Thai Pop Music:
Corpus Analysis and Second Language Learning

Teddy Bofman (T-Bofman@neiu.edu)
and
Paul Prez (P-Prez@neiu.edu)
Northeastern Illinois University

Abstract

This paper examines a corpus of 400 popular songs downloaded from www.ethaimusic.com on the Internet. The techniques of corpus linguistics were used to analyze this body of music in an effort to better prepare language teachers to use these songs in the classroom.

Corpus linguistics is the study of language using large bodies of text. While linguists have always been interested in data, that interest waned considerably during the heyday of Noam Chomsky. The modern computer has refocused our attention and allowed us to examine large bodies of authentic data.

The analysis of pop songs that we present allows us to focus on key linguistic features found in this music. Using Oxford WordSmith Tools 4.0, we have looked at word frequencies and concordances in order to get an overall picture of the language found in these songs. We examine features such as pronoun usage, references to emotions and body parts, elaborate expressions, slang spellings, Sanskrit vocabulary and English terminology found in these texts.

Our findings provide guidance to language teachers in a number of ways. For example, they point out which vocabulary is high frequency and therefore essential to teach. They also indicate major collocational usage, essential knowledge for the language learner. The lyrics we analyzed allow students to engage in data-driven learning. These same tools applied to other genres could have tremendous impact on the way Thai and other languages are taught in the United States.

This study examines a corpus of four hundred popular Thai songs found on the Internet at www.ethaimusic.com. The techniques of corpus linguistics were used to examine this body of music in an effort to help language teachers better prepare to teach this material. We believe that studies such as this one could be applied to various types of discourse such as newspaper articles,
academic writing, business reports, news broadcasts, and sitcoms, as a means of helping language instructors gauge appropriate levels of material and also as a means of helping teachers choose what to target before, during, and after actual lessons.

Our discussion covers several areas, including a brief description of corpus linguistics, the tools we have used in our corpus study, the kinds of information we have learned about the corpus we examined, and finally, suggestions for working with this corpus, given the information we have gleaned from our research.

What is corpus linguistics? It is the study of language using large corpora or bodies of text, whether written, spoken, or signed (Hunston 2002). Linguists have been using corpus techniques for decades. Bofman’s doctoral dissertation, for example, completed at the University of Michigan in 1978, was a corpus study of a sixty-page segment of the Ramakian, the national epic of Thailand. In this study, she examined pragmatics and word usage without the aid of a computer. A portion of that study has been made available on www.seasite.niu.edu/thai. During the heyday of Noam Chomsky interest in language data waned to some extent, with the focus of study based on the intuitions of native speakers, rather than looking at bodies of naturally-occurring data.

Advances in computer technology have reignited an interest in corpus studies, which allow researchers to look at very large bodies of data in an efficient manner. Corpus studies value the use of authentic texts. The songs we have looked at, popular Thai music found on the Internet, are just that. Authentic texts, valuable to all teachers, are of particular importance to teachers who are non-native speakers and may have doubts about their intuitions.

By using corpus techniques to examine a body of texts, teachers are better able to judge whether or not the material is at the appropriate level for the class being taught. Corpus techniques also enable teachers to determine what aspects of the language need to be targeted in actual teaching. Corpus linguistics is one of many tools available to teachers. It is not meant to replace existing techniques, but to complement them.

Corpus linguists investigate computer-readable corpora with the aid of analytical software. Ideally, such software provides the types of analyses that will highlight and profile the features and characteristics of a corpus that are of interest to the individual researcher. For the present study, we used Oxford WordSmith Tools 4.0, published by Oxford University Press. This software package comprises a suite of analytical tools each of which aids in building an overall picture of a corpus. The WordSmith tools that were of most interest to us were WordList and Concord. WordList generates word frequency lists in ascending or descending order as well as in alphabetical order. This same tool also provides statistical data regarding range of use of each lexical item. This information is useful because a word that occurs thirty times in a single song is of less pedagogical significance than a word that occurs thirty times but in a number of songs throughout the corpus. The WordSmith tool, Concord, provides lexis-in-context data, i.e., concordances. Making extensive use of these two functions and the various options included in each, we gained insight into the language of Thai pop music that would be difficult to obtain relying on intuition alone.
Why did we choose this particular corpus to study? The most obvious reason, perhaps, was its accessibility. We had access to hundreds of songs, amounting to roughly 103,000 running words. The corpus provides authentic material using modern colloquial Thai. It is not material written with the goal of teaching in mind. Rather, it is material that stems from the culture itself. Of equal importance, perhaps, is the fact that the topic is inherently of interest to college-age students who enroll in language classes.

