LUNDS
UNIVERSITET

Investigating effects of L1 influence and phrase type on L2 processing of contextualized phraseological units – A self-paced reading experiment

Name: Marcus Olofsson

Autumn Semester 2015

Supervisor: Henrik Gyllstad

Centre for Languages and Literature

Lund University

Abstract

This essay investigates the online L2 processing of restricted collocations and figurative idioms. Earlier research has shown there to be a difference in how these word combinations are processed depending on a variety of factors. In this study a self-paced reading experiment is administered to 15 second language learners of English (L1 Swedish) where their reading time is measured by DMDX software. I look for differences in reading time depending on what type of condition the word combinations adhere to. Four critical conditions are tested, namely: congruent restricted collocations, incongruent restricted collocations, congruent figurative idioms and incongruent figurative idioms. The test items consist of 12 verb + object noun word combinations per condition. An additional 12 items consisting of infelicitous verb + object noun word combinations are also tested to serve as a base line. All the test items are incorporated into sentences. The participants are presented with these sentences segment-by-segment using DMDX software. They are also presented with 60 fillers which consist of randomly selected sentences where none of the segments presented adhere to any of the conditions tested. Although some earlier research observed differences in processing cost between different phraseological conditions, in this study no statistically significant differences were observed for mean reading times cross the four critical conditions. The results yielded by the experiment in this study show that there might be more to processing cost of collocations and idioms than previously disclosed.

Table of Contents

1. Introduction

According to Howarth (1996, 1998) it has been generally accepted that students learning English on a higher level need to master a broad variety of word combinations. These word combinations are believed to be processed as a whole by native speakers (NS), but when it comes to second language (L2) learners there is not sufficient research with regards to how they process word combinations, or as some linguists call them phraseological units (PUs), to determine how phraseological proficiency develops in L2 learners of English (Howarth, 1996, 1998). There is a wide variety of PUs. The different types of PUs include *collocations, idioms, metaphors, compounds* and *proverbs*. There are collocations and idioms of different kinds such as *blow a fuse* (verb + object noun), *under attack* (preposition + object noun) and *make a bet* (verb + object noun) (Allerton et al., 2004). Howarth (1996, 1998) divides these PUs into four categories: *free combinations*, *restricted collocations*, *figurative idioms* and *pure idioms*. These will all be described in detail in subsection 2.1 in this essay.

When studying these PUs there are several different frameworks and ways of describing different phenomena within the field of phraseology. This study will rely on the notion that the representation in the first language (L1) will influence how a corresponding PU in the L2 is processed. Gyllstad and Wolter (2016) state that there are two factors that empirically have been found to affect the processing efficiency of L2 word combinations, namely congruency and frequency of input. Congruent L2 word combinations are combinations where the L1 combination corresponds with the L2 word combination in a word-for-word translation and in terms of core meaning (p. 298). An example of such a congruent combination is *book a table* which in Swedish would be *boka ett bord*. Both constructions consist of a verb, an indefinite article and a noun, and they also correspond word-for-word and in meaning. An example of an incongruent combination is *catch a break* which in Swedish would be, if translated word-for-word, *fånga en paus*. Both the English and the Swedish construction consist of a verb, an infinite article and a noun. However, they do not correspond in meaning. The Swedish combination *fånga en paus* has no meaning as a whole and is not frequently used in the Swedish language.

According to several studies congruency has an effect on the processing of PUs (Wolter & Gyllstad, 2011, 2013; Gyllstad & Wolter, 2016; Yamashita & Jiang, 2010). According to Wolter & Gyllstad (2011, 2013) the incongruent PUs that they tested were processed more slowly than the congruent ones.

Studies looking at the processing of different types of phraseological units have to a great extent relied on data from experiments where decontextualized presentation formats were used (Gyllstad & Wolter, 2016; Gyllstad, 2016). However, so far very little research has been carried out in which more ecologically valid methods were used, where participants are asked to process collocations and idioms in context. This is the gap this study will attempt to fill.

In the current study I will investigate the processing efficiency of L2 learners of English (L1 Swedish). I will use Howarth's descriptive framework (described below), and I will look at both congruent and incongruent items to further explore the effects of congruency on L2 processing of PUs. The study will focus on contextualized restricted collocations and figurative idioms and how these are processed by L2 learners of English.

The research question this study will try to answer is: Is there a difference in how efficiently L2 learners process congruent and incongruent items when contextualized, and is there a difference in processing cost between restricted collocations and figurative idioms when contextualized?

This essay will start by describing the theoretical background and the phraseological typology to get a handle on the framework this study is built on. Previous research will be presented and discussed to further illustrate the relevance of the experiment administered in the current study. After that I will focus on the experiment design and the participants of the study. When we are familiar with the experiment and how it was designed, we will move on to the results section and then a discussion section where the results will be analyzed and discussed. Finally, there will be a conclusion section where I discuss what implications this study may have on future research.

2 Theoretical Background

2.1 Phraseological Typology

The field of phraseology is a subfield of linguistics where word combinations (also referred to as multi-word units, phraseological units and lexical bundles) such as *collocations, idioms, metaphors, proverbs* and *compounds* are researched. The field is, in an article titled *Phraseology* by Cowie (1994) published in *The Encyclopedia of Language and Linguistics*, defined as "the

study of the structure, meaning and use of word combinations" (p. 3168). In the research of word combinations there are two major traditions, the frequency based tradition and the phraseological tradition. The frequency based tradition only takes into account the number of times words co-occur in corpora. In this tradition we do not take into consideration the semantic meaning of the word combination but only rely on the number of co-occurrences. Here collocations such as *blow a fuse* and *blow the gaff* would be considered just collocations. In the phraseological tradition researchers look at how the co-occurring words correspond to different meaning and level of semantic transparency (Howarth, 1996, 1998; Gyllstad & Wolter, 2016; Barfield & Gyllstad, 2009; Granger & Paquot, 2008). In the phraseological tradition *blow a fuse* and *blow the gaff* would be considered two different kinds of PUs. *Blow a fuse* would be considered a restricted collocation and *blow the gaff* a pure idiom (Howarth, 1996, 1998).

In this study I am working within the Howarthian framework of phraseology. Howarth (1996, 1998) draws on the phraseological tradition where linguists such as Cowie and Mel'čuk have been prominent contributors. They in turn drew on the Russian tradition of phraseology which was developed during the 1940s – 1960s (Cowie, 1998; Barfield & Gyllstad, 2009). Howarth (1996, 1998) developed a continuum where the different PUs are put into different conditions depending on the type of combination and their level of idiomaticity. Howarth's continuum is presented in Table 1.

	Free combinations	Restricted collocations	Figurative Idioms	Pure idioms
Lexical composites (Verb + Noun)	*blow a trumpet*	*blow a fuse*	*blow your own trumpet*	*blow the gaff*
Grammatical composites (Preposition + Noun)	*under the table*	*under attack*	*under the microscope*	*under the weather*

Table 1. *Visualization of Howarth's (1996) Collocational Continuum.*

The first condition, free combinations, is composed by words used in their literal senses and are freely exchangeable. An example of a free combination is *blow a trumpet*, a verb + object noun combination where both components can be easily exchanged (e.g. blow a trumpet, *blow a bubble, play a trumpet*).

