^M/ เหลืองอร่าม /leuaa^R ngaaw^M raam^F/, เขียวขจี /khiaao^R kha^L jee^M/, etc.
- 8) The word group that co-occurs with “แข็ง” /khaeng^R/ (hard), “หยาบ” /yaap^L/ (rough), “อ่อน” /aawn^L/ (soft) entails: แข็งใจ /khaeng^R jai^M/, แข็งแรง /khaeng^R raaeng^M/, หยาบคาย /yaap^L khaai^M/, หยาบช้า /yaap^L chaa^H/, อ่อนกว่า /aawn^L gwaat^L/, อ่อนข้อ /aawn^L khaaw^F/, etc.
- 9) The word group that co-occurs with “เสีย” /siai^R/ (lose) is: เสียใจ /siai^R jai^M/, เสียหาย /siai^R daai^M/, เสียเด็ก /siai^R dek^L/, เสียท่า /siai^R thaa^F/, เสียหน้าตา /siai^R nam^H dtaa^M/, เสียเปรียบ /siai^R bpriiap^L/, เสียรางวัล /siai^R rang^M wat^H/, เสียหน้า /siai^R naa^F/, etc.
- 10) The word group that co-occurs with “เสียง” /siiang^R/ (sound) includes: เสียงดู /siiang^R doo^L/, เสียงแหลม /siiang^R laaem^R/, เสียงทุ้ม /siiang^R thoom^H/, เสียงแตก /siiang^R dtaaek^L/, เสียงอ่อน /siiang^R aawn^L/, etc.
- 11) The word group that co-occurs with “หน้า” /naa^F/ (face) entails: หน้าจิต /naa^F jeuut^L/, หน้างอ /naa^F ngaaw^M/, หน้าเด็ก /naa^F dek^L/, etc.
- 12) The word group that co-occurs with “ยิ้ม” /yim^H/ (smile), “ร้องไห้” /raawng^H hai^F/ (cry), “หัวเราะ” /huua^R raw^H/ (laugh) is: ยิ้มหวาน /yim^H waan^R/, ยิ้มแย้มแจ่มใส /yim^H yaaem^H jaem^L sai^R/, ร้องไห้แง /raawng^H hai^F ngaaw^M ngaae^M/, ร้องไห้

สะอึกสะอื้น /raawng^H hai^F/ sa^L euk^L sa^L euun^F/, หัวเราะจ้า /huaa^R raw^H raa^F/,
หัวเราะร่วน /huaa^R raw^H ruan^F/, etc.

13) The word group that has specific usage such as ขมวดคิ้ว /kha^L muaat^L khiu^H/,
ขยับหู /kha^L yee^F boo^L ree^L/, ต่อมุหรี /dtaaw^L boo^L ree^L/, etc.

14) Idioms such as ใจปลาซิว /jai^M bplaa^M siu^M/, ปากหอยปากปู /bpaak^L haawy^R
bpaak^L bpuu^M/, เสียงนกเสียงกา /siiang^R nohk^H siiang^R gaa^M/, etc.

7.2 Data Collecting

We were interested in studying the collocations frequently used in these three books according to the study approach of Lewis (2000) which stated, “a collocation is the frequency of words that co-occur naturally, and it is the word combination that when one word starts, the other word normally follows or it can be expected to follow”. Base words of five sensory perceptions: sight, taste, smell, hearing, and touch are selected. For example, “smell”, “heart”, “mouth”, “speak”, “thin”, “fat”, “taste”, “color”, “hard”, “rough”, “soft”, “lose”, “sound”, “face”, “smile”, “cry” and “laugh” are used as base words which will collocate with other words.

Moreover, specific collocations and idioms will be included. These words reflect the culture, feelings, emotions, beliefs, and life style of the characters in the book, which can be applied in real life communication.

The data collection of collocations in Thai and those in Vietnamese concerning the five sensory perception words was done with two columns as follows:

Column 1: The data of collocations in Thai language that co-occur with base words of “กลิ่น” /glin^L/ (smell), “ใจ” /jai^M/ (heart), “ปาก” /bpaak^L/ (mouth), “พูด” /phuut^F/ (speak), “ผอม” /phaawm^R/ (thin), “สูง” /suung^R/ (tall), “อ้วน” /uaan^F/ (fat), “รส” /roht^H/ (taste), “สี” /see^R/ (colour), “แข็ง” /khaeng^R/ (hard), “หยาบ” /yaap^L/ (rough), “อ่อน” /aawn^L/ (soft), “เสีย” /siaa^R/ (lose), “เสียง” /siiang^R/ (sound), “หน้า” /naa^F/ (face), “ยิ้ม” /yim^H/ (smile), “ร้องไห้” /raawng^H hai^F/ (cry), “หัวเราะ” /huaa^R raw^H/ (laugh) including the specific words and idioms

Column 2: Making comparisons between collocations in Vietnamese and those in Thai.