What we found in this body of songs is abundant linguistic creativity, a feature that is the hallmark of classical Thai literature and obviously of modern Thai as well. Additionally, the songs are a rich source of cultural information. They make reference to modern life in Thailand, providing an excellent starting point for discussion and explanation of current life in the country. To the best of our knowledge, there is still no frequency list of vocabulary for Thai comparable to the General Service List (GSL) that exists for English (West 1953). These lists provide a rank order of vocabulary in terms of usefulness and frequency of usage. These high frequency vocabulary items are the basic building blocks of language and should be a major focus of language pedagogy. It is our hope that as other genres of Thai are analyzed using corpus techniques, a general service list of Thai will emerge. Our study is perhaps a first step in that direction.

Our vision is that studies such as ours need to be applied to texts representative of various domains in a given language. Corpus studies of Buddhist sermons, classical poetry, newspaper articles, and other types of texts would provide language teachers with information about various genres that could be used as authentic material for the classroom. By sharing our information about these corpora, teachers of Thai, Indonesian, Tagalog, and other Southeast Asian languages would have pedagogical knowledge of a rich and varied supply of authentic material at their disposal.

The website www.ethaimusic.com allows the viewer to see the song written in Thai, transliterated into English, and translated by non-native speakers of English into English. One can hear the song sung and can also consult an online glossary of the vocabulary occurring in the lyrics. These features alone make the music highly accessible to students, whether for classroom study or independent work.

The Thai language is written without spaces between words. In order to be processed digitally, the text has to be segmented into words. This fact provides a tremendous challenge to those who do corpus-based studies of Thai. We went through each song and manually inserted spaces as word boundaries. Decisions were made in a variety of ways, using our knowledge of the language as a starting point. The online glossary provided by the website gave us additional clues. Dictionaries were consulted as well as native speakers. We found that native speakers did not always agree as to what constitutes a word in Thai. In fact, from one day to the next, a native speaker might have a different opinion as to whether something was or was not a single word in Thai. This process of segmentation was laborious and time-consuming and likely to be the source of inconsistencies. However, this particular study does not depend on having perfect counts; instead, we look at trends and therefore feel confident that our data is acceptable. It is to be hoped that as corpus-based work on Southeast Asian languages enters into its own, this particular problem of segmentation will be resolved by the community of Thai scholars.
A positive aspect of the segmentation process is that it allows the researcher to gain initial familiarity with the corpus. It is through this first reading that the researcher begins to make mental notes of the nature of the text. What kinds of vocabulary are most frequent? What are the basic grammatical features used in these songs? What themes are found within the corpus? These questions formed the basis of our analysis.

Throughout the segmentation process, the researchers made notes of trends in the text. In this particular corpus we noted, for example, that, not surprisingly, pronoun usage was largely informal, and we planned to track this usage. We saw that the mood within the songs was largely melancholy; many words describing sadness, confusion, and loneliness were used. While we noted a relatively limited mention of parts of the body, we did note numerous references to nature. We saw abundant usage of doublets, reduplication, internal rhyming, alliteration, and elaborate expressions, structures found in so many Southeast Asian languages. These features will be discussed subsequently.

Once the segmentation was completed, we inserted tags for features that we wished to track throughout the text. One such feature, for example, was doublets or word pairs. Each time a doublet was encountered, we tagged it in the following manner: <doub>. Tagging allowed us to generate a list of all the doublets occurring within the corpus. Armed with such a list, the language teacher can target which doublets need to be pre-taught for maximum comprehension of the texts. Choosing which items to tag depends on the nature of the corpus and the interests of the researcher.

In general, the work of preparing and analyzing a corpus is not a straightforward, linear process. With each reading and rereading, researchers add to their knowledge of the corpus and thus might add to the areas they wish to explore.

After the corpus has been prepared, the next step is to run a frequency list of the vocabulary found in the text. This list is crucial in determining the level of the material. It also provides data allowing the teacher to know which vocabulary the student will most frequently encounter. It is this vocabulary that should be targeted in teaching.

We then generated an alphabetical listing of the vocabulary. This list allows the teacher to look up any given lexical item to determine its frequency within the corpus. Parenthetically, version 4.0 of WordSmith has a less-than-ideal system of alphabetizing Thai. Since it starts with the symbol farthest to the left, it first alphabetizes all words that have vowels that occur to the left of the first consonant, rather than integrating these words the way Thai is actually alphabetized. For example, the word รื้อ/rian/ ‘to study’ occurs before กลาง/klaaŋ/ ‘in the middle’ in this list. However, the opposite would be true in a Thai dictionary. This is a relatively minor inconvenience, and it is easily dealt with. Secondly, there are over 450 different words written in English throughout the corpus. These words are alphabetized first, before any of the Thai words. They pose no particular challenge to the learner of Thai, since they are English and they are written in English.
Yet another feature of WordSmith, Concord, allows the researcher to make concordances, which are listings of all occurrences of a given word, with text appearing to the left and text appearing to the right of that word. The benefits of concordances are considerable, including the ability to generate collocational data. Collocates are words that tend to co-occur. By examining concordances, students are able to see authentic examples of how a word is actually used. Word-in-context examples expand the learners’ depth of vocabulary knowledge. With concordances, the teacher does not have to create artificial examples to use in teaching. From these concordances, language exercises can be constructed using material actually found in the texts.