Restricted collocations consist of one component that is used in a specialized manner, "often a figurative sense only found in the context a limited number of collocates" (Howarth, 1998, p. 28). A good example of a restricted collocation is *make money* which, according to *Oxford Dictionary of English Idioms* (ODEI), means "earn money, or acquire it by other means" (p. 375). This is a clear example where one of the components of the PU *make* have a specialized meaning and it is used figuratively. In this PU *make* does not refer to the act of 'manufacturing' but rather to the act of earning or acquiring. Figurative idioms are PUs that "have a metaphorical [meaning] in terms of the whole and have a current literal interpretation" (Howarth, 1998, p. 28). *Break the ice* is a good example of a figurative idiom since it is possible to literally *break the ice* but in its idiomatic use it means to "do or say something to remove […] social awkwardness or tension" (ODEI, 1993, p. 77). Pure idioms are PUs that "have a unitary meaning that cannot be derived from the meaning of the components" (Howarth, 1998, p. 28). The pure idioms "are the most opaque and fixed [condition]" (Howarth, 1998, p. 28). An example of a pure idiom is *blow the gaff*, which is impossible to figure out by looking at its components. The meaning of this pure idiom is to "let out a secret" (ODEI, 1993, p. 72), which has nothing to do, literally, with any blowing of a gaff. A *gaff* is, according to *the Longman Dictionary of Contemporary English*, "the place where someone lives [or] a stick with a hook at the end used to pull big fish out of the water" (LDOCE, 2014).

In the current study I also research the effect of congruency. Congruency between L1 and L2 is achieved when the L1 version of a PU is translatable word-for-word from the L2, in terms of core meanings. For example, in Swedish we have the PU *få kalla fötter* which translates to *get cold feet* which in turn is an English PU. Both versions of the PU would in Howarth's framework be called a figurative idiom. They both carry the same meaning and are word-for-word translatable between the languages. These are the criteria that need to be fulfilled to call a PU congruent (Yamashita & Jiang, 2010; Wolter & Gyllstad, 2011, 2013; Nesselhauf, 2003). Nesselhauf (2003) also takes into consideration the "general syntactic rules" (p.236) of the L1 and L2 when defining congruency. In the present study a PU such as *say a prayer* will be deemed incongruent since the Swedish translation *be en bön* does not translate word-for-word (as it would render **säga en bön*). In the present study we only look at restricted collocations and figurative idioms since free combinations are, with the definition we rely on in the present study, virtually impossible to find incongruent and that there are very few congruent pure idioms. How

the items were chosen will be described at length in section 3.1 (the Item Development section) of this essay.

2.2 Processing of phraseological units

In 1993 a study was conducted by Bahns & Eldaw on 58 German students of English as a foreign language (EFL). The study was done to determine the students' ability to produce English collocations. Bahns & Eldaw did so by running two experiments, one translation task where 34 of the participants were tested and one cloze task where the remaining 24 were tested. They also investigated the connection between general vocabulary and collocational knowledge. In this effort their results showed that general vocabulary and collocational knowledge do not develop simultaneously but rather independent from each other. The results of the translation task and the cloze task showed that the German EFL students were more likely to be able to correctly produce a correct English collocation when the collocations were, what Bahns & Eldaw called, readily paraphrased. They did not further investigate what these readily paraphrased collocations had in common but concluded that, if such collocations exist, learners should focus on the ones that are not readily paraphrased (Bahns & Eldaw, 1993).

Another linguist who also investigated German EFL learners' ability to produce collocations was Nadja Nesselhauf (2003). In her article she investigated the challenges advanced EFL-learners encounter when producing collocations in English. In her investigation she looked at 32 essays written by German-speaking advanced EFL learners. She categorized the types of collocational mistakes made by the students. In doing so she realized that the L1 influence was a factor in at least half of the mistakes made. Nesselhauf therefore applied the idea of congruent word combinations and categorized the collocations accordingly. Her definition of a congruent collocation was the following: "[O]nly combinations that [sound] natural in both languages if they [are] rendered word for word [are] regarded as congruent (considering, however, general syntactic rules of the two languages)" (p. 236). Working with this definition, she realized that non-congruent (incongruent) combinations were where the EFL-learners frequently ran into challenges when producing English collocations (Nesselhauf, 2003).

The congruency-effect on collocational processing has since been investigated in several studies (e.g., Yamashita & Jiang, 2010; Wolter & Gyllstad, 2011, 2013; Gyllstad & Wolter,

2016). Yamashita & Jiang (2010) investigated L1 influence on L2 acquisition. Their study included three participant groups: 20 NS of English, 24 Japanese English as a second language (ESL) learners (ESL students are students learning English in an English speaking country) and 23 Japanese EFL learners (EFL students are students learning English in their home country i.e. Japanese students learning English in Japan where English is not an official language). These groups were presented with an acceptability task where they had to deem 96 English expressions (24 congruent collocations, 24 incongruent collocations and 48 implausible word combinations), presented by DMDX software (described in detail in section 3.3 of this essay) on a computer screen, acceptable or not acceptable by pressing a YES or a NO button on the computer keyboard. Yamashita & Jiang measured reaction time and error rate for the congruent and the incongruent items and found that the NS group was unaffected both in reaction time and in error rate when responding to congruent or incongruent collocations. The ESL group had a higher error rate on the incongruent collocations than the congruent ones while their reaction time was equal on the congruent and on the incongruent collocations. The EFL group, however, showed both a higher error rate and a longer reaction time on the incongruent collocations than on the congruent ones. With these results Yamashita & Jiang point out that there is a long-lasting effect when it comes to congruency on L2 speakers. There was a difference in error rate between the ESL and the EFL groups. This, according to Yamashita & Jiang, shows that the effect of input also has an effect on acquisition.

Two studies that also investigated the effect of congruency on processing were Wolter & Gyllstad (2011) and Wolter & Gyllstad (2013). Wolter & Gyllstad (2011) investigated the influence students' L1 have on their collocational knowledge. In this study two tests were administered to 31 advanced Swedish learners of English. The test items used in this study were chosen in an attempt to use semitransparent word combinations i.e. not free combinations nor pure idioms. One of the tests was a primed lexical decision task. This test was also administered to 37 native speakers of English who served as a control group. The other test was a "test of receptive collocational knowledge" (p. 430). The tests focused on three conditions. Collocations with Swedish to English equivalents, collocations acceptable in English only and one condition that consisted of baseline items. The results showed that the L1 may have influence on the L2 collocation processing ability of a non-native speakers NNS. Wolter & Gyllstad (2013) followed up with a similar study that explored the effect of frequency on the processing of collocations,

both congruent and incongruent. In their study they used 25 L2 learners of English (L1 Swedish) with a high proficiency level in English and a control group consisting of 25 NS. They tested both groups' ability to process collocations with an acceptability judgment task. The items tested in this study were chosen within the framework of the frequency based tradition of collocation. The results of this study showed that there is an effect on processing time linked to the frequency of the collocation. It also showed that there seems to be an advantage when collocations are congruent as opposed to incongruent.

In Gyllstad & Wolter (2016) they compared processing cost between free combinations and restricted collocations. Like in the current study they used the phraseological approach, more specifically drawing on Howarth's collocational continuum, when developing items for the study. To compare the processing cost between restricted collocations and free combination Gyllstad & Wolter administered a semantic judgment task with decontextualized experimental items where two participant groups were tested. One of the participant groups consisted of 27 L2 learners of English (L1 Swedish) and the other group consisted of 38 native speakers of English. The results of the study showed that "semantic transparency affect processing of word combinations" (p. 317). More specifically, there was a difference in how efficiently both participant groups processed free combinations in relation to collocations. They also concluded that "it would seem that lack of congruency between the L1 and the L2 is the biggest stumbling block in L2 collocational processing" (p. 318).

