Dual Lexical Activation in Bilinguals: a Tale of Two Verbs

Pieter Hussaarts
MA Thesis

Written under the supervision of Sergey Avrutin
Second reader: Luisa Meroni

At Utrecht University
11-06-‘17

Abstract

This study investigates the lexical organisation in the bilingual brain by the hand of an information theoretical approach. Reaction times of English-Dutch bilingual participants in response to high and low entropy verb paradigms were measured and compared to the performance of monolingual speakers of English. By selecting verbs that have different entropy values across the bilinguals' languages, it was possible to measure if both language were activated during lexical decision making. The results seem to indicate that, in the bilingual brain, the inflected forms of English verbs and their Dutch counterparts are stored in a single paradigm that becomes fully active during lexical decision making.

1. Introduction

One of the largest uncharted territories within linguistics is perhaps the mental lexicon. Even though we are still very far from understanding how exactly this lexicon is organised, studies have begun to lift the very tip of the veil. The present study proposes an experiment that will contribute to the hand of linguistics in exposing the countenance of that elusive bride, focussing specifically on bilingual lexical organisation.

One aspect studies on bilingual lexical processing seem to agree on is that in many instances, the lexicons of both languages are activated simultaneously in production or comprehension tasks (Bijeljac-Babic, Biardeau & Grainger, 1997; Jared & Kroll, 2001; van Heuven, Dijkstra, & Grainger, 1998). However, most available literature focusses exclusively on the dual lexical activation of concrete nouns. To add to this, the current study will shift the focus to investigating the dual lexical activation of verbs. In Germanic languages, such as will be the subject of this study, noun systems are relatively simple as compared to verb systems. A verb only exists as an inflected member of a verb paradigm, which is assumed to be co-activated in its entirety (van Ewijk & Avrutin, 2016). The level of activation required to activate an entire verb paradigm of one language and co-activate the verb paradigm of another language is thus assumed to be more taxing than the dual activation of nouns; as such, the results might differ from previous findings for dual lexical activation of nouns.

However, as will become apparent from a review of the existing literature, studying the processing of verbs is a challenging task given the current experimental methods, hence, a different approach is required. This approach will be an information theoretical one, considering its usefulness in studying verbs has previously been attested (cf. van Ewijk & Avrutin, 2016).

As studies on bilingualism have not (widely) employed information theory yet, the goal of the current study will be two-fold. It aims to answer the following questions: Can

information theory be used as a predictive tool for bilingual studies? And can it be used to find a difference in the mental lexical organisation of monolinguals and bilinguals (i.e., will bilinguals show signs of dual lexical activation? And if yes, what does this imply for their mental lexical organisation?).

The subsequent section will review the literature relevant for this study, followed by the method that will be employed. Finally, the results will be presented after which the paper concludes with a general discussion and the conclusions.

2. Theoretical framework

2.1. Information theory

Information theory was originally proposed by Shannon (1948) as a means of quantifying the complexity of information. Though originally designed for the purposes of studying communication operations and signal processing, Shannon's theory has now found its way into many more fields, amongst which quantum computing, thermal physics and linguistics.

They key value of information theory is the ability to convert probability into information in bits (I) and, based on the distribution of the information load of each member in a system, calculate the entropy of the system (H), which serves as a measure of uncertainty. Higher entropy indicates that the distribution of the members in a system is relatively even, hence, there is more uncertainty. Lower entropy means that the members in a system are distributed more unevenly, hence there is less uncertainty.

To put it in simple terms; imagine you have two boxes, both containing an amount of apples and oranges. Box 1 contains 5 grapefruits and 5 oranges, box 2 contains 9 grapefruits and 1 orange. When taking a random fruit from the first box, you are maximally uncertain whether you will end up holding an grapefruit or an orange because the distribution of grapefruits and oranges is perfectly equal. This box thus has high entropy (in fact, as will be explained later, the entropy value of this box is actually the maximum).

When picking a random citrus from the second box, you will feel rather confident you come out holding a grapefruit as the distribution is very unequal, hence less uncertainty exists. This is an example of a low entropy system.

2.1.1. More on information theory

The above illustrates very shortly the essence of information theory. However, for the present study it is important to go into further detail about the precise calculations.