For English, there is a published list of the two-thousand most “useful” words, listed in the order of their frequency. This list, called the General Service List, was compiled in 1953 by Michael West. Although it has not been updated, it is still widely consulted by linguists and educators. Not surprisingly, a large number of the most frequently used words are function words, the words that are part of the grammar, such as ‘the,’ ‘be,’ ‘of,’ ‘and,’ ‘a,’ ‘to,’ and ‘in.’ They are an absolute necessity to learn, and need to be taught early on.

Our corpus comprises 4,931 types, or different words. Thirty-nine types make up 50.02% of the total running words. These thirty-nine words are basic vocabulary items and are likely to be learned in a first-year Thai class. In order to understand this corpus, it is crucial that students have a command of these words. These thirty-nine words are ‘you, informal,’ ‘I, informal,’ ‘that, which,’ ‘future,’ ‘causative, to,’ ‘so,’ ‘go,’ ‘love,’ ‘that, direct or indirect quotation marker,’ ‘have,’ ‘person,’ ‘get, get to do, can’ ‘be,’ ‘come,’ ‘each other,’ ‘heart,’ ‘know,’ ‘man’ ‘it,’ ‘but,’ ‘this,’ ‘do, make,’ ‘person,’ ‘know,’ ‘who,’ ‘still,’ ‘as far as,’ ‘etc.’ ‘do, make,’ ‘must,’ ‘day,’ ‘good,’ ‘-ness,’ ‘we,’ ‘with,’ ‘tell,’ ‘then, past tense marker,’ ‘that,’ ‘in,’ and ‘like, kind.’

What is important here is the fact that fifty percent of the vocabulary in this corpus is likely to be covered very early on in a Thai class. This fact alone suggests that the corpus is highly suitable for a beginner’s class, since the majority of its vocabulary consists of words that are of high frequency.

Having established the accessibility of the corpus to beginning learners, let us look at other significant features of these songs. The corpus we studied consists of four hundred songs and roughly 103,000 running words. Well over half of the songs are sung by males. Roughly three-fourths of the songs are slow ballads, a fact which makes this corpus particularly useful for relative beginners in a language, since the slow pace makes listening comprehension easier. Those songs that are not slow ballads include rap music, jazz, hip hop and dance music.

The songs are overwhelmingly melancholy in nature. They deal with themes such as lost love, shyness, romantic yearning, and scorn. Typical lines from these songs are “We’ve been apart for a long time,” “I will dry your tears,” “Do you miss me when you are lonely?” and “I know you
don’t need me.” A small number of these songs (less than one quarter) deal with other topics, such as fun, sexuality, and politics.

Words dealing with sadness far outnumber words dealing with happiness in this corpus. Words meaning ‘sad,’ ‘lonely,’ ‘disheartened,’ ‘confused,’ and ‘worried’ appear in abundance throughout the corpus and need to be targeted in teaching. In contrast, words for ‘happiness,’ such as ยินดี/yindii/ and สุข/sùk/, are sparsely scattered throughout the corpus.

Nature is a theme that one might expect to find in a corpus of songs dealing primarily with love. The most frequently mentioned word related to the theme of nature is ฟ้า/fáa/ ‘sky’ or ‘heaven,’ occurring by itself ninety-nine times. It also occurs in compounds such as ท้องฟ้า/thóng fáa/ ‘sky’ (11 times) and นางฟ้า/nâa fáa/ ‘angel’ (6 times). Also frequent is the word น้ำ/náam/ ‘water,’ (50 times), occurring also in compounds such as น้ำตา/námtaa/ ‘tears’ (91 times), น้ำใจ/n mácay/ ‘considerate of the feelings of others’ (13 times), and น้ำลาย/n málaay/ ‘saliva’ (7 times). The Thai word ลม/lom/ ‘air’ or ‘wind,’ occurs by itself forty-eight times and in compounds such as ลมหายใจ/lomhǎaycay/ ‘breath’ (8 times). Nature phenomena include ดาว/daaw/ ‘star,’ (34 times), จันทร์/can/ ‘moon,’ (26 times), อาทิตย์/aathít/ ‘sun,’ (3 times), นา/naa/ ‘rice field,’ (36 times), น้ำ/thông/ ‘water,’ (50 times), ทะเล/thálee/ ‘sea,’ (5 times), นก/nók/ ‘bird,’ (14 times), ปลา/plaa/ ‘fish,’ (8 times), งู/ŋuu/ ‘snake,’ (10 times), และ/meew/ ‘cat,’ (5 times), and แมว/mεε/ ‘cat,’ (5 times), and หมา/mǎa/ ‘dog,’ (1 time). Clearly these words, many of which are high frequency words, need to be addressed pedagogically.