Gyllstad (2016) is a study which is similar to the present study where processing cost of free combinations, collocations and idioms are compared in a semantic judgment task. The participants of the study (30 NS & 21 NNS) were asked to deem whether decontextualized word combinations were "meaningful and natural". The results of this study showed that the results found in Gyllstad & Wolter (2016) holds as far as the NNS goes. The different conditions in Howarth's (1996, 1998) continuum are reflected in how these are being processed by NNS. However, no such difference was found in the NS group. This study compared the processing of three conditions in Howarth's (1996, 1998) continuum. Even though the subjects of the current study consist of NNS only, the results of Gyllstad (2016) are still a significant point on which to set the bearing for the current study.

The present study aims to make the test environment more ecological than the studies described above. One way to do so is to contextualize the test items in a self-paced reading task

(described below). In a study by Tremblay et al. (2011), lexical bundles (strings of frequently co-occurring words such as *in the middle of the*) were studied in an effort to ascertain whether they are stored and processed holistically. They studied processing advantage of lexical bundles with three differently designed self-paced reading tasks where non-lexical bundles (strings of less frequently co-occurring words such as *in the front of* the) were used to compare processing cost with lexical bundles. In their study they compare the results of three different ways of presenting the items in the self-paced reading tests. The first method used to present the lexical bundles was a word-by-word presentation where participants were shown the sentences containing the lexical bundle one word at a time. The second version was a portion-by-portion presentation where the sentences were chopped up in portions, whereof one portion contained the lexical bundles Tremblay et al. wanted to study. In the third version they presented the whole sentence. The portion-by-portion experiment was designed to provide the participants with a naturalistic way of reading and proved to yield the most significant results when it came to difference in processing between the lexical bundles and the non-lexical bundles.

Taking stock of the above reviewed studies, there have been several efforts to show the effects of frequency and/or congruency in phraseological processing research. It seems, when looking at previous research, that frequency and congruency both have an effect on processing ability of non-native speakers. It seems, however, that there is a gap in the literature regarding learners processing ability of PUs where they are contextualized. In all the studies I have encountered, decontextualized items were used to measure processing cost of PUs. In the current study I will, therefore, contextualize our items in an effort to make the experiment more ecologically valid. Meaning, the items will be presented to the participants in a way they would experience them when processing text in real life. For example, when reading a book or any other type of text PUs will generally be contextualized.

3. Methods & Materials

3.1 Item Development

In the present study I wanted to run a self-paced reading experiment where I manipulated two of the Howarthian conditions, restricted collocations and figurative idioms. Each of the conditions

were divided into two sub-conditions, i.e. congruent items and incongruent items. This resulted in four critical conditions and one baseline condition.

The first step in developing the items for this study was compiling a list of possible candidates for each condition. This was done by drawing on previous research as well as on several idiom dictionaries, and when needed corpora. It was decided early in the process that a verb + object noun construction would be beneficial since those seemed to be the PUs that were most frequently occurring when looking for congruent items. The list started off as a list of eight conditions but in the process of compiling items I realized that there were very few congruent pure idioms and not enough incongruent free combinations. Restricted collocations as well as figurative idioms were ideal to use since both contain a good variety congruent and incongruent combinations.

The first condition consisted of 12 congruent restricted collocations (CRC). These CRCs were chosen using Howarth's collocational continuum as a reference and the items were also checked for lemmatized frequency (FQ) of occurrence in the Corpus of Contemporary American English (COCA). To control for congruency Korp – the corpus infrastructure of Språkbanken (KORP) - was consulted and all the items were cross-checked with their counterpart in Swedish. All the items were checked with both the verb and object noun lemmatized and with all possible combinations of pronouns, if they contained one. This was done with the intention that the items would be chosen in a way where the frequency of exposure to the items was considered as thoroughly as possible within the limitations of the corpus.

The second conditions were incongruent restricted collocations (ICRC). These ICRCs were controlled in the same manner as the CRCs. However, instead of making sure that they were congruent with KORP I used it to make sure that the items translated word-for-word did not exist in Swedish. This step was taken to make sure that there would be no cross-linguistic influence affecting processing in a similar way as congruent items.

The third conditions were congruent figurative idioms (CFI). These items were carefully chosen with the same criteria as the CRCs. They were also controlled for congruency in Hübinette & Odenstedt's book *Ord och Inga Visor*, which is a book containing 2000 Swedish idioms translated to English. Since the book only contained 2000 idioms I could use it for the items that existed in the book. The items not found in the book were therefore controlled with

KORP, but in a deeper way where the items were crosschecked with sentences in the corpus containing their congruent Swedish version making sure they were used idiomatically.

The fourth group of items were incongruent figurative idioms (ICFC). These items were controlled in the same manner as the ICRCs.

When all the items were controlled for frequency and for length I could move on and weed out the items that were less similar in length and frequency. During this process I decided that a total of 12 items per condition would be sufficient to run the experiment. When the 12 items for each condition were finalized I started working on the baseline items. These items were constructed in a way where they would be slower to process since they were word combinations not existing in the corpora which makes them highly unlikely to have occurred in speech or text. An example of this type of item is *sail a cougar* which does not exist in the COCA. I also checked if its Swedish word-for-word version, *segla en puma*, existed in the Swedish corpus KORP; it did not. For lemmatized frequency and length see Table 2.

Table 2. Means for experimental items (standard deviation)

Condition	CRC	ICRC	CFI	ICFI	Baseline items
FQ COCA	2.6 (0.7)	2.4 (0.6)	2.5 (0.3)	2.2 (0.4)	0
FQ KORP (congruent items)	2.8 (0.8)		2.9 (0.5)		
Length	10.5 (1.4)	10 (1.4)	10.6 (0.6)	10.0 (1)	10.1 (1.4)

Note. All COCA & KORP values are log normalized and rounded up to one decimal. Length is rounded up to one decimal.

An attempt was made to keep the mean frequencies and lengths of the word combinations in the respective conditions as close as possible. A one-way ANOVA comparison of the mean values shown in Table 2 revealed that for word frequencies (log-normalized), there was no difference, F (3, 44) = 1.21, p = .316. Please note that the baseline items were excluded from this analysis, as they were made-up, nonsensical word combinations. A similar ANOVA analysis for word length showed no difference either, F (4, 55) = 0.91, p = .465. Thus, the items in the critical conditions do not statistically differ in terms of mean frequencies and length.

When all the items were set they were contextualized by putting them into short sentences. This resulted in a list of 60 sentences all containing test items. The sentences were constructed in a manner where the test items appeared second in the SPRT (see Figure 1). I initially intended to control for spillover effect (see: Jiang, 2012, p. 174-75). This would mean that I would have to control and measure reaction time (RT) on the segment following the stimuli as well as the stimuli itself. However, due to the limited timeframe of the study I had to leave that out in the current experiment.

To control that all the participants understood and were reading properly we introduced 20 comprehension questions. These consisted of simple YES-or-NO questions that related to the sentence that the participants had just read. An example of sentence and a comprehension question that followed used in the current study is: [sentence] *Jim started to grow a beard to change his appearance.* [comprehension question] *Did Jim stop shaving?* In total, 20% of the sentences were followed by comprehension questions. According to Jiang (2012) this is the most widely used method to control for comprehension in SPRT research. Ideally one would have a comprehension question for all the sentences. However, if this makes the experiment last for too long it is also satisfactory to use less. Researchers sometimes use as few as one question per five stimuli (p. 176).