From the data collecting of collocations concerning the five sensory perception words in the three books, I found that there were 14 groups of frequently used base words which co-occurred with 338 collocators, and there were a total of 798 co-occurrences. The detailed information is specified in Table 2 below.

Table2: The number of sampling groups concerning the five sensory perception words

Base words	Number of words	Number of times
1. "กลิ่น" (smell)	19	21
2. "ใจ" (heart)	10	24
3. "ปาก" (mouth)	15	15
4. "พูด" (speak)	49	83
5. รูปร่าง : "ผอม สูง อ้วน" (Figure: thin, tall, fat)	22	42
6. "รส" (taste)	13	22
7. "สี" (color)	39	78
8. สัมผัส : "แข็ง หยาบ อ่อน" (Touch: hard, rough, soft)	21	52
9. "เสีย" (lose)	12	81
10. "เสียง" (sound)	20	165
11. "หน้า" (face)	38	86
12. อารมณ์: "ยิ้ม ร้องไห้ หัวเราะ" (Emotion: smile, cry, laugh)	34	75
13. คำเฉพาะ (specific words)	26	34
14. สำนวน (idioms)	20	20
Total	338	798

One observation from Table 2 is that the base word concerning of "รูปร่าง" /ruup^F raang^F/ (Figure) comprises the words "ผอม" /phaawm^R/ (thin), "สูง" /suung^F/ (tall), "อ้วน" /uaan^F/ (fat), "สัมผัส" /sam^R phat^L/ (Touch) comprises the words "แข็ง" /khaeng^R/ (hard), "หยาบ" /yaap^L/ (rough), "อ่อน" /aawn^L/ (soft), "อารมณ์" /aa^M rohm^M/ (Emotion) comprises the words "ยิ้ม" /yim^H/ (smile), "ร้องไห้" /raawng^H hai^F/ (cry), "หัวเราะ" /huua^R raw^H/ (laugh). This study found 548 collocations. The detailed information was as follows is identified in Table 3.

Table3: The number of sampling groups of collocations

Base words	Number of words
1. "กลิ่น" (smell)	24
2. "ใจ" (heart)	48
3. "ปาก" (mouth)	38
4. "พูด" (speak)	55
5. รูปร่าง : "ผอม สูง อ้วน" (Figure: thin, tall, fat)	35
6. "รส" (taste)	16
7. "สี" (color)	71
8. สัมผัส : "แข็ง หยาบ อ่อน" (Touch: hard, rough, soft)	37
9. "เสีย" (lose)	28
10. "เสียง" (sound)	48
11. "หน้า" (face)	56
12. อารมณ์: "ยิ้ม ร้องไห้ หัวเราะ" (Emotion: smile, cry, laugh)	36
13. คำเฉพาะ (specific words)	28
14. สำนวน (idioms)	28
Total	548

7.3 Data Analysis

In this study document research was used by researching the related literature concerning collocations focusing on the components, meanings, usages, structure and classifications of collocations including research documents related to contrastive analysis consulted for their guidelines to help with research implementation. The contrastive analysis between collocations in Thai and those in Vietnamese was divided into three parts.

Part 1: The contrastive analysis on the components, meanings, and usages of collocations concerning the five sensory perception words. Hausmann's (1984) study was applied in the analysis.

Part 2: The contrastive analysis on the lexical collocations concerning the five sensory perception words was conducted by using Benson and Ilson (1986) in the analysis.

Part 3: The contrastive analysis on classifications of collocations concerning the five sensory perception words was conducted by applying Lewis (2000) in the analysis.

In concluding our findings, we chose a descriptive approach to report on the meanings and usages of collocations in Thai and those in Vietnamese together with the contrastive analysis of the components, structures and classifications of collocations in order to bring out the ratio of similarities and differences by counting the number of collocations and calculating for percentage value.

8. Research results and Discussions

The research result from Part 1: There were a total of 548 collocators studied, it was found that the percentages of components, meanings and usages in Thai and Vietnamese have more similarities than differences. In other words, there were there were 404 similar collocations components or 73.72 in percentages, and there were 144 different collocations or 26.28 in percentages. Moreover the comparison between Thai and Vietnamese, it was found that the base word “สี” /see^R/ has the most number of similarities which are 66 words or 12.04 in percentages. For the base word “พูด” /phuut^F/ it was the secondary rank in similarities. There were 52 words or 9.49 per cent, and for the base word “ใจ” /jai^M/ there were the least number of similarities. There were 8 words or 1.46 percent.

