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

การศึกษาการเปรียบเทียบคําปรากฏร่วมในภาษาไทยและภาษาเวียดนาม
กานตชนิตา คุณธรรมสกุล, อัจฉรา วงศโสธร และสุวรีย ยอดฉิม

บทคัดยอ
การศึกษาระดับนี้ มีวัตถุประสงคเพื่อวิเคราะหปรศึกษาการเปรียบเทียบความเหมือนและความแตกตางของคําปรากฏร่วมวิถีในภาษาไทยและภาษาเวียดนามที่เกี่ยวกับประสาทสัมผัส 5 คือ รูป รส กลิ่น เสียง สัมผัส โดยการนับจํานวนและคิดเปนรอยละ ขอมูลมีจํานวนทั้งสิ้น 548 คําปรากฏร่วมวิถี โดยแบงการวิเคราะหออกเปน 3 สวน คือ สวนที่ 1 ศึกษาสวนประกอบคําปรากฏร่วมวิถีในภาษาไทยและภาษาเวียดนามที่เกี่ยวกับประสาทสัมผัส 5 คือ รูป รส กลิ่น เสียง สัมผัส โดยการนับจํานวนและคิดเปนรอยละ ซึ่งมีจํานวนทั้งสิ้น 548 คําปรากฏร่วมวิถี โดยแบงการวิเคราะหออกเปน 3 สวน คือ สวนที่ 2 ศึกษาโครงสรางคําปรากฏร่วมวิถีในภาษาไทยและภาษาเวียดนามที่เกี่ยวกับประสาทสัมผัส 5 คือ รูป รส กลิ่น เสียง สัมผัส โดยการนับจํานวนและคิดเปนรอยละ ซึ่งมีจํานวนทั้งสิ้น 548 คําปรากฏร่วมวิถี โดยแบงการวิเคราะหออกเปน 3 สวน คือ สวนที่ 3 ศึกษาโครงสรางคําปรากฏร่วมวิถีในภาษาไทยและภาษาเวียดนามที่เกี่ยวกับประสาทสัมผัส 5 คือ รูป รส กลิ่น เสียง สัมผัส โดยการนับจํานวนและคิดเปนรอยละ ซึ่งมีจํานวนทั้งสิ้น 548 คําปรากฏร่วมวิถี โดยแบงการวิเคราะหออกเปน 3 สวน คือ สวนที่ 1 ศึกษาสวนประกอบคําปรากฏร่วมวิถีในภาษาไทยและภาษาเวียดนามที่เกี่ยวกับประสาทสัมผัส 5 คือ รูป รส กลิ่น เสียง สัมผัส โดยการนับจํานวนและคิดเปนรอยละ ซึ่งมีจํานวนทั้งสิ้น 548 คําปรากฏร่วมวิถี โดยแบงการวิเคราะหออกเปน 3 สวน คือ สวนที่ 2 ศึกษาโครงสรางคําปรากฏร่วมวิถีในภาษาไทยและภาษาเวียดนามที่เกี่ยวกับประสาทสัมผัส 5 คือ รูป รส กลิ่น เสียง สัมผัส โดยการนับจํานวนและคิดเปนรอยละ ซึ่งมีจํานวนทั้งสิ้น 548 คําปรากฏร่วมวิถี โดยแบงการวิเคราะหออกเปน 3 สวน คือ สวนที่ 3 ศึกษาโครงสรางคําปรากฏร่วมวิถีในภาษาไทยและภาษาเวียดนามที่เกี่ยวกับประสาทสัมผัส 5 คือ รูป รส กลิ่น เสียง สัมผัส โดยการนับจํานวนและคิดเปนรอยละ ซึ่งมีจํานวนทั้งสิ้น 548 คําปรากฏร่วมวิถี โดยแบงการวิเคราะหออกเปน 3 สวน คือ สวนที่ 1 ศึกษาสวนประกอบคําปรากฏร่วมวิถีในภาษาไทยและภาษาเวียดนามที่เกี่ยวกับประสาทสัมผัส 5 คือ รูป รส กลิ่น เสียง สัมผัส โดยการนับจํานวนและคิดเปนรอยละ ซึ่งมีจํานวนทั้งสิ้น 548 คําปรากฏร่วมวิถี โดยแบงการวิเคราะหออกเปน 3 สวน คือ สวนที่ 2 ศึกษาโครงสรางคําปรากฏร่วมวิถีในภาษาไทยและภาษาเวียดนามที่เกี่ยวกับประสาทสัมผัส 5 คือ รูป ร
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” (คำประดิษฐ์เฉพาะ) /khamM bpraM gohtL ruaamF champho/ or (สหบท) /saL haL bohtL/. However, for this study, the researcher uses “collocation” (คำปรากฏร่วมวิธี) /khamM bpraM gohtL ruaamF 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 /fohnR dtohk nak/ (The rain falls heavily.)
2) แดดออกจ๊า /daaet aawk jaa/ (The sun shines brightly.)
3) ลมพัดแรง /lohm phat raaeng/ (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/ for the collocators and adverb “heavily” (หนัก) /nak/. They will not use “shine” for the verb and “brightly” (จ้า) /jaa/ 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/ in Thai, the verb that follows has to be “fall” (ตก) /dtohk/. The words that follow might be “an adjective- persistent” (ครึ่ม) /khreum/, “another verb – come soon” (มาทีด) /ma meuut/ or “a prepositional phrase – from the North” (จากทางเหนือ) /jaak thaang neua/. Likewise, the verb “fall” (ตก) /dtohk/ might follow other headwords, not only “the rain” (ฝน) /fohn/; for example, สีตก /see dtohk/ (The color is washed out.), or หุ้นตก /hoon dtohk/ (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 đặc ²/ce₂ dak⁵/ is an equivalent of “ชามเข้ม” /chaaM khem⁵/. The structural order is the same: chè /ce₂/ is a noun meaning “ชา” /chaaM/, and đặc /dak⁵/ is an adjective meaning “เข้ม” /khem⁵/. 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” /seeR dam⁵/ (the black color), which has many shades of meaning such as “ดำลั่น” /dam⁵ khlap⁵/ “ดำลักษ์” /dam⁵ la⁵ luay/, “ดำierarchical” /dam⁵ kham⁵/, “ดำกล้า” /dam⁵ khaam⁵/, “ดำบิด” /dam⁵ bpit⁵/ bpee⁵/, “ดำบี” /dam⁵ bpee⁵/, “ดำมะมี่” /dam⁵ thaH meuun⁵/, 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 den ¹/ddam⁰/ meaning “the black color.” There are six words which mean “the black color” such as huyền /hwian²/ (glossy black), mun /mun⁰/ (Ebony black), mực /muk⁵/ (the black ink), ô /ı⁰/ (shiny black), hồng /hɒŋ⁴/ (smoky black), and thâm /tham⁵/ (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: ตอค หุ่ยน /tɔk hwian2/, มตหุ่ยน /mat hwian2/, หมวม /mew2 mun0/, ชอ มุค /co1 muk5/, นกข ö /ŋua5 ṭ0/, บะหงส /bo2 ṭ0ŋ1/, อะทัม /ʔaw1 tham4/, etc. It can be observed that หุ่ยน /hwian2/ is applied to the color of hair, or eyes, มว/ mun0/ to a cat’s hair, มุค /mu1k5/ to a dog’s hair, ṭö /ṭ0/ to a horse’s hair, หงส /hong/ to the colors of smoke, ashes or coal, and ทัม /tham4/ to the color of a mouth and eyes, and clothes such as ตรี ทัม /trɔi tham4/ (gloomy black sky), อะ or วัย ทัม /aw1/ vaj1 tham4/ (blackish shirt/cloth),  môi ทัม /moi1 tham4/ (color of lips) and มัท ทัม /mat1 tham4/ (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 are