Given that so many of these songs are about love, an American audience might expect body parts such as eyes and lips to play a large role in the lyrics. But in this corpus, fewer than twenty terms for body parts are used, and used sparingly. They include เอว/eew/ ‘waist,’ แขน/khεn/ ‘arm,’ หลัง/lǎŋ/ ‘back,’ ปาก/pàak/ ‘mouth,’ ฟัน/fan/ ‘tooth,’ เขี้ยว/khîaw/ ‘fang,’ หน้า/nâa/ ‘face,’ หู/hǔu/ ‘ear,’ อก/ök/ ‘chest,’ คอ/khóo/ ‘neck,’ หัว/hǔa/ ‘head,’ แก้ม/kεm/ ‘cheek.’ Overwhelmingly, the body part most frequently mentioned in the corpus is ใจ/cay/ ‘heart;’ it occurs alone roughly one thousand times, and in compounds roughly two thousand times. These compounds frequently refer to moods and personality traits, such as ใจดี/caydii/ ‘kindness’ and ใจหาย/cayhǎay/ ‘frightened.’ On the other hand, จิตใจ/citcay/, rather than being a physical body part, refers to the mind. In contrast, the next most frequently mentioned body part is มือ/máa/ ‘hand,’ which occurs fewer than one hundred and fifty times. Clearly, an essential focus of teaching should center around the word ใจ/cay/ and its compounds.

A great deal of subtlety is expressed through the use of pronouns in Thai. Pronoun choice indicates levels of formality and intimacy and thus needs to be carefully understood when listening to love songs. It comes as no surprise that the pronouns typically found in this corpus tend towards the informal. The two most frequently occurring pronouns – เธอ/thəə/ ‘you, informal’ and ฉัน/chán/ ‘I, informal’ – occur over four thousand times (4,356) and over three thousand times (3,081), respectively. The other pronouns which need to be taught are คุณ/man/
‘it’ (973 times), ฉัน /raw/ ‘we’ (598 times), ฉัน /khāw/ ‘he’ or ‘she’ (501 times), ผม /phōm/ ‘I, masculine, formal’ (129 times), คุณ /khun/ ‘you, formal’ (127 times), น้อง /nōng/ ‘I, you, younger sibling, informal’ (35 times), ฉัน /ay/ ‘I, informal’ (12 times), น้อง /ngūw/ ‘I or you, literally rat, informal’ (16 times), เหล่า /kuu/ ‘I, informal’ (8 times), เธ่ /phīi/ ‘I or you, elder sibling, informal’ (58 times), and ยู /yu/ ‘you, informal’ (1 time).

Two highly productive grammatical morphemes occur in abundance in this corpus among those that occur in spoken and written Thai. They are ความ /khwaam/, which changes verbs to nouns and which occurs 661 times, and การ /kaan/, which corresponds roughly to a gerund in English and occurs 48 times. Specific examples are ความรัก /khwaam rák/ ‘love’ and การคิดถึง /kaan khítthung/ ‘feeling homesick.’ While not terribly difficult, these two constructions might not be covered in a first year Thai class; consequently, they would need to be presented before the corpus is used in a beginners’ class.

Two structural features which we did not tag are worth mentioning. The first is the structure X negative Y, such as นอนไม์หลับ /n̄n̄ mây làp/ ‘lie down, not sleep,’ where the second verb is the unfulfilled result of the first. It is our impression that this structure occurs with moderate frequency, and thus it would be worth mentioning before teaching this corpus.

The second structure, also not tagged, is serial verbs. It is our impression that this structure is not very prominent in the corpus. Were we to do another round of tagging, we would definitely look at these structures to assess their significance within the corpus, rather than relying on our impressions.

Lewis (1993) talks of the advantages of teaching language in meaningful chunks. Thai lends itself well to this strategy. A ubiquitous feature of Thai is the pairing of words, often in rhyming doublets or words paired through alliteration. Sometimes the words are synonyms. At other times a word may occur only as a part of a pair and does not appear to have a meaning of its own. An example of a rhyming doublet is เปรียบเทียบ /prīap thīap/, in which both words singly and together mean ‘compare.’ An example of a pair using alliteration is มากมาย /mâak maay/, in which both words, singly and together, mean ‘a lot.’ The word เหมาะ /mâak/ is heard more frequently in everyday Thai than its pair สม่ำ /maay/. There are well over one thousand such doublets occurring throughout the corpus, in almost every song. The fact that these words often rhyme or share alliteration makes it easier to recognize a doublet. It pays for the classroom teacher to teach these word pairs. They are important chunks found in everyday spoken language as well as in poetry and song. Ability to recognize and use them makes one a more comfortable and more natural user of Thai.