2.1.1.1. Information load

To calculate entropy, the first thing we require is the information load of each member of the set. The information load (I), expressed in bits, is a measure of how much information an individual item carries. The lower the probability of an item, the higher I in bits, i.e. the more information it is assumed to carry. It is calculated by taking the positive $\log_2$ of the probability of an item, as shown in (1), where *p* is probability and x represents the item.

(1) $I_x = -\log_2 p(x)$

By taking the minus $\log_2$ of the probability, the information is represented in bits. A classic method of illustrating the workings of information load and entropy is by taking the example of a coin flip. Assuming the coin has a 50/50 chance of coming up heads or tails, the probability of either option is .5. The information in bits is $-\log_2(.5) = 1$. Reversing this calculation would look like this: 2^{-1}. So essentially $-\log_2$ calculates the power to which 2 must be raised to obtain the given number.

The importance of the $\log_2$ transformation is that the probability of our coin toss is now represented in binary bits of information. These bits are the same as the bits that pertain to computers. In (2) below a representation of an 8-bit code unit is given.

(2)	a.	0	0	0	0	0	0	0	0
	b.	128	64	32	16	8	4	2	1

8 bits of information thus represent 8 units of binary information (1 or 0). If a 0 is set to one, this means that the number it represents (here indicated directly below in (2b)) is present in the information the combination of these bits represents. 8 bits of information can thus maximally represent 256 combinations (ranging from 0 (0.0.0.0.0.0.0.0) to 255 (1.1.1.1.1.1.1.1). The coin toss can end in either of two ways: heads (1) or *not* heads (0), hence it contains 1 bit of information.

The split with the binary notion in computers occur quickly when we introduce more complicated data. Any side of a fair die when rolled, for example, has a chance of 1/6 of coming up. This translates to roughly 2.585 bits of information. The floating point is introduced because information load represents *precisely* how many bits of information are conveyed. 2 binary bits can maximally represent 4 options (0 (0.0) to 3 (1.1)), whereas 4 bits can maximally represent 8 options (0 (0.0.0) to 7 (1.1.1)). Logically, as 6 options is somewhere between 4 and 8 options, its counterpart in bits is somewhere between 2 and 3.

2.1.1.2. Entropy

The above explained how to calculate the information in bits of a single item in a set. However, for the purpose of the current study the combined information in bits of each member of a set is of vital importance. This measure is called entropy.

Continuing on the example of the die, the entropy of the set of possible outcomes can be calculated using the formula below in (3):

(3) $$H = -\Sigma_{\varepsilon}\, p(x_{\varepsilon}) \log_2 p(x_{\varepsilon})$$

Example (3) states, '*H* is the sum of the probabilities multiplied the $\log_2$ of the probabilities'. For the die example that translates to (1/6) x $\log_2(1/6)$ for each side of the die. In plain terms,

the calculation concerns the sum of each item's information load times the fraction of the set that item accounts for. It is, then, no surprise that in this particular example the entropy in bits is the same as the information load of any given side of the die, as the calculation is the sum of six times one sixth of the information load. This, however, illuminates a fundamental principle of information theory: the maximum entropy of a set is achieved when all members are distributed evenly. This principle is very logical considering entropy is a measure of uncertainty: the most uncertainty is experienced when all possible events in a set are equally likely.

For entropy to decrease then, the distribution of its members needs to become more unequal. Imagine if our die were loaded on both the top (1) and the bottom (6); the possibility of either of these coming up increases, and the entropy decreases. Estimating 1 and 6 are twice as likely to come up in a roll, their probability would shift to .25, whereas the probability of 2 through 5 is now .125. The I of 1 and 6 is now $-\log_2(.25)= 2$, and $-\log_2(.125)= 3$ for 2 through 5. (4) illustrates how the entropy (*H*) is calculated from here.

(4) a. $H = .25 \times 2 + .25 \times 2 + .125 \times 3 + .125 \times 3 + .125 \times 3 + .125 \times 3$

b. $H = 2.5$

As mentioned in the previous section, the information load of a given item in the evenly distributed set, which in that case equals the entropy, was 2.585. By making the distribution more uneven in (4) the entropy lowered to 2.5.