1) “The collection of Nam Cao’s short stories” Nam Cao (รวมเรื่องสั้นของนางก้าว) /ruamM reuuangF sanF khaawngR naamM gaaoM/; is a great and famous literary piece of Vietnam, written by Trần Hữu Trí, 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 Sathiarnkoses-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 (บันทึกของดั่งถุ ีเจิ่ม) /banM theukH khaawngR dangL thuyL jeermL/; 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 (ขอตั๋วหนึ่งใบกลับไปสูวัยเด็ก) /khaawR dtuaaR neung1 baiM glap1 bpa1 suu1 waiM dek1/; 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 /phaawF khruaM/, พ่อบ้าน /phaawF baanM/, พ่อพระ /phaawF phraM/, พ่อพุทธธรรม /phaawF boonM thamM/, etc.” (2) medium-strength collocations—different words with the same meaning can co-occur with other words; such as “กินข้าว /ginM khaaoF/, ทานข้าว /thaanM khaaoF/, รับประทานข้าว /rapM bpraL thaanM khaaoF/, บูรพา /buaatL seeR saL/, etc. and (3) strong collocations—limited and specific words which cannot be changed such as “ขี้ผึ้ง /kheeF muuk/ not “อุจจาระผึ้ง” /ootF jaaM muuk/ and “แกงขี้เหล็ก” /gaaengF kheeF lekL/ not “แกงอุจจาระเหล็ก”, /gaaengM ootL jaaM raM lekL/ 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 คุณเก็บเสื้อไวไหน /M L F H R/ which tone indication, there are 5 tones: 1) mid tone; χ(M) (เสียงสามัญ) /siiangR saaR manM/, 2) low tone; χ(ฮือแล้ว) /siiangR aehkL/, 3) falling tone; χ(F) (เสียงโท) /siiangR tho:hM/, 4) high tone; χ(H) (เสียงตรี) /siiangR dtreeM/, 5) rising tone; χ(R) (เสียงจัตวา) /siiangR jatL dtaL waaM/. 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; χ⁰ ba (three), 2) the huyén tone starts low falling (breathy); χ¹ ba (lady) 3) the sâc tone starts as mid rising, tense; χ² ba (governor) 4) the hoi tone starts a mid falling (-rising),harsh; χ³ ba (poison) 5) the ngâ tone is mid rising glottalized; χ⁴ b a (residue) 6) the nãng tone starts midfalling, glottalized, short; χ⁵ 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