Also ubiquitous in both Thai and in our corpus, but less of a challenge to students, is the use of reduplication. Adjectives and adverbs can be repeated, making the meaning less precise or intensifying the meaning. Nouns, too, are sometimes repeated, as a form of pluralization. In some cases, the word is written twice, while in others it is written once and followed by ฯ, a marker which signals reduplication. Once students recognize the marker, they need give little
additional attention to this form other than to determine whether the reduplication is meant to intensify or make the meaning less precise.

Although a relatively small portion of the corpus, English words written in Thai letters pose a challenge to non-Thai readers. Scattered throughout the corpus are over one hundred words of this nature. They include words like /mítnay/ ‘midnight’ /lipsatik/ ‘lipstick,’ /buummeerɛəŋ/ ‘boomerang,’ /beebi/ ‘baby,’ and /khɛɛ/ ‘care.’ Many of these words, though clearly not all, are easily associated with western culture and modernity. What is important is that, even for relatively experienced students of the language, these words are difficult to recognize when encountered in the text. There is nothing that marks them as non-Thai, and one tends to try to interpret them along Thai lines until the realization hits that the word is really an English word. A worthwhile exercise to give before these songs are studied would be one that matches the English words written in the two different alphabets, Thai and English. That way, when encountered, the student is already familiar with the word and less likely to be caught off guard. An exercise on this phenomenon can be found in the Appendix.

Sanskrit words are sparsely scattered throughout this corpus. The word /kam/ ‘karma,’ for example, occurs thirteen times. The word /yütìtham/ ‘justice’ occurs three times, and /àtsacan/ ‘miraculous,’ occurs just once. These words are best dealt with as they arise within a given song, since they are relatively infrequent. Given the highly modern content of this particular corpus, it is interesting to note that certain features found in classical Thai poetry are also found in this modern genre. Classical Thai poetry is known for an abundance of internal rhyme, where multiple words rhyme in a given line and sometimes extend into the next line as well. Our corpus is filled with examples of internal rhyme such as /mii cay kèp wáy hây thəə/, where /cay/, /wáy/, and /hây/ all rhyme.

Similarly, alliteration, where multiple words begin with the same sound, is rampant throughout the corpus. An example of this phenomenon is /thúk thɔɔy kham thii thəə ɔɔy/, where /thúk/, /thɔɔy/, and /thii/, all begin with a /th/ sound.

A hallmark feature of many Southeast Asian languages, including Thai, is the use of elaborate expressions, in both spoken and written language. This form of word play gives beauty and elegance to the language, and demonstrates a certain linguistic virtuosity. Two patterns are highly typical, ABAC and ABCB, although variations on these two patterns can and do occur. An example of the ABAC pattern found in this corpus is /yîŋ klây yîŋ chít/, where the first and third syllables, /yîŋ/, are repeated. An example of the ABCB pattern found in the corpus is /tua eeŋ thîaw eeŋ/, where the second and fourth syllables, /eeŋ/, are repeated.

The frequent use of doublets, reduplications, internal rhyme, alliteration, and elaborate expressions all point to the need for the classroom teacher to discuss these features before introducing the songs in this corpus. Students who go on to study advanced Thai are likely to come across these forms. These popular songs provide a reason to introduce these features of the language before the student reaches advanced levels of language study.
A final feature worth noting involves the written form of certain words in this corpus. Throughout the corpus we have found variant spellings for certain words. They are slang equivalents, such as รู้/รู้ / instead of หรู/รู้ ‘or’ and อย่าง/อย่าง instead of อย่าง/อย่าง/อย่าง /อย่าง /อย่าง/อย่าง/อย่าง /อย่าง/อย่าง/อย่าง/อย่าง ‘how.’ While not terribly frequent, they can be confusing to the learner, and they are worth pointing out to students of these songs.

Having explained the process and some of our findings, let us look at some language exercises based on high-frequency vocabulary highlighted in this study.

***
Appendix I
EXERCISES

Pt. I. Match the Thai (1-4) with the English equivalent (A-D).

1. แล้วเธอจะเกลียดไหม  [C]
2. เธอให้สัญญาว่าเธอจะมีฉันรักกันตลอดในชาตินี้  [A]
3. อยากจะขอให้เธอตอบท่าน  [B]
4. หากเธอได้เห็นว่าในบางทีก็มีที่หวั่นไหว  [D]

A. You promised that we would love each other for the rest of our lives.
B. I want you to think it over.
C. Would you hate me?
D. If you found that sometimes there would be uncertainty…

Pt. II. Use your knowledge of Thai to help you answer these questions.

1. เธอให้สัญญาว่าเธอจะมีฉัน รักกันตลอดในชาตินี้ (You promised we’d love each other for the rest of our lives.)

Which word means promise? (50 times)

2. แล้วเธอจะเกลียดไหม (Would you hate me?)

Which word means hate? (29 times)

3. อยากจะขอให้เธอตอบท่าน (I want you to think it over.)