Another important principle is highlighted here when looking at the individual probabilities and their respective information load: the lower the probability, the higher the information load, i.e., the more unlikely an event is, the more informative it is. When framed in an example, this is a logical assumption. Imagine Johnny threw a football on the roof 1000

times. 999 times, Mark caught the ball at the other end. In one particular instance, the ball had decided to cease abiding by the laws of physics and floated off. This singular instance certainly appears more informative than the other 999.

One final important characteristic is illustrated in (4): as the probability of an item or event decreases, the information load increases, but the result of multiplying the probability by the $\log_2$ probability (information load) decreases. This is what causes the entropy to lower as the distribution of a set becomes more uneven.

2.1.2. Information theory in linguistics

In recent years, the application of information theoretical approaches in psycholinguistic studies has seen an increase. One of the earliest examples of this is a study by Kostiç (1995). Kostiç calculated the information load of Serbian inflected nouns using a slightly altered equation than the one given in (1). His method, illustrated in (5), calculated the probability not only based on the frequency of a specific inflection, but rather based on the frequency of the infection relative to the morphological paradigm (i.e., every possible inflection of that noun). Furthermore, the frequency was divided by the number of functions that specific inflection carries. For example, phonologically, the inflection –s after a noun can have two functions. It can indicate plural (the dogs) or possession (the dog's). In this case the frequency would thus be divided by two.

(5) $$I_e = -\log_2\left(\frac{F_e/R_e}{\sum_e F_e/R_e}\right)$$

In (5), F denotes frequency and R denotes the number of functions. In the current study, R will denote the number of functions of each inflected member of a verb paradigm, as illustrated below in table 1.

Table 1. *Number of functions for each inflected member of the verb* walk

Inflected form	Functions	Number of functions
(to) walk	1st/2nd sg. pres., 1st/2nd/3rd pl. pres., infinitive	6
walks	3rd person singular present	1
walking	present continuous[1]	1
walked	1st/2nd/3rd sg. past., 1st/2nd/3rd pl. past., past participle.	7

Calculating the information load of an inflectional ending is achieved by dividing the frequency count by the number of functions, dividing the outcome by the sum of every item's frequency divided by number of functions and taking the $-\log_2$ of the outcome.

Kostiç (1995) used the inflected nouns in the paradigm in a visual lexical decision task. In this experimental design, participants are tasked with deciding, as quickly as possible, whether a word presented on the screen is an existing or a non-existing word. The reaction times of the participants are measured. It is assumed that when a certain stimulus requires more mental processing, the reaction time will be higher. In his experiment, Kostiç found a clear correlation between reaction times and information load; the higher the information load, the slower the reaction times. A higher information load thus requires more processing time.

In a similar experiment, Baayen, Feldman and Schreuder (2006) applied different measures to response data of an earlier visual lexical decision task (cf. Balota, Cortese, Sergent-Marshall, Spieler & Yap, 2004). One of these measures was inflection entropy, which is obtained by calculating the information load of each member in the paradigm (by the use of the formula in (1)) and subjecting the outcomes to Shannon's (1948) entropy

[1] Note that some functions, for example the gerund (I like walking) are not included. As section 3 will explain, the frequencies are obtained through a corpus by searching for the number of occurrences of *verb* forms, which does not include verb forms used as, for example, nouns.

formula in (3). The resulting entropy value is taken to be representative of the complexity of lexical access to the members in a given paradigm.

Baayen et al. found that higher inflectional entropy lead to faster reaction times in comprehension tasks. The explanation given by the authors was that higher entropy represents a more informationally rich inflectional paradigm; one that has more connections to other items in the mental lexicon, which facilitates reaction times.

The studies above are just a few of many that have used information load and inflectional entropy to study lexical retrieval concerning nouns (cf. De Jong, Feldman, Schreuder, Patizzo & Baayen, 2002; Kuperman, Bertram & Baayen, 2008; Milin Filipovic Durdevic & del Prado Martín, 2009). Far fewer studies have focussed on verbs in their lexical decision experiments. In a series of papers, Tabak et al. (2006; 2010) studied verbs in both a lexical comprehension and production task. Interestingly, they found that, though form frequency and inflectional entropy both facilitated reaction times, the measures were not correlated, meaning they function independently and that inflectional entropy does not merely predict what frequency already does. Another interesting finding was that, though high entropy had a facilitatory effect on reaction times in lexical decision tasks, it had an inhibitory effect in production tasks.