After all the sentences and the control questions were deemed grammatically correct and spellchecked, we moved on to creating filler items for the experiment. The 60 sentences which were used as fillers were added to make sure the participant's reading behavior did not get affected as they were administered the SPRT. According to Jegerski (2014) there is no clear percentage of filler sentences in an SPRT experiment decided by the research community, but that a minimum of 50% of the total number of stimuli is recommended to achieve an adequate amount of non-target sentences to avoid effects on the results (p. 32; see also: Jiang, 2012, p. 175-76). (For detailed information on items, sentences and comprehension questions see Appendix 1 & 2).

3.2 Participants

The participants of this study included 15 university students of English at Lund University with Swedish as their first language (L1). They all lived in Sweden at the time of the study, hence they all used Swedish on a daily basis. According to Jiang (2012) there are RT experiments published

where only 12-20 participants were tested. He states that the "RT data does not change significantly after a certain number of participants are tested" (p. 46). He continues to state that adding participants after a certain amount of participants have been tested does not alter the results significantly. Therefore, the 15 participants tested in this study were considered enough to arrive at stable results. Before the participants were administered the self-paced reading task (further described below), they were all asked to evaluate their English proficiency. This was done by administering a questionnaire where they were given a self-rating scale between 1-10 where one equaled no proficiency and ten native-like proficiency.

Adding a proficiency test was deemed logistically difficult because of the fact that the SPRT was quite extensive. Thus, as a second best alternative, self-reported values were relied on in the present study. Gyllstad (p.c.) found that self-reported proficiency correlated with vocabulary size scores at .46 for non-native speakers of English in Gyllstad & Wolter (2016). Another factor to why I used the self-evaluation method was due to the limited timespan and proportions of the study. The participants were also asked to give account of as to what extent they had been students of English as a second language and what their age of onset of acquisition (AoA) was. I worked with a definition of AoA where the time of the AoA is whenever the participants were first exposed to the English language i.e. if they watched TV shows or movies that were in English. The mean AoA was 4.3 years of age, and the mean amount of time the participants had been instructed in English in school was 10.9 years. This information was collected in a questionnaire before each participant was administered the self-paced reading task (SPRT). For detailed results of the questionnaire see below in Table 3.

Table 3. Means of Participants' Bio and Proficiency. (standard deviations)

Variable	*Participants (n = 15)*	
Age (range 20-30)	22	(2.5)
Gender (m/f)	7/8	
Self-Rated English Speaking (1-10)	7.3	(1.3)
Self-Rated English Listening (1-10)	8.5	(1.2)
Self-Rated English Reading (1-10)	8.3	(1.1)
Self-Rated English Writing (1-10)	7.7	(1)
AoA (years)	4.3	(2.5)
Years of Instructed L2 Learning	10.9	(3)

Note. English proficiency self-ratings are based on a 1-10 scale (1 = none, 10 = native-like) All values are rounded up to the closest one decimal.

In most studies where RT-research is done on non-native speakers a control group of native speakers is used to further validate the results (e.g., Gyllstad & Wolter, 2016; Wolter & Gyllstad 2011, 2013; Yamashita & Jiang, 2010). Due to the limited timeframe and the limited resources of the present study I only had the possibility of testing the non-native speakers.

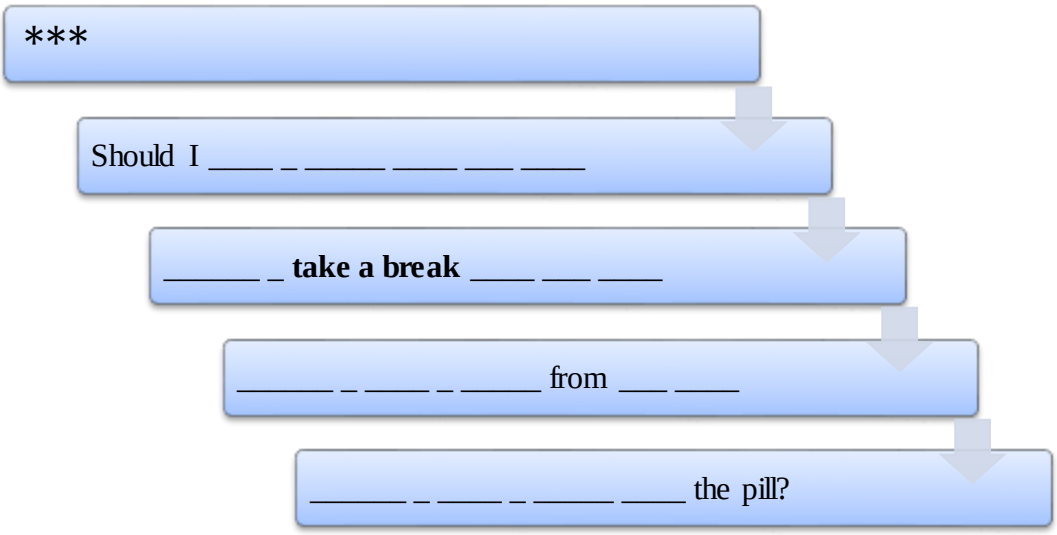
3.3 Procedures

In this study a self-paced reading task (SPRT) was administered to control for differences in RT for different word combinations, more specifically between CRCs, ICRCs, CFIs and ICFIs. Using SPRTs in linguistic research is a commonly used technique for testing reaction time. According to Jiang (2012), a SPRT is built on stimuli consisting of sentences presented word-by-word or (as in the current study) as segment-by-segment. To navigate an SPRT a participant is asked to click a button, either on the keyboard of the computer on which the SPRT is run, or (as in the current study) on a computer mouse connected to the computer running the SPRT. What I measure in an SPRT is the RT which is the time it takes for the participants from the moment they are presented with the item until they press the button to proceed. (p. 170-171)

The experiment was administered using DMDX software. DMDX is software developed by K. Forster & J. Forster at the University of Arizona. When using DMDX, which is a Windows based display program that measures RT with millisecond accuracy, one will have to construct an item file that is readable by the software (Forster & Forster, 2003).

The actual item file that was used in this study was written in such a way that every participant was exposed to the same sentences and control questions. However, the order in which these items and the accompanying control questions occurred on the screen was randomized by the software. This was done in an attempt to minimize the risks of getting similar results due to the order in which the items were presented. Drawing on Jiang's (2012) moving window model (see: Jiang, 2012, p. 172) and on Jegerski's noncumulative model (see: Jegerski, 2014, p. 22-23), the test items were presented segment-by-segment with the previous segment disappearing when the next appeared (visualized in Figure 1), and before the first segment was presented to the participants there was a focal point (also visualized in Figure 1) to ensure that the participants were looking at the right area of the screen when the first segment was presented. Each sentence was followed by a comprehension question or by a preparatory question (READY?).

Figure 1. *Visualization of presentation of sentences in DMDX.*



Note. The test item is the second segment in order in the sentence (bold font in the figure, but not in experiment).

3.4 Data analysis

Upon analyzing the RTs from each participant I divided the results into 5 different brackets, one for each condition and one for the control items. I then proceeded to calculate the mean RT from each condition and for the control items. When the mean was calculated I used the data to create a mean for each condition calculated on the combined results of all 15 participants, one for each condition, and compared those to each other. None of the 60 fillers were analyzed. Ideally one would control for spillover effects in the item following the test items, but because of the limited time frame of the study this was not done. Furthermore, the experiment design did not allow for doing so in an efficient manner. I controlled the error rate of the participant's answers on the control questions and made sure I did not have to omit anyone due to their lack of focus and ability to understand what they read. Because of a technical error in the programming of the DMDX file the CRC condition had one item omitted. For more information on the items used in the experiment see Appendix 1.