On the other hand, it was found that the base word “ใจ” /jai^M/ has the most number of differences which are 40 words or 7.3%. For the base word “ปาก” /bpaak^L/ it was the secondary rank in differences. There were 28 words or 5.11%, and for the base words “กลิ่น” /glin^L/, “รูปร่าง” /ruup^F raang^F/, “อารมณ์” /aa^M rohm^M/, and “สีนวน” /sam^R nuaan^M/, had the least number of differences. There were 2 words or 0.37%.

From part1 I found that some base words in Vietnamese language convey the same meaning and can be substituted as four collocators that co-occur with the base word “ใจ” /jai^M/

(heart) in Vietnamese: 1) *tâm* /tɛm⁰/ (center) 2) *lòng* /lɔŋ²/ (heart, barrel, bore, and bosom) 3) *bụng* /buŋ²/ (stomach, belly, abdomen, maw, and paunch) 4) *gan* /ga n⁰/ (liver, daring and near to). All 4 words when co-occurring with other words have the meaning of “ใจ” /jai^M/, but depending on the appropriateness of which word to collocate with it. However, some base words may have other meanings when they are by themselves.

The research result from Part 2: There were a total of 548 collocations and I found that there are a total of 8 collocation structures: 1) a noun with a noun, 2) a noun with an adjective, 3) a noun with a verb, 4) a verb with a verb, 5) a verb with an adverb, 6) a verb with a noun, 7) an adjective with an adjective and 8) an adjective with a noun. I found that there were more similarities than differences in the collocation structures in Thai and Vietnamese. There were 422 similar collocation structures. It was 77.01%, and there were 126 different collocation structures which were 22.99%. The collocation structure that has the most similarities is the structure of an adjective with an adjective. There are 91 words or 16.61%. The secondary rank in similarity is the structure of a verb with a noun. There are 60 words or 10.95%, and for the collocation structure of an adjective with a noun that the least similarity, there are only 12 words or 2.19%.

On the other hand, it was found that the collocation structure that has the most differences is the structure of a noun with an adjective. There are 47 words or 8.58%. The secondary rank in differences is the structure of a noun with a verb. There are 28 words or 5.11%, and for the collocation structure of a verb with a verb and a verb with an adverb that the least differences, there are only 2 words or 0.37%.

From part2 found that some structures in Vietnamese language, the base word follows the collocators such as the word “ปากมาก” /bpaak^L maak^F/ in Thai matches that in Vietnamese *lắm mồm* /lom¹ mom²/ and the word “พูดน้อย” /phuut^F naawy^H/ in Thai matches that in Vietnamese *ít nói* /it¹ noj¹/.

The research result from Part 3: in total 548 collocations I found that there were collocations used in Thai and Vietnamese which were similar and different. In the contrastive analysis there were totally 3 types of collocations. The similarities of collocation types that had match types such as Thai—Weak Collocations with Vietnamese—Weak Collocations, Thai—Strong Collocations with Vietnamese—Strong Collocations and Thai—Medium-strength Collocations Vietnamese—Medium-strength Collocations.

From this study it was found that there were more similarities than differences of collocation types. There were 372 of the same type of collocations or 67.88%, and there were 176 of different type of collocations or 32.12%. Furthermore, the collocations in Thai and Vietnamese I found that the classification of collocations that had the most similarities was

Weak Collocations— words which can co-occur with other words. There were 327 collocations or 59.67%. The second most similarities occurred with Strong Collocations—limited and specific words which cannot be changed; with 44 collocations or 8.03%. For the Medium-strength Collocations—different words with the same meaning can co-occur with other words there was only one similarity or 0.18%.

In the constrastive analysis I found that differences in that there was no complementary match types such as Thai—Weak Collocations but Vietnamese—Strong or Medium-strength Collocations. In this study of the classification of collocations I found that the most differences occurring with Weak Collocations There were 171 collocations or 31.21%. The second most differences was Medium-strength Collocations, with 4 collocations or 0.73%. For the Strong Collocations there was only one difference or 0.18%.