In her study, Van Ewijk (2013) ran similar experiments. An important difference in her study was that the calculation of the inflectional entropy included the measure proposed by Kostiç (see (5)). Van Ewijk reasoned that the distribution of the information load of the inflected members of a verb paradigm are representative of the memory trace for that item (see the next section for a more elaborate discussion on lexical access). A higher information load means the presence of a stronger memory trace, which requires less activation to reach the threshold of being activated. According to her, the memory trace is affected by the number of contexts in which a given inflected verb form is used, so the frequency of the form

alone is not sufficient to account for the information load of an item. The formula used by Van Ewijk to calculate inflectional entropy is presented below in (6) and will be adopted in this paper as well.

(6) $$H = -\sum {}^{F_e}\!/_{R_e} \log_2 {}^{F_e}\!/_{R_e}$$

The study found similar results to that of Tabak et al.: higher entropy had a facilitatory effect during comprehension tasks but an inhibitory effect during production tasks. The explanation Van Ewijk (2013) provides for this elegantly stipulates the most crucial difference in the type of lexical retrieval performed in each task. She defines lexical retrieval in production as a conceptually driven task: one starts out with a concept, finds the appropriate overarching verb (lemma), and has to select the correct inflected member for the context. In the case of a high entropy verb paradigm, the distributions of probabilities are all very similar, hence it requires more processing time to select the correct form from amongst its competitors. Lexical retrieval in a comprehension task, on the other hand, Van Ewijk defines as a perceptually driven task. The correct form is heard and consequently activated in the long-term memory, after which activation spreads to the entire paradigm of the verb. In a high entropy environment, the connections between the verb forms are short and strong, so

the activation of the entire paradigm happens more quickly. Figure 1 below illustrates what a high entropy paradigm and a low entropy paradigm may look like.

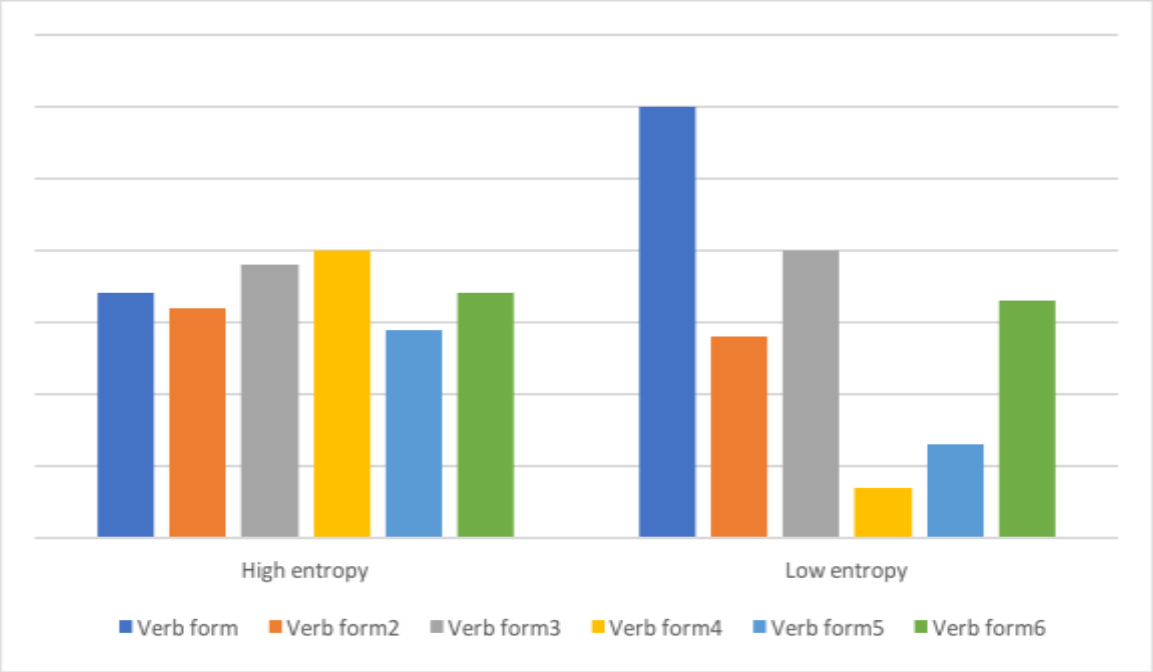

Fig. 1: *Example of a high and low entropy verb paradigm*

The high entropy paradigm clearly has short distances between each verb form, whereas the low entropy paradigm features much larger distances. As Van Ewijk (2013) puts it,