4. Reading Time Results

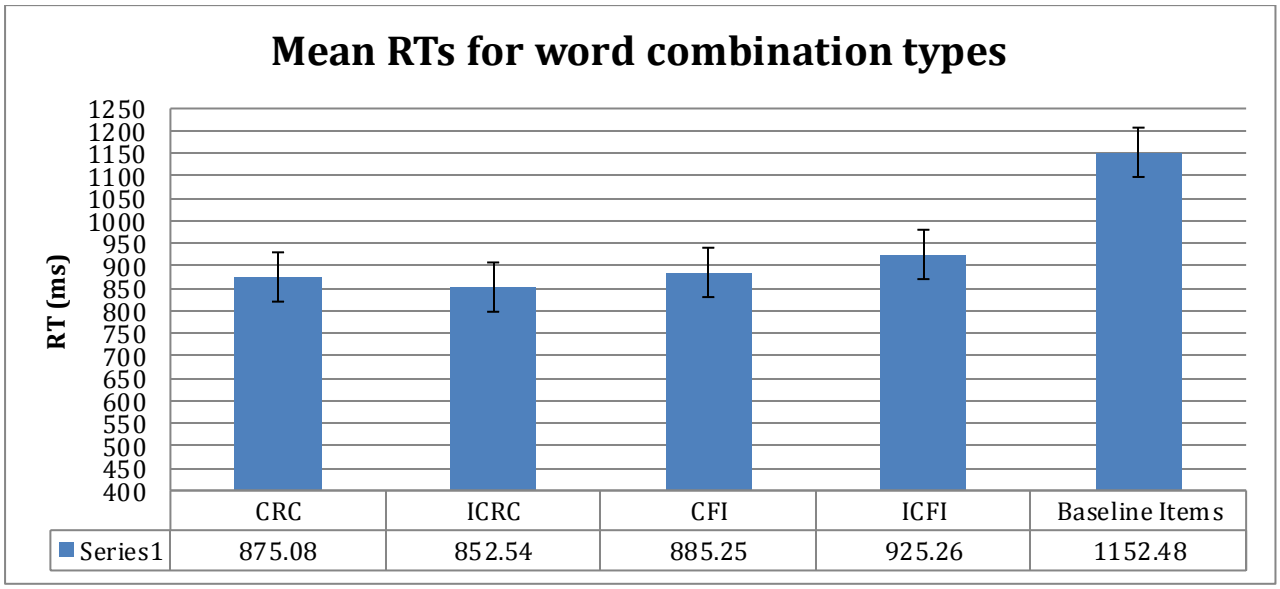
The mean RTs and the accuracy of how the participants answered the comprehension questions are presented in Table 4. The RT-results are also visualized in Figure 2. In Figure 2 and in Table 4 we can see that the mean reading times for the different conditions differ slightly numerically. We can also, however, notice that there is a slight trend of the RT going up as the conditions change from the RC-conditions to the FI-conditions. A repeated measures ANOVA was used to analyse the data from the self-paced reading experiment. Log-normalized values were used. In the ANOVA, the assumption of sphericity was violated, and therefore a Greenhouse-Geiser estimate was used to adjust the degrees of freedom for the *F*-ratios. The ANOVA results show that reading times were significantly affected by the type of word combination, F (1.88, 26.30) = 10.39, $p < .01$, $\eta = .43$. Subsequent pairwise comparisons revealed that the only mean reading times that were different from one another were the baseline items compared to the CRC (at $p = .042$), ICRC (at $p = .011$), and CFI (at $p = .008$) items. The ICFI items, however, were not significantly different from the baseline items (at $p = .074$), although they approached significance. For no other pairwise comparisons were statistically significant differences observed. This means that none of the collocation and idiom conditions differed from each other significantly in terms of mean reading times.

Table 4. Mean reading times and (standard deviations) in the five conditions, and accuracy answering control questions.

Condition	Participants	(n = 15)
CRC	875.08	(315.70)
ICRC	852.54	(309.87)
CFI	885.25	(333.96)
ICFI	925.26	(358.15)
Baseline	1152.48	(438.29)
Accuracy	87%	(11%)

The accuracy rate of the comprehension questions being as high as 87% shows that the participants were focused throughout the experiment. It also shows that the participants understood what they read and that their proficiency in English was adequate to participate in the present study.

Figure 2. *Visualization of mean Reading Time results in the 5 conditions presented in milliseconds.*



5. Discussion

The purpose of this study was to collect data regarding processing of word combinations with an approach aimed at bringing in more ecological validity by using contextualized rather than decontextualized items. Then compare it to the results of earlier studies, such as Wolter & Gyllstad (2011, 2013) and Gyllstad & Wolter (2016), that show that there is a difference in processing time depending on what condition an item adheres to in Howarth's continuum and if the item is congruent or not. Do these results hold up when test items are contextualized?

In this section I will discuss the layout of the current experiment and compare it to earlier studies that researched processing of word combinations. Then I will discuss why contextualizing the items might have had an effect on the processing time. I will also put congruency against frequency to discuss whether frequency has more effect on processing time for learners that have studied English on a higher level and if when a person reaches a certain proficiency level, frequency over-rides congruency.

First we will look at the layout of the experiment itself. One of the challenges when contextualizing your test items is predicting how the context itself affects the processing cost for the test subjects. Something to consider in future studies is making the context perfectly

homogeneous by controlling the pre- and post-test item sequences for frequency and length. Doing so would make it possible to measure spill-over effect.

This was something the present study did not consider. If the context itself (the pre- and post-test item sequences) was controlled for frequency, length and all the parameters that the test items were controlled for, my hypothesis is that the results would be easier to interpret. One could, for example, easily control for spillover effects in a way that the current study was not able to do. This is a factor that will have to be further researched. It is clear that when looking at the results of previous studies (e.g., Yamashita & Jiang, 2010; Wolter & Gyllstad, 2011, 2013; Gyllstad & Wolter, 2016) there is a difference in how proficiently different types of word combinations are processed and that congruency has an effect on processing cost. However, none of the previous studies have had the test items contextualized. It presented a challenge that I did not predict upon deciding to contextualize the items, the effect of context on processing time. This is something that needs to be further discussed and ultimately taken into consideration when conducting experiments where the test items are contextualized. It seems we should, when contextualizing our test items, also take into consideration and control the context.

Another factor that might have an effect on the RT could be the presentation technique. In Tremblay et al. (2011) 3 different methods of presentation were used. In the first experiment they used a word-by-word presentation technique. The results showed online processing advantage for lexical bundles over non-lexical bundles. However, the difference was close to insignificant. This could according to Tremblay et al. have to do with the unnaturalness of reading a text word-for-word. With that in mind they designed their second experiment which was built on portion-by-portion (in the current study referred to as segment-by-segment) presentation method. In the third experiment they presented the items sentence-by-sentence. When we look closer at Tremblay et al. (2011), where full sentences were presented in one of their experiments and all the sentences were controlled for length and frequency, we can show that an effort was made to control the context itself. This is something that should be applied to all experiments where context is a factor. However, in the present study no such effort was made. Ideally, when context is a part of the experiment, one should aim to run parallel experiments like Tremblay et al. did in their study to be able to summarize and compare results of the same test items depending on presentation. If such an effort is made it would further strengthen the validity of the results.