From part3 I found that some types in Thai and Vietnamese are different such as in Thai Medium-strength Collocations—different words with the same meaning can co-occur with other words; “กลิ้งเต๋า” /glin^L dtao^L/ or “กลิ้งรั๊กแร่” /glin^L rak^H rae^H/ is grouped in Medium-strength collocations because it is a collocation that can be substituted but in Vietnamese it is grouped in Strong Collocations because it can be only one word *mũi hôi nách* /muj² hoj⁰ nat²/.

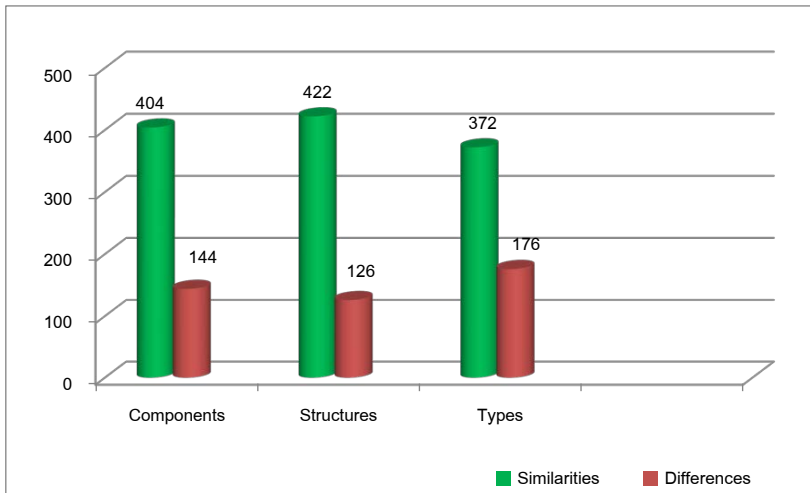
This study aimed to carry out a contrastive analysis to find the similarities and the differences between collocations in Thai and Vietnamese that convey the five sensory perceptions are sight, taste, smell, hear, and touch.

The study results were intended to be used as guidelines for making correct, appropriate and natural word choices by counting and calculating for percentage value. From the contrastive analysis of a total of 548 collocations occurring in Thai and Vietnamese, I found that there was a ratio of more similarities than there were differences.

In detail, there were a total of 14 groups of collocation components. There were 404 collocations or 73.72% that had similarities in components, and there were 144 collocations or 26.28% with different components.

There were totally 8 groups of collocation structures. There were 422 collocation structures or 77.01% with similar structures while there were 126 collocation with different structures or 22.99%.

Chart 1: The ratio between the similarities and differences of collocations of 3 parts of analysis



9. Conclusions

From the contrastive analysis of 548 collocations occurring in Thai language and Vietnamese language, it was found that there were more similarities than differences. More than 70% of the similarities were in components and structures and more than 60% of the similarities were in classifications.

From this study it can be concluded that in the Vietnamese language, there are many base words that convey the same meaning and can be substituted. However, some base words may have other meanings when they are by themselves.

In addition, some collocation structures are different because there are some phrases in the Vietnamese language where the base words follow the collocators due to the influence of Chinese Syntax. However, the classifications in Thai and Vietnamese are similar although there are many more base words in Vietnamese.

References

- Bahns, J. and Eldaw, M. (1993). Should we teach EFL students collocations? *System* 21(1), 101-114.
- Benson, M., Benson, E. and Ilson, R. (1997). *The BBI Dictionary of English Word Combinations*. Amsterdam: John Benjamins.
- Cowie, A. P. and Lewis, J. W. (1974). *The Oxford Advanced Learner's Dictionary of Current English* (3rd Edition). Oxford: Oxford University Press.
- Duong, Thanh Bình. (1971). *A Tagmemic Comparison of the Structure of English and Vietnamese Sentences*. The Hague: Mouton.
- Đài, Xuân Ninh. (1978). *Hoạt Động của Từ tiếng Việt* [Functions of Vietnamese Lexemes], Hanoi: Khoa-học Xã-hội.
- Đinh, Văn Đức. (1986). *Ngữ pháp Tiếng Việt: Từ loại* [Vietnamese Syntax: Word Classes]. Hanoi: Đại-học và Trung-học Chuyên-nghiệp.
- Fontenelle, T. (1994) "What on earth are collocations?" Retrieved in Aug 5, 2014, from www.ugr.es/~inped/.../What_on_Earth_are_Collocations.pdf
- Hausmann, Franz Josef. (1989). "Le dictionnaire de collocations." In Hausmann, F. J. et al. (eds.). *Wörterbücher, Dictionaries, Dictionnaires, erster Teilband* (pp.1010 – 1019). Berlin: de Gruyter.
- Herbst, T. (1996). What are collocations: sandy beaches or false teeth? *English Studies* 4, 379-393.
- Hill, J. (2000). Revising priorities: From grammatical failure to collocational success. In Lewis, M. (Ed.). *Teaching collocation: Further development in the lexical approach* (pp. 47-69). Oxford: Oxford University Press.
- Hill, J., Lewis, M., and Lewis, M. (2000). Classroom strategies, activities and exercises. In Lewis, M. (Ed.). *Teaching Collocations: Further Developments in the Lexical Approach* (pp. 88-117). Boston: Thomson.
- James, C. (1980). *Contrastive Analysis*. London: Longman.
- Jumsai Na Ayudhyam, Art-ong. (1997). *Five Human Values and Human Excellence Integration of Human Values in Mathematics and Sciences*. Bangkok: The Institute of Sathya Sai Education.
- Lê, Khả Kế. (1991). *Từ Điển Anh-Việt* [English-Vietnamese Dictionary]. Hanoi: Thành phố Hồ Chí Minh.
- Lennon. (1996). *Lexical Collocations and their Relation to the Online Writing*. Retrieved Nov, 22, 2014, from <https://e-flt.nus.edu.sg/v4n22007/hsu.pdf?q..collocations..to>.