Which word means think it over? (21 times)
4. หากเธอได้เห็นว่าในบางทีก็มีที่หวั่นไหว
   (If you found that sometimes there would be uncertainty…)

Which word means uncertainty?  (34 times)

CLOZE EXERCISE

Fill in the blanks in the sentences with the correct word from the list of 5 below.

ทบทวน คบ สัญญา หวั่นไหว เกลียด

1. แล้วเธอจะ เกลียด ไหม
   Would you hate me?

2. เธอให้ สัญญา ว่าเธอจะมีฉัน รักกันตลอดในชาตินี้
   You promised we’d love each other for the rest of our lives.

3. อยากจะขอให้เธอ ทบทวน.
   I want you to think it over.

4. หากเธอได้เห็นว่าในบางทีก็มีที่หวั่นไหว.
   If you found that sometimes there would be uncertainty…

PRONOUNS

Choose the best pronoun.  (Answers in red.)

1. นั่นรัก__________  (I love you.)
   ผม (I) เรัก (we) เธอ (you)

2. __________ เป็นสาวอีสาน (I’m a girl from the northeast.)
   ผม (I-masculine) เธอ (you) น้อง (I-younger one/term of endearment)
PHRASE COMPLETION

Join phrase segments from columns A and B.

<table>
<thead>
<tr>
<th>COLUMN A</th>
<th>COLUMN B</th>
</tr>
</thead>
<tbody>
<tr>
<td>คนแรก</td>
<td>กว่าใครๆ</td>
</tr>
<tr>
<td>ไม่ต้องรักฉันมาก</td>
<td>ไหมช่วงราว</td>
</tr>
<tr>
<td>ยืมเป็นพี่ชายได้</td>
<td>ที่รัก</td>
</tr>
</tbody>
</table>

Write out the complete phrase.

1. The first person I’ve ever loved.
   คนแรกที่รัก

2. May I borrow you to be my big brother for a while?
   ยืมเป็นพี่ชายได้ไหมช่วงราว

3. No need to love me more than anyone else.
   ไม่ต้องรักฉันมากกว่าใครๆ

CONCORDANCE

Study the concordance lines. Then answer the questions that follow the list.

1. นั่น  ก็คิด  ถึงเธออยู่ทุกเวลา
2. เขาไม่  ก็คิด  มีใจ
3. นั่นไม่  ก็คิด  ว่าเธอจะมีอีกคน
4. ยังคง  ก็คิด  ถึงเธอตลอดมา
5. นั่น  ก็คิด  ว่าคงหน้าตาเหมือนเธอ
6. เพราะ  ก็คิด  ถึงเธอไม่อยากไปไหน
7. ไม่  ก็คิด  ไปจริงจริง
8. ไม่  ก็คิด  ว่าเธอจะมีอีกคน
1. In this concordance what words typically followed คิด? ถึง ว่า

2. What kind of word typically followed ถึง? A pronoun

3. Is it formal or informal? Informal

4. What kind of structure typically followed ว่า? A clause

6. How would you translate ว่า into English? That

WORD PAIRS

Make word pairs using items from the list below. (Answers in red.)

ดาย นอน ตาย สน อุ่น

มาก ย (a lot)

แน่น นอน (certain)

อบ อุ่น (good-natured)

เดียว ดาย (lonely)

สับสน (confused)

FEELINGS EXPRESSED WITH ใจ

Underline the best answer.

1. Your lover just left you. How do you feel?
  คิด (happy)
ประทับใจ (impressed)
หมดใจ (disheartened)

2. Your lover has found somebody new. How do you feel?
ช้ำใจ (hurt)
จริงใจ (sincere)
ไว้ใจ (trusting)

3. You are not sure what to do. How do you feel?
วุ่นวายใจ (confused)
ไว้ใจ (trusting)
ยุ่งใจ (secure)

OPPOSITES
Draw lines to connect the opposites in columns A and B.

COLUMN A                  COLUMN B
จน                  (poor)                   ยินดี                  (happy)
เลว                  (bad)                    รัก                  (love)
เกลียด               (hate)                    ดี                  (good)
เศร้า                 (sad)                    รวย                  (rich)

FINDING ENGLISH WORDS
Each of these lines has an English word written in Thai letters. Underline the English word.
Choose from the following:

professor, care, technology, concert, BMW, World Trade, cocaine
1. โปรดอย่าทำลายด้วยเทคโนโลยี
2. มันไม่มีเครื่องและไม่หวั่น
3. เวิร์ลเทรดก็ดูดี
4. เปลี่ยนให้เขาได้เน้นปิดอัลบั้มเบิ้งยู
5. อยากเปลี่ยนโคเคนเป็นแตงโม
6. ได้เป็นโปรเฟเรซ_refs
7. ขอให้ชีวิตสดใสยิ่งกว่าแสงไฟคอนเสิร์ต