Another factor to take into consideration is the participants' familiarity with the test items (subjective frequency). In the present study one limitation is the use of corpora. According to Wolter & Gyllstad (2013) there is a limited unanimity between corpus frequency and actual frequency of input. There is also the language user themselves to consider. Any two users are likely to have different activities that they are part of and other cultural differences that affect their frequency of input in a variety of ways (Wolter & Gyllstad, 2013, p. 457). One way to control that could be to run a familiarity test before the items were chosen for the study. Wolter & Gyllstad (2011) did such a subjective familiarity rating test which was administered to a group of 20 participants (see. Wolter & Gyllstad, 2011, p. 436-437). Gyllstad & Wolter (2016) also mentions the significance of frequency and how familiarity (how well known a PU is to an individual) adds a processing advantage to the PUs when presented to the participants (p. 316). Gyllstad (2016) which is a study partially similar to the present study, also administered a familiarity rating test where the participants rated items in the administered experiment. This is something that ideally should be controlled for in future studies of this kind.

Another result that Gyllstad (2016) yielded was that the difference in processing between the different conditions in Howarth's (1996, 1998) was insignificant for the NNS group of the Gyllstad (2016) study. Similar results were found in the current study. Something that was surprising in the current study, however, was that congruency did not have a positive effect on the processing cost of the word combinations. This is something that does not concur with earlier results from studies like Wolter & Gyllstad (2011, 2013) where congruency had a boosting effect in the processing cost of word combinations. This could be interpreted as the more proficient the NNS is overall, the less processing cost there will be independent the type of word combination. This could also explain why the results for the NNS in the current study did not show any significant difference in processing between the congruent and the incongruent word combinations. The fact that neither the difference between congruent or incongruent items nor semantic transparency was statistically significant might also have something to do with the fact that the items in the current study were contextualized. Since there are, to my knowledge, no similar studies where contextualized word combinations have been researched, we have nothing to compare to. Hence, I cannot state with any certainty that contextualizing word combinations has an equalizing effect on processing between the Howarthian conditions or between congruent and incongruent word combinations. This is something that needs to be further looked into before

any assumptions can be made. Another thing to factor in is the fact that the participants in the present study had a very low error rate. A mean accuracy as high as 87% could have an effect on the results as it shows a high general ability in the participant group. The results of the self-rated proficiency level of the participants were also high. What this seems to show is that the participants' previous knowledge of English collocations and overall English language proficiency was high enough to possibly decrease the effect normally observed when using congruent items vs. incongruent items in a study. However, this is something that would have to be further researched, and it could be done by administering a proficiency test in a way similar to previous studies such as Wolter & Gyllstad (2011, 2013) to get a comparable result. A standardized English proficiency test containing phraseological proficiency testing would benefit future similar research.

6. Conclusion

The aim of this study was to investigate if the results from previous studies researching the processing of different word combinations held up when contextualizing the word combinations tested. The congruency effect was absent in the results as well as the difference in processing cost between the different conditions. This could have something to do with the way the word combinations were presented and the fact that they were contextualized. It seems that when processing RCs and FIs online the context possibly adds an equalizing effect to the processing ability and evens it out between the conditions. To properly test if this equalizing effect exists, I recommend that future studies take into consideration the fact that contextualizing the test items might affect the RT results. I would like to see an experiment where the same PUs were tested both contextualized and decontextualized to further research this effect. I will conclude this essay the way I started, with a couple of unanswered questions which I hope future research will consider trying to answer.

1. In what way does the contextualization of word combinations affect the processing speed?
2. Does context equalize the congruency effect and the difference in processing cost between the different phraseological (Howarthian) conditions?

References

Allerton, D., Nesselhauf, N., & Skandera, P. (Eds.). (2004). *Phraseological units: Basic concepts and their application*. Basel: Schwabe.

Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocations? *System, 21*(1), 101-114. doi:10.1016/0346-251x(93)90010-e

Barfield, A., & Gyllstad, H. (Eds.). (2009). *Researching collocations in another language: Multiple interpretations*. New York: Palgrave Macmillan.

Borin, L., Forsberg, M., & Roxendal, J. (2012). Korp – the corpus infrastructure of Språkbanken. Available online at https://spraakbanken.gu.se/korp/

British National Corpus (BYU-BNC). (n.d.). Retrieved December 11, 2015, from http://corpus.byu.edu/bnc/

Cowie, A. P. (1994). Phraseology. In R. E. Asher & J. Simpson (Eds.), *The Encyclopedia of Language and Linguistics* (pp. 3168-3171). Oxford: Pergamon Press.

Cowie, A. P., & Mackin, R. (1983). *Oxford dictionary of English idioms* (Vol. 2). Oxford: Oxford University Press.

Davies, Mark. (2008-) The Corpus of Contemporary American English: 520 million words, 1990-present. Available online at http://corpus.byu.edu/coca/.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*(1), 116-124. doi:10.3758/bf03195503

Granger, S., & Paquot, M. (2008). Disentangling the phraseological web. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 27-49). Amsterdam: John Benjamins.

Gyllstad, H. (2016, April). *Comparing L1 and L2 phraseological processing: Free combinations, collocations and idioms.* Poster presented at the AAAL conference in Orlando, FL, USA, 9-12 April, 2016

Gyllstad, H. & Wolter, B. (2016). Collocational processing in light of the Phraseological Continuum Model: Does semantic transparency matter? *Language Learning, 66*(2), 296–323. DOI: 10.1111/lang.12143

Howarth, P. (1996). *Phraseology in English academic writing: Some implications for language learning and dictionary making*. Tübingen: M. Niemeyer.

Howarth, P. (1998). Phraseology and Second Language Proficiency. *Applied Linguistics, 19*(1), 24-44. doi:10.1093/applin/19.1.24

Hübinette, L., & Odenstedt, B. (1988). *Ord och inga visor: 2000 svenska idiom i engelsk översättning*. Lund: Studentlitteratur.

Jegerski, J. (2014). Self-Paced Reading, In Jegerski, J. & VanPatten, B. (Eds.), *Research methods in second language Psycholinguistics,* (pp. 20-49). New York, NY: Routledge.

Jiang, N. (2012). *Conducting reaction time research in second language studies*. New York: Routledge.

Nesselhauf, N. (2003). The Use of Collocations by Advanced Learners of English and Some Implications for Teaching. *Applied Linguistics, 24*(2), 223-242. doi:10.1093/applin/24.2.223

Oxford English Dictionary Online. (2000). Oxford: Oxford University Press. Retrieved December 11, 2015, from http://www.oed.com/

Pearson Education Limited, (2014). *The Longman Dictionary of Contemporary English* (Sixth Edition) [Mobile application software]. Retrieved from http://itunes.apple.com

Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing Advantages of Lexical Bundles: Evidence From Self-Paced Reading and Sentence Recall Tasks. *Language Learning, 61*(2), 569-613. doi:10.1111/j.1467-9922.2010.00622.x

Wolter, B., & Gyllstad, H. (2011). Collocational Links in the L2 Mental Lexicon and the Influence of L1 Intralexical Knowledge. *Applied Linguistics, 32*(4), 430-449. doi:10.1093/applin/amr011

Wolter, B., & Gyllstad, H. (2013). Frequency Of Input And L2 Collocational Processing. *Studies in Second Language Acquisition, 35*(03), 451-482. doi:10.1017/s0272263113000107

Yamashita, J., & Jiang, N. (2010). L1 Influence on the Acquisition of L2 Collocations: Japanese ESL Users and EFL Learners Acquiring English Collocations. *TESOL Quarterly, 44*(4), 647-668.