![Conceptual Framework](image)

**Figure 1**: Conceptual Framework

**Selected sampling** on the basis of purposive sampling which, printed material of three books translated from Vietnamese language into Thai. To analyze the collocations concerning the five sensory perception words on: sight, taste, smell, hearing and touch.

**Data analysis**

1) On the meaning aspect it is done by the contrastive analysis of combination. (Hausmann, 1984).

2) On the usage aspect it is conducted by

- analyzing collocation structure according to the structural classification used by Benson. (Benson and Ilson, 1986).
- classifying collocations according to the classification system developed by Lewis (2000).

To study the ratio of similarities and differences of collocations involving the five sensory perception words.

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, "ผัสสสะ" /phatL saL/ means contact, percussion and touch which cause feeling. "ผัสสสะ" /phatL saL/ is the interaction of three things which are internal senses (eyes, tongue, nose, ears, body and mind), external senses (sight, taste, smell, hearing, โผฏฐัพพะ /pho:htR thapH phaH/, ธรรมารมณ์ /thamM maaM rohmM/) and spirit (http://th.wikipedia.org/wiki/).

In marketing, Martin Lindstrom, 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: "กลยุทธสร้างแบรนด์ทรงพลังด้วยรูป รส กลิ่น เสียงและสัมผัส" /gohnM laH yootH saangF braaenM sohngM phaH laungM duayF rohtH glinL siiangR laeH samR phatL/ (http://www.martinlindstrom.com/).

Therefore, We selected the sampling “Base” (คำศัพท์ฐาน) /khamM sapL thaan/ 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:

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

1) The word group that co-occurs with "กลิ่น" /glin²/ (smell) entails: กลิ่นหอม /glin¹ haawm²/, กลิ่นเหม็นเขียว /glin¹ men³ khiaao³/, กลิ่นสาบ /glin¹ saap²/, etc.

2) The word group that co-occurs with "ใจ" /jai¹/ (heart) entails: ใจ กว่า /jai¹ gwaang²/, ใจดำ /jai¹ dam²/, ใจดี /jai¹ dee²/, etc.

3) The word group that co-occurs with "ปาก" /bpaak²/ (mouth) entails: ปากกล้า /bpaak² glaa²/, ปากแข็ง /bpaak² khaeng²/, ปากดี /bpaak² dee²/, etc.