Appendix II
TAGGING EXAMPLES

ไม่ต้องแงะ<redup>ติด ๆ<redup>กันไว้<redup>
<elab>จะ กอด ไม่ กอด<elab>
กอด ตัน นะ ตาม<Sanskrit>ธรรมเนียม<Sanskrit>คน
ระหว่าง เรา<doub>มากมาย<doub>
วัน<eng>วันนี้<eng>จะ ส่ง ดอกไม้ เรา จะ ไป ให้ใคร
<englet>Welcome to 5 club<englet>ยินดี
เก็บไว้ ใน<body>ยก<body>นี้ มัน อีด อยู่ ใน นี้
ไม่ว่าเธอจะอยู่ที่ใดก็ยังส่งใจไป

ถอยคำที่เธอเคยบอกฉัน

ELABORATE EXPRESSIONS
ยิ่งใกล้ยิ่งชิด
A B A C
OR

ALTERNATE SPELLINGS
รึหรือ
ไม่ไหม
ยังไงหรืออย่างไร

Appendix III

Word Frequency List:
First Four Hundred Words

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>เธอ</td>
<td>7</td>
<td>ก็</td>
</tr>
<tr>
<td>2</td>
<td>ไม่</td>
<td>8</td>
<td>รัก</td>
</tr>
<tr>
<td>3</td>
<td>ที่</td>
<td>9</td>
<td>ไป</td>
</tr>
<tr>
<td>4</td>
<td>ทั้ง</td>
<td>10</td>
<td>ว่า</td>
</tr>
<tr>
<td>5</td>
<td>ให้</td>
<td>11</td>
<td>มี</td>
</tr>
<tr>
<td>6</td>
<td>จะ</td>
<td>12</td>
<td>คน</td>
</tr>
<tr>
<td>7</td>
<td>เดี</td>
<td>13</td>
<td>ได้</td>
</tr>
<tr>
<td>8</td>
<td>เป็น</td>
<td>14</td>
<td>ใจ</td>
</tr>
<tr>
<td>9</td>
<td>นี้</td>
<td>15</td>
<td>มา</td>
</tr>
<tr>
<td>10</td>
<td>อยู่</td>
<td>16</td>
<td>กัน</td>
</tr>
<tr>
<td>11</td>
<td>อย่างไร</td>
<td>17</td>
<td>แท้</td>
</tr>
<tr>
<td>12</td>
<td>ทำ</td>
<td>18</td>
<td>ยิ่ง</td>
</tr>
<tr>
<td>13</td>
<td>รัก</td>
<td>19</td>
<td>หรือ</td>
</tr>
<tr>
<td>14</td>
<td>เป็น</td>
<td>20</td>
<td>นั้น</td>
</tr>
<tr>
<td>15</td>
<td>ใจ</td>
<td>21</td>
<td>นี้</td>
</tr>
<tr>
<td>16</td>
<td>มา</td>
<td>22</td>
<td>อยู่</td>
</tr>
<tr>
<td>17</td>
<td>กัน</td>
<td>23</td>
<td>อยาก</td>
</tr>
<tr>
<td>18</td>
<td>แท้</td>
<td>24</td>
<td>ทำ</td>
</tr>
</tbody>
</table>
ความรักในทุกวันต้องการความคงไว้ไหมบอกเคยในอย่างเขาขอและหัวใจคิดทุกของจริงถึงเหมือนไว้คำตัวอะไรเพราะจากถ้าดูไหนอย่าเพียงเอาเจอเมื่อใช้นานพูดยอมทีเจ็บสักเองพอมองเดียวลืมหรือตาฝนมันกลับรอทั้งเข้าใจผูยิ่งเรื่องตรงเห็นเวลาเดินเปลี่ยนหน้าคือหากพบอาจต่อทางกว่าเปล่าช่วยไกลครั้งอีกจนสิ่งผ่านมากทิ้งหมดคอยเกินเหลือใดยังไงคืนชอบแม่มือหาเข้าเดิมเพื่อทนข้างรับหนึ่งเก็บผมคุณนอนจำถามเพื่อนหญิงชีวิตหวังเหงาซักฟังลงสายฟ้าชายหายแม่น้ำปล่อยลมต้องการไงหวง
| 161 | จบ | 195 | ไร้دين | 230 | คุย | 265 | นิด |
| 162 | ตาย | 196 | กอด | 231 | พ้อ | 266 | ว่าง |
| 163 | น้ํา | 197 | เกิด | 232 | เสีย | 267 | สายใจ |
| 164 | แสน | 198 | ก่อน | 233 | ใช้ | 268 | เชือ |
| 165 | ตอน | 199 | ร้อง | 234 | บ้าน | 269 | เดิน |
| 166 | ลา | 200 | ฝาก | 235 | สวย | 270 | กล้า |
| 167 | กอด | 202 | ผลงาน | 236 | อย่างไร | 271 | ร้อง |
| 168 | ชั่ว | 203 | ไน | 237 | แผล | 272 | สัญญา |