Appendix 1 – Experiment items and Sentences used in the SPRT Experiment

	Congruent RC			**Test Item's Log Normalized FQ**	**Test Item's Length w/o space**
Should I	*take a break*	from the	pill?	*3.708675793*	*9*
Everyone needs to	*make money*	and I'm	not against that.	*3.731105051*	*9*
The best way to	*keep the peace*	is to	redefine our terms.	*2.727541257*	*12*
She actually did	*get a degree*	from	the Sorbonne?	*2.348304863*	*10*
The person that will	*save the world*	is the	one that shares the pain.	*2.785329835*	*12*
The first thing I do is	*book a table*	at the	restaurant across the street.	*1.255272505*	*10*
Film watchers	*take notice*	of any	work by George Lucas.	*2.964730921*	*10*
A girl can	*show interest*	first,	but waits for me to take it from there.	*2.477121255*	*12*
We are willing to	*make a bet*	that	they will win the race next Friday.	*2.068185862*	*8*
I had to	*call a meeting*	during	the game in Texas.	*2.23299611*	*12*
It is a golden opportunity to	*give advice*	about	good nutrition.	*2.73239376*	*10*
I simply had to	*show progress*	on my	dissertation.	*1.949390007*	*12*
	Incongruent RC			**Test Item's Log Normalized FQ**	**Test Item's Length w/o space**
She does not need to	*throw a fit*	to get	her way.	*1.977723605*	9
They don't even	*pay attention*	to you	sometimes.	*3.860577551*	9
Who would	*play a joke*	on me	on my birthday?	*1.785329835*	9
Ask a coworker to	*lend a hand*	when	projects pile up.	*2.51054501*	13
They planned to	*bring peace*	to a	troubled country.	*2.716837723*	10
In order to	*break a record*	you have to	be very good at what you do.	*1.69019608*	12
Johanna arranged to	*pay a visit*	early	in the afternoon.	*1.949390007*	8
Some people	*take comfort*	in the	argument that torture never works.	*2.572871602*	11
They seem to	*have power*	over me	in ways they do not know.	*2.931966115*	9
Jim started to	*grow a beard*	to change	his appearance.	*2.117271296*	10
As an infant, she would	*catch a cold*	every time	the wind changed.	*2.1430148*	10
I was even afraid to	*keep a diary*	because	everyone would look over my shoulder.	*2.264817823*	10
	Congruent FI			**Test Item's Log Normalized FQ**	**Test Item's Length w/o space**
She had to move fast not to	*get cold feet*	before	she had asked him out.	*2.012837225*	11
A local charity is working to	*build bridges*	between	different ethnic groups in the area.	*2.394451681*	12

Cont. of Appendix 1

It'll be better if we	clear the air	before	we move on with the project.	*2.424881637*	11
I am doomed to	tread water	until	I get an opportunity to try for another job.	*2.494154594*	10
The Republicans will surely	gain ground	in the	Southern states.	*2.669316881*	10
Most of my talent will	gather dust	because	I never really have an opportunity to perform.	*2.190331698*	10
He tried to	*break the ice*	but she	was way too cold.	*2.354108439*	11
Sometimes people think they	*have a point*	when all	they really have is confusion.	*3.036229544*	10
I hope I don't	*drop the ball*	before	I get this project done.	*2.491361694*	11
He came to	*save the day*	while	everyone else was sleeping.	*2.537819095*	10
They had to	*follow suit*	otherwise	the gang members would get suspicious.	*3.084933575*	10
I will not	*lift a finger*	until I	get what I want.	*2.220108088*	11
	Incongruent FI			**Test Item's Log Normalized FQ**	**Test Item's Length w/o space**
I never seem to	*catch a break*	with	this guy.	*2.025305865*	11
They didn't tell me they would	*cut corners*	building	this house, yet they did.	*2.527629901*	10
I told him not to	fan the flames	any more	or the situation would get out of hand.	*2.214843848*	12
Don't you dare to	*show your face*	here	ever again.	*2.378397901*	12
Another one will	*bite the dust*	if we	don't stop this madness.	*2.045322979*	11
Don't come over here and	*pick a fight*	because	it will end very badly for you.	*2.382017043*	10
She likes to	*make a scene*	whenever	she gets bored.	*2.209515015*	10
You can go ahead and	*take your pick*	from	this huge box of chocolates.	*2.421603927*	12
They always seem to	*cross swords*	whenever	she is around.	*1.681241237*	11
That should	*do the trick*	and if	it doesn't I don't know what will.	*2.926342447*	10
They really would	*drop a bomb*	if they	told everyone that the mayor was resigning.	*2.004321374*	9
John had to	*drop a dime*	on his	dealer if his sentence were to be reduced.	*1.531478917*	9
	Baseline Items			**Test Item's Log Normalized FQ**	**Test Item's Length w/o space**
The bird went to	fry grass	in the	hills beyond the trees.	0	8
I have to	buy a flea	if I'm	going to make it on time.	0	8
Some people have to	swing cattle	to get	their cars started.	0	12
It is hard to feel when you	grasp spiders	in another	person's house.	0	10
He had come to	smash an elk	for the	people in the auditorium.	0	10

Cont. of Appendix 1

Someone had told him to	dig a hose	to make	up for lost time.	0	8
John had a dream to	sail a cougar	across	the Atlantic.	0	11
It is sometimes hard to	conduct a moose	when	the roads are icy.	0	11
The black man tried to	fake a goose	but his	wife caught him in the act.	0	10
Anywhere where you can	advise fish	there	is usually a desert.	0	10
Freedom is dependent on if people	create banana	well	enough together.	0	12
When I	blend cinema	I usually	leave out the eggs.	0	11

Appendix 2 – Sentences & Comprehension Questions

S: Should I take a break from the pill? Q:Is this a question about birth control?
S: Johanna arranged to pay a visit early in the afternoon. Q: Was Johanna going to visit in the afternoon?
S: He tried to break the ice but she was way too cold. Q: Was she a warm and open person?
S: They had to follow suit otherwise the gang members would get suspicious. Q: Did they do what the gang member did?
S: Another one will bite the dust if we don't stop this madness. Q: Will someone die unless the madness stops?
S: We are willing to make a bet that they will win the race next Friday. Q: Do they think that they will win the race?
S: Thy planned to bring peace to a troubled country. Q: Did they want to bring war to a peaceful country?
S: Jim started to grow a beard to change his appearance. Q: Did Jim stop shaving?
S: As an infant, she would catch a cold every time the wind changed. Q: Did she get sick often as a child?
S: She would have to move fast to not get cold feet before she asked him out. Q: Was she nervous to ask him out?
S: I am doomed to tread water until I get an opportunity to try for another job. Q: Does he feel stuck in his job?
S: The republicans will surely gain ground in the Southern states. Q: Will the Republicans do well in the Southern states?
S: Don't you dare show your face here ever again. Q: Does he want to see him again soon?
S: They didn't tell me they would cut corners building this house, yet they did. Q: Did they take shortcuts while building the house?
S: You can go ahead and take your pick from this huge box of chocolates. Q: May they choose whichever piece of chocolate they want?
S: She actually did get a degree from the Sorbonne. Q: Did she go to school at the Sorbonne?
S: Film watcher take notice of any work by George Lucas. Q: Were the film watchers fans of George Lucas?
S: They always seem to cross swords whenever she is around. Q: Did they stay good friends when she was around?
S: John had to drop a dime on his dealer if his sentence were to be reduced. Q: Did John have to give up his dealer?
S: I was ever afraid to keep a diary because everyone would look over my shoulder. Q: Did the person feel comfortable writing in their diary?