> … closeness in probability distribution means more equality among base levels of activation. In high entropy families activation of the family members boosts the activation level of the target, which means it is faster in reaching the threshold for retrieval, leading to shorter response latencies. (p. 53)

2.2. Lexical activation

This section will discuss the current theories on how lexical items are retrieved from the memory and how and why this might differ between certain groups.

2.2.1. Lexical retrieval

Though the organisation of the mental lexicon and the process of lexical retrieval are still largely a mystery, most models distinguish two steps in the process of lexical retrieval: the retrieval of the lemma (the 'base' of the word whose semantic information conveys the concept one is trying to express), followed by the selection of the lexeme (the appropriate form for the context in which it will appear (e.g. nominative, genitive) and for the speaker's intended meaning (e.g. person, tense, number, etc.). The speed of this process is dependent on the strength of the memory trace. Alexandrov, Boricheva, Pulvermüller and Shtyrov (2011) explain that "memory traces for words are frequently conceptualized neurobiologically as networks of neurons interconnected via reciprocal links developed through associative learning in the process of language acquisition" (p. 1). Thus, the more strongly connected the network for a certain lexical item, the faster it can be retrieved. The assumption in an information theoretical framework is that the strength of these connections is determined by the distribution of the information load of the lexemes within the paradigm of the lemma, which is derived from the frequency of these lexemes.

The remaining question, however, is how neurological network is organised in the presence of two languages. Though the existing body of work does not seem to have found a hint to this answer yet, previous studies have already uncovered a lot on how two languages in the bilingual brain seem to interact. The following section will discuss a number of these studies.

2.2.2. Bilingual lexical organisation

Within bilingual studies, the general consensus is that "semantic representations are shared across languages and […] these are connected to separate word-level representations in each

language." (Gollan, Montoya, Fennema-Notestine, & Morris, 2005). Some phenomena common in bilinguals are indicative of this theory. For example, bilinguals experience a dysfluency in lexical retrieval more frequently than monolinguals. This dysfluency is commonly known as a tip-of-the-tongue state, where subjects feel they are nearly successful in retrieving a certain word and can sometimes provide characteristics of the word (such as the initial sound) but, for an extended period, fail to recall the full form (Gollan et al., 2005). This finding has been documented for bilinguals of many different language pairs, such as Hebrew-English, Spanish-English and Tagalog-English (Gollan & Silverberg, 2001; Gollan & Acenas, 2004).

Numerous studies have also argued that bilinguals do not only have a shared semantic space, but that both languages are in fact activated simultaneously when processing language (dual lexical activation). A study by Hermans, Bongaerts, de Bot and Schreuder (1998) aimed to answer the question of whether bilingual subjects are able to supress activation in their first language when tasked with naming pictures in their second language[2]. To test this, they devised an experiment where Dutch-English bilingual participants were shown a picture of, for example, a mountain which they had to name in their second, less dominant, language (English). However, upon being shown the picture, they were also presented with an auditory stimulus which served as an interference stimulus. Several varying interference stimuli (ISs) were used in order to test different aspects, but the most striking one was the phono-Dutch ISs. These stimuli were English words that were phonologically similar to the target word in the non-target language, e.g., in the case of the target word *mountain*, the IS was *bench*, which is similar to the Dutch word for mountain: *berg*. Hermans et al. found that in the face

[2] Note that first and second language here are used to refer to both of a bilingual's native languages separately. This is different from the traditional use in which a first language refers to a person's native language and second refers to the person's non-native language.

of these ISs, naming times for bilinguals were slowed, which they concluded was strongly suggestive of both the target and non-target language being activated simultaneously.