4) The word group that co-occurs with "พูด" /phuut²/ (speak) entails: พูดหัวใจ /phuut² waan²/, พูดแข็ง /phuut² khaeng²/, พูดมัน /phuut² man²/, etc.

5) The word group that co-occurs with "ผอม" /phaawm²/ (thin), "สูง" /suung²/ (tall), "อ้วน" /uaan²/ (fat) entails: ผอมห่าง /phaawm² haaeng²/, ผอมกระชับ /phaawm² greg⁴/, ผอมบาง /phaawm² baang²/, สูงชัน /suung²  chan²/, สูงโปร่ง /suung² bpro:hng²/, ด้านพื้น /uaan²  thuan²/, ด้านพื้น /uaan² phee²/, etc.

6) The word group that co-occurs with "รส" /roht⁴/ (taste) entails: รสกลมกลม /roht⁴ glohm² glohm²/, รสดี /roht⁴ dee²/, รสฝาด /roht⁴ faat²/, etc.

7) Word group that co-occurs with "สี" /see²/ (colour) includes: ขาวนวล /khaao³ nuaan²/, เหลืองอร่าม /leuua³ ngaaw² raaam²/, เขียวขจี /khiaao³ kha³ jee²/, etc.

8) The word group that co-occurs with "แข็ง" /khaeng²/ (hard), "หยาบ" /yaap²/ (rough), "อ่อน" /aawn²/ (soft) entails: แข็งใจ /khaeng² jai¹/, แข็งแรง /khaeng² raaw²/, หยาบคาย /yaap² khaai²/, หยาบชื้น /yaap² chaa²/, อ่อนกว่า /aawn² gwaa²/, อ่อนนิ่ม /aawn² khaaw²/, etc.

9) The word group that co-occurs with "เสีย" /siia²/ (lose) is: เสียใจ /siia² jai¹/, เสียตา /siia² dtaa²/, เสียร่าง /siia² rangM²/, เสียประทุน /siia² rangM² watH²/, เสียตา /siia² naa²/, etc.

10) The word group that co-occurs with "เสียง" /siiang²/ (sound) includes: เสียงดุ /siiang² doo²/, เสียงแหลม /siiang² laaem²/, เสียงทุ่ม /siiang² thoom²/, เสียงแตก /siiang² dtaaek²/, เสียงของ /siiang² aawn²/, etc.

11) The word group that co-occurs with "หน้า" /naa²/ (face) entails: หน้าจิ้ม /naa² jeuut²/, หน้าตัก /naa² ngaaw²/, หน้าเด็ก /naa² dek²/, etc.

12) The word group that co-occurs with "ยิ้ม" /yimH²/ (smile), "ร้องไห้" /raawng⁴ hai²/ (cry), "หัวเราะ" /huaaR rawH²/ (laugh) is: ยิ้มหวาน /yimH² waan²/, ยิ้มแย้มแจ่มใส /yimH² yaaem⁴ jaem³ sai³/, ร้องไห้ข้องใจ /raawng⁴ hai² ngaaw² ngaaw²/, ร้องไห้
The word group that has specific usage such as ขมวดคิ้ว /khaL muaatL khiuH/, ขยี้บุหรี่ /khaL yeeF booL reeL/, ต่อบุหรี่ /dtaaawL booL reeL/, etc.

Idioms such as ใจปลาซิว /jaiM bplaaM siuM/, ปากหอยปากปู /bpaakL haawyR bpaakL bpnM/, เสียงนกเสียงกา /siiangR siiangR gaaM/, 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 /glinL/ (smell), ใจ /jaiM/ (heart), ปาก /bpaakL/ (mouth), พูด /phuutF/ (speak), ผอม /phaawmR/ (thin), สูง /suungR/ (tall), อ้วน /uaanF/ (fat), รูส /rohtH/ (taste), ผิว /seeR/ (colour), แข็ง /khaengR/ (hard), หยาบ /yaapL/ (rough), อ่อน /aawnF/ (soft), สี /seeR/ (lose), เสียง /siiangR/ (sound), หน้า /naaF/ (face), ยิ้ม /yimH/ (smile), ร้องไห้ /raawngH haiF/ (cry), หัวเราะ /huaaR rawH/ (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.
Table 2: The number of sampling groups concerning the five sensory perception words