- Lewis, M. (Ed.). (2000). *Teaching Collocations: Further Developments in the Lexical Approach*. Boston: Thomson.
- Lindstorm, M. (2014). *Brand sense: build powerful brand through touch, taste, smell, sight and sound*. Retrieved Oct, 25, 2014, from <http://www.martinlindstrom.com/>
- McCarthy, M. and O' Dell, F. (2005). *English Collocations in use* Retrieved Jan, 22, 2014, from <http://creativecommons.org/publicdomain/mark/1.0/>
- Moon, R. (1997). Vocabulary connections: multi-word items in English. In Schmitt, N. and McCarthy, M. (eds.). *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press.
- Moon, R. (1998). *Fixed Expressions and Idioms in English*. New York: Oxford University Press.
- Nguyễn, Đình-Hòa. (1972). 'Vietnamese categories of result, direction and orientation,' In Smith, M. E. (ed). *Studies in Linguistics: Essays in honor of George L. Trager*. The Hague: Mouton.
- Nguyễn, Đình-Hòa. (1975). 'Lexical and syntactic borrowing in modern Vietnamese'. In van Opphen, H. H. (ed). *Linguistic Borrowing*. Austin: The U. of Texas Center for Asian Studies.
- Sinclair, J. M. (2004). The search for units of meaning. *Trust the Text: Language, Corpus and Discourse* (pp. 24-48). London: Routledge.
- Thai-language.com. (2000). Retrieved Feb 12, 2014, from <http://www.thai-language.com/forums/t/linguistics/transliteration/t12511>.
- Trần Văn Trọng. (2004). *Cẩm Nang Đàm Thoại Thông Dụng Việt Nam-Thái Lan Nhà Xuất Bản Văn Hóa Thông Tin*
- Trinh, F. N. (2002). *English and Vietnamese Collocations: A Contrastive Analysis*. Retrieved Jan, 15, 2014, from http://www.researchonline.mq.edu.au/vital/access/manager/Repository/mq:9610?f0=sm_creator%.
- Whitman, R. L. (1970). Contrastive Analysis: Problems and Procedures. *Language Learning Research Club*. Michigan: University of Michigan.
- Wikipedia, the free encyclopedia. (2014). Retrieved Apr, 22, 2014, from <https://en.wikipedia.org/wiki/Sense>
- บรรจง วัชรประภา. (2543). *บทเรียนสนทนา ภาษาเวียดนาม-ไทยด้วยตัวเอง*. กรุงเทพฯ: โรงเรียนดอนบอสโก. ราชบัณฑิตยสถาน. (2546). *พจนานุกรมฉบับราชบัณฑิตยสถาน*. กรุงเทพฯ: นานมีบุ๊คพับลิเคชั่น.
- วงศ์ วรรณพิเชฐ. (2542). *พจนานุกรมตัวอย่าง [English by Example: A Dictionary of English]*. กรุงเทพฯ: บริษัทไทยเว็ส พับลิเคชั่นส์ จำกัด.
- สถาบันศึกษา ภาคเอเชียอาคเนย์. (2533). *พจนานุกรมไทย-เวียดนาม*. กรุงเทพฯ: สังคมศาสตร์.
- เหิงยง จี ธง. (2001). *บทเรียนสนทนา ภาษาเวียดนาม-ไทยด้วยตัวเอง*. กรุงเทพฯ: บริษัทกิววิเวลส.