Another picture naming experiment by Gollan, Montoya, Fennema-Notestine and Morris (2005) found that bilinguals who reported one of their languages to be notably more dominant than the other are still slower in picture naming tasks than monolinguals, which is in line with what dual lexical activation would predict. Interestingly, when tasked with classifying the pictures (e.g., deciding between whether the object shown is 'natural' or 'human-made') bilinguals had response times equally fast to monolinguals. The difference thus seems to lie solely in the lexical domain. Apart from these picture naming tasks, many studies have also investigated dual lexical activation using word naming tasks. These studies also uniformly found that bilinguals have difficulty supressing activation of the non-target language (Bijeljac-Babic, Biardeau, & Grainger, 1997; Jared & Kroll, 2001; van Heuven, Dijkstra, & Grainger, 1998).

Lastly, Marian and Spivey (2003) examined dual lexical activation in Russian-English bilinguals using an eye-tracking experiment. The participants were presented with an English word and four objects. The target object was described with the English stimulus, two other objects were fillers and the forth was an object the Russian name for which was phonologically related to the English target (the "between-language competitor"). Marian and Spivey found that the bilingual participants spent a good portion of the time observing the between-language competitor, leading them to conclude that there was parallel activation of both the bilinguals' languages.

2.3. Discussion

The previous section has shown that there is no scarcity of evidence for dual lexical activation. The findings of Gollan et al. (2005) and Hermans et al. (1998) suggest that the

phonological characteristics of a word are capable of activating a competitor in the other language. Many of the other studies seem to indicate that phonological similarity is not even a necessity for activating other-language competitors. However, how dual lexical activation generally affects a bilingual's performance and how the two languages are precisely organised in the mental lexicon of a bilingual largely remains a mystery.

The first section of this theoretical framework provided us with the means to measure the complexity of verb paradigms. Importantly, it should be noted that the entropy of a verb paradigm in language A is not necessarily equivalent to language B; as the frequency of the verb forms in a paradigm and their number of functions change, so will the inflectional entropy. This is key to the experimental design of this study.

By measuring bilinguals' reaction times of verbs that have different inflectional entropy in their two languages, it should be possible to gain insight into how the two languages are organised and interact with each other. The stimuli chosen for the experiment are verb forms that either belong to a high entropy paradigm in the target language and a high entropy paradigm in the non-target language, or a low entropy paradigm in the target language and a high entropy paradigm in the non-target language. The following section will justify this choice. By measuring the reaction times of bilinguals and comparing this to the reaction times of monolinguals, to whom only the first in the sets of paradigms applies, we will be able to compare whether and how the second paradigm has influenced the bilinguals' performance. At this point, we will formulate three hypotheses about the possible outcomes:

Increased workload hypothesis

Assuming that activating additional members in the verb paradigm slows reaction time by virtue of demanding more processing, thus resulting in higher reaction times, the bilinguals, overall, will be slower than the monolinguals. There will be a notable

difference between the different conditions, both within the bilingual group and as compared to the monolingual group.

The high entropy English and high entropy Dutch (H-H) verb pairs will see the fastest reaction times from bilinguals, though still slower than monolinguals, as the addition of a parallel verb paradigm will slow the processing.

The reaction times of bilinguals for the L-H condition are expected to be higher (i.e., slower reactions) compared to the monolinguals. The monolinguals will be slower in this condition than the H-H condition, but faster still than the bilinguals, as the latter group has and additional verb paradigm to process.

To summarise, the increased workload hypothesis simply states that more to process means more processing time and expects the bilinguals to be slower in both conditions.

Additional support hypothesis

The second hypothesis assumes that the additional activated language will not by default slow processing, but instead provide either a facilitatory or inhibitory effect, depending on whether the additional language has a high or low inflectional paradigm for the specific verb.

With this hypothesis in mind, we can thus assume that in the H-H condition, bilinguals will receive additional support from the indirectly activated language, resulting in lower reaction times than the monolinguals in this condition.

In the L-H condition, the task of identifying the verb will be facilitated by the activation of an additional high inflectional entropy verb paradigm for the bilinguals. The bilinguals are thus expected, by this hypothesis, to show lower reaction times than the monolinguals