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

One observation from Table 2 is that the base word concerning of "รูปร่าง" /ruupF raangF/ (Figure) comprises the words "ผอม" /phaawmR/ (thin), "สูง" /suungR/ (tall), "อ้วน" /uaanF/ (fat), "สัมผัส" /samR phatL/ (Touch) comprises the words "แข็ง" /khaengR/ (hard), "หยาบ" /yaapL/ (rough), "อ่อน" /aawnL/ (soft), "อารมณ์" /aaM rohmM/ (Emotion) comprises the words "ยิ้ม" /yimH/ (smile), "ร้องไห้" /raawngH haiF/ (cry), "หัวเราะ" /huaaR rawH/ (laugh). This study found 548 collocations. The detailed information was as follows is identified in Table 3.

Table 3: The number of sampling groups of collocations

<table>
<thead>
<tr>
<th>Base words</th>
<th>Number of words</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. &quot;กลิ่น&quot; (smell)</td>
<td>24</td>
</tr>
<tr>
<td>2. &quot;ใจ&quot; (heart)</td>
<td>48</td>
</tr>
<tr>
<td>3. &quot;ปาก&quot; (mouth)</td>
<td>38</td>
</tr>
<tr>
<td>4. &quot;พูด&quot; (speak)</td>
<td>55</td>
</tr>
<tr>
<td>5. รูปร่าง : &quot;ผอม สูง อ้วน&quot; (Figure: thin, tall, fat)</td>
<td>35</td>
</tr>
<tr>
<td>6. &quot;รส&quot; (taste)</td>
<td>16</td>
</tr>
<tr>
<td>7. &quot;สี&quot; (color)</td>
<td>71</td>
</tr>
<tr>
<td>8. สัมผัส : &quot;แข็ง หยาบ อ่อน&quot; (Touch: hard, rough, soft)</td>
<td>37</td>
</tr>
<tr>
<td>9. &quot;เสีย&quot; (lose)</td>
<td>28</td>
</tr>
<tr>
<td>10. &quot;เสียง&quot; (sound)</td>
<td>48</td>
</tr>
<tr>
<td>11. &quot;หน้า&quot; (face)</td>
<td>56</td>
</tr>
<tr>
<td>12. อารมณ์: &quot;ยิ้ม ร้องไห้ หัวเราะ&quot; (Emotion: smile, cry, laugh)</td>
<td>36</td>
</tr>
<tr>
<td>13. คำเฉพาะ (specific words)</td>
<td>28</td>
</tr>
<tr>
<td>14. ส่วนวน (idioms)</td>
<td>28</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>548</strong></td>
</tr>
</tbody>
</table>
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/ has the most number of similarities which are 66 words or 12.04 in percentages. For the base word “พูด” /phuut/ it was the secondary rank in similarities. There were 52 words or 9.49 per cent, and for the base word “ใจ” /jai/ 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/ has the most number of differences which are 40 words or 7.3%. For the base word “ปาก” /bpaak/ it was the secondary rank in differences. There were 28 words or 5.11%, and for the base words “กิ่ง” /glin/, "รูป รูป" /ruup raang/, "อารมณ์" /aa rohm/, and "ส่วนวัน" /sam nuaan/, 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/.

244
(heart) in Vietnamese: 1) tâm /tαm/ (center) 2) lòng /lɔŋ/ (heart, barrel, bore, and bosom) 3) bụng /bʊŋ/ (stomach, belly, abdomen, maw, and paunch) 4) gan /gaŋ/ (liver, daring and near to). All 4 words when co-occurring with other words have the meaning of "ใจ" /jaiM/, 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” /bpaak¹ maak⁷/ in Thai matches that in Vietnamese làm môm /lom¹ mɔm⁷/ and the word “พูดน้อย” /phuut⁶ naawy⁴/ in Thai matches that in Vietnamese ít nói /it nɔi/ mj⁷/.

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; “กึ่งหน้าตา” /glinL dtaoL/ or “กึ่งหน้ารักแวว” /glinL rakM raehM/ 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 /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%.
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