| 169 | แท้ | 204 | โลก | 238 | อย่าง | 273 | ก่อน |
| 170 | _FA | 205 | กลัว | 239 | เท่าไร | 274 | อาจจะ |
| 171 | มาก | 206 | นิง | 240 | เท่าไร | 275 | เหยี่ยม |
| 172 | เสียใจ | 207 | ลง | 241 | สำาคัญ | 276 | แทน |
| 173 | แบบ | 208 | สุขสันต์ | 242 | _RAD | 277 | ทะเล |
| 174 | ส่ง | 209 | เจ้า | 243 | ตอก | 278 | งาย |
| 175 | ผิด | 210 | ไหม | 244 | สุ dụng | 279 | น่าย |
| 176 | เท่าไหร่ | 211 | ถูก | 245 | ผิด | 280 | หยุด |
| 177 | ใหม่ | 212 | ส่ง | 246 | ชั่ว | 281 | ดัง |
| 178 | ก่า | 213 | บอก | 247 | ภาพ | 282 | ร้าว |
| 179 | ใจ | 214 | อี้สิก | 248 | ทรงก | 283 | ถูก |
| 180 | แจ้ง | 215 | หมาย | 249 | หัวสนใจ | 284 | เข้น |
| 181 | ถาม | 216 | เตาหนึ่ง | 250 | เสียง | 285 | ขาด |
| 182 | ต่าง | 217 | เหต | 251 | เท่า | 286 | ยู่ |
| 183 | พร้อม | 218 | ไร | 252 | ช่าง | 287 | ซิ่ง |
| 184 | เร่ง | 219 | นิ้วทา | 253 | นี่ | 288 | สาเร็จ |
| 185 | ดัง | 220 | พัน | 254 | สัมผัส | 289 | ที่มัน |
| 186 | ห่าง | 221 | เท่า | 255 | สุข | 290 | ไป |
| 187 | กอด | 222 | เท้าไหร่ | 256 | เส้น | 291 | ดวง |
| 188 | ถึง | 223 | เทมิยา | 257 | รวม | 292 | หา |
| 189 | ลง | 224 | ที | 258 | นั้น | 293 | แผล |
| 190 | หัก | 225 | ที่ | 259 | งาม | 294 | แผล |
| 191 | ออก | 226 | รู้ | 260 | เทพสุด | 295 | จง |
| 192 | เพลง | 227 | เที่ยง | 261 | ปวด | 296 | น่าที |
| 193 | เสียง | 228 | บ้าน | 262 | ปวด | 297 | น่า |
| 194 | แสบร | 229 | หลับ | 263 | ส้น | 298 | หลอก |
| 195 | ไร้دين | 230 | คุย | 265 | นิด | 299 | เด้า |
เลือก ปิด ใหญ่ 300 336 นิด หัว ทาย 371 หยิบ
โทร พวง ทาย 301 337 กลาง 372 ทา
บัง อ้าย 302 338 ผม 373 พา
ตก ลึก 303 339 เกิน 374 มากยิ่ง
ส่ง 304 340 เล่น 375 ยุ
ติด เล็บ 305 341 เรือน 376 เกลี้ยง
ทัน 306 342 แตกไม 377 เนื้อ
ยึก เล่น 308 343 แตก 378 แกแบ่ง
สาย 309 344 พอ 379 หมวด
หน้า 310 345 ปุ 380 แยก
กลาง 311 346 ออก 381 ขอ
กลาย 312 347 หว่า 382 จั่ง
นั่ง 313 348 หยัก 383 ขับ
ปาก 314 349 เจาะ 384 ตึงเต็ง
เข้า 315 350 เจาะ 385 ถือ
เค 316 351 โรค 386 ท้อง
โค 317 352 โปร่ง 387 ยุ
อื่น 318 353 คอ 388 บาง
เลือก 320 354 สน 389 บาน
กาย 321 355 บาน 390 วาง
ผูก เลือน 322 356 ท่า 391 หมวดโย
แกลง 323 357 ชี้ 392 อาการ
น้อย 324 358 ล้ม 393 ดีม
นับ 325 359 ลิ้นชี 394 มือ longing
มั่ว 326 361 โกรธ 395 กล้า
รูจัก จับ 327 362 จ้าน 396 ทาง
หายใจ 328 363 ขี้ 397 ขัด
หยิบ 329 364 ดัน 398 ดวงใจ
แซ่ 330 365 พัด 399 ต้น
ไฟ 331 366 ต่อ 400 บังเอิญ
ไฟ 332 367 พบ 401
ควร 333 368 อุ่น 402 ทุกฟัน
งาม 334 369 เดือ 403
เมื่อไร 335 370
Acknowledgements

We would like to thank Carol J. Compton, Rhodalyne Crail, John Hartmann and Amara Vatanakul for their kind assistance with this article.

References

Electronic References


Print References