Appendix 3 – Filler Items

Sometimes things	*that seem very*	hard are	actually easy.
When I try to make up	*sentences my*	mind seems to	not work very well
I guess the only	*way to pull this off*	is to	steal all the items.
Ancient cultures	*sometimes had*	very strange customs	including human sacrifice
Some people drink	*tea in the morning*	and other people	drink coffee.
The Y2K was	*a phenomenon*	that scared	a lot of people.
The feminist party has	*a lot of followers*	who are	women
Some politicians tell	*the truth*	but most of them	do nothing but lie
Owning a car is	*something that will*	cost more	than you can imagine
In Los Angeles there	*are a lot of*	different areas	each with their own dialect.
The magicians apprentice	*did not*	make it	as a wizard on his own.
A lot of Americans	*get their news*	form sources such	as Fox News
I like to go out	dancing on	Friday	nights
It was the first time	*she had gotten*	a spray tan	on a Wednesday.
Not a day goes by	*that I don't*	wonder if	I could've done something differently.
I helped her move in,	*said goodbye*	to the girl	and drove back.
I went over	*to see*	Marina two or three	times a week.
She and the child	were living	in a hippie	commune in New Mexico.
I got the idea	*that everybody*	in the world	was doing it.
But he didn't know	*how to make*	a woman	happy.
Henry poured a drink	*and looked out*	the window	at the hot and bare street.
It had	*been a long haul*	and he was still	up against the wall.
Death was next,	*death was*	always	there.
There was a photo	*of him,*	dead, right	after the fix.
Humanity may	*have been*	suffering	but not him.
I smoked cheap cigars	*and drank*	and listened to	classical music until dawn.
I put my	*mouth and body*	against	hers.
She worked at	*the mound of clay*	with a wooden	tool tipped with a loop of wire.
She held the	*beer bottle*	out at arm's length	with one hand.
We had sandwiches ,	*pickles, chips*	and two different kinds	of soft drinks.
He had a long	*slim neck and*	a gold earring hung	from the left ear.
She came back	*with a pencil*	and a paper	in her hands.
Then she goes	*to bed and*	eat cookies	and watches TV.
I waited until she was finished,	*and then I*	leaned across the	table and gave her a kiss.
It was good	*to have a place*	to go when	things went bad.
I began	*receiving letters*	from a girl	in New York.

Cont. of Appendix 3

I answered the	*letters*	and was always happy	to find one of hers in my mailbox.
Most people	*are much better*	at saying things	in letters.
They are wild,	crazy, like some	animal peering out	of the forest on fire.
We had a second	drink and	went down	for the luggage.
The next morning	we all sat	around the	campfire.
I went	to bed	and thought	about her hair.
Lydia picked up	the typewriter	and ran out into	the center of the street with it.
I parked a block	below Santa Monica and Western,	got out and	looked around.
We decided to	spend the holidays	together at	my place.
Thirty or	forty minutes later	there was	a knock on my door.
The place I	was living in	at the time	did have some qualities.
Most people left	notes in the	empty bottles	explaining what they wanted.
I've read all	your essays about	the presidents	visit last week.
He was small and	thin with barely	any hair on	his head.
Then I caught	a glimpse of	my reflection	staring in at them.
I was like some kind	of jungle animal	drawn to the light	and looking in.
I drained my	glass and got	up and found another	bottle.
You are	becoming what	you've	always hated.
The first realtor	we stopped	at was	in Santa Monica.
He thought	that we would	waste his	valuable time.
I drove the car along,	hardly knowing	where I	was going.
I was hungover	and needed	a shave	that morning.
I paid the barkeep	and we backed off	our stools	and moved toward the door.
Then I would sleep	with the	dark blue	walls, healing.

Appendix 4 – Individual mean RTs for the experimental conditions

	P1	P2	P3	P4	P5
Congruent RC	642.40	530.77	798.58	876.89	634.69
Incongruent RC	724.82	470.03	654.43	997.10	633.77
Congruent FI	874.57	428.22	597.65	982.29	735.24
Incongruent FI	865.52	482.55	669.25	1064.07	650.74
Baseline	859.98	587.74	895.09	1204.95	1898.43
	Accuracy				
Comprehension	90%	85%	80%	85%	95%

P6	P7	P8	P9	P10	P11	P12
1271.20	1103.66	925.93	1769.37	1052.62	505.34	648.50
1242.13	1168.03	873.66	1653.12	1023.36	442.23	718.95
1318.39	1008.16	968.32	1673.77	1271.92	404.85	581.16
1468.78	1193.56	1134.82	1721.57	1248.69	400.61	686.98
1513.41	1238.66	1324.44	1968.89	1662.86	585.35	797.82
50%	95%	80%	85%	85%	95%	100%

P13	P14	P15	Mean TOTAL RT	SD
729.99	821.89	814.34	875.08	315.70
682.32	673.27	830.89	852.54	309.87
816.05	689.19	928.98	885.25	333.96
794.37	711.75	785.58	925.26	358.15
1150.89	625.44	973.26	1152.48	438.29
			Mean	**SD**
95%	90%	90%	87%	11%

Appendix 5 – Participant Questionnaire

Research Participant Information and Consent Form

1. EXPLANATION OF THE RESEARCH and WHAT YOU WILL DO

You are being asked to participate in a research study of how learners of English process English words and word combinations.

There are 2 things I will ask you to do:

1) provide language background details in a questionnaire
2) do a reading task (on a computer)

The whole test is expected to take approximately 20 minutes.

2. YOUR RIGHTS TO PARTICIPATE, SAY NO, OR WITHDRAW:

Participation in this study is voluntary. You have the right to say no. You may change your mind at any time and withdraw.

Your signature below means that you voluntarily agree to participate in this research study. Your data may be used in a report/article but your name or other details that could identify you will never be mentioned; you will remain anonymous.

_____ _____

Signature Date

Questionnaire

1. Bio data
Today's date (year/month/day): ___ ___ ___ ___ / ___ ___ / ___ ___

Name (surname + given name): _____
Gender: ☐ Male ☐ Female ☐ Other
Personal code number (personnummer): ___ ___ ___ ___ ___ ___ - ___ ___ ___ ___
E-mail address: _____
Phone: _____

I am: ☐ Left-handed ☐ Right-handed
I have normal or corrected-to-normal (glasses or lenses) vision ☐ yes ☐ no

2. Language background
A. Please list the languages you know in order of acquisition (your native language first)

1	2	3

3. English language proficiency

Here you will be asked to do a self-assessment of your **English** proficiency. On a scale of 1-10, state your own English proficiency in the four language skills by putting a cross in the corresponding box.

Skills	**None** **1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Native-like** **10**
Speaking										
Listening										
Reading										
Writing										

B. L2 / Second language:_____

1. When was your first exposure to this language? Age _____

2. Number of years of instructed learning (taught in school) Years ____ Months ____
Age period __ __ - __ __

3. Total time spent in a country where this language is the main language (y/m) __ __ / __ __

4. How often do you use this language?
☐ every day ☐ several times a week ☐ once a week ☐ once a month
☐ other: _____

C. L3 / Third language: _____

1. When was your first exposure to this language? Age _____

2. Number of years of instructed learning (taught in school) Years ____ Months ____
Age period __ __ - __ __

3. Total time spent in a country where this language is the main language (y/m) __ __ / __ __

4. How often do you use this language?
☐ every day ☐ several times a week ☐ once a week ☐ once a month
☐ other: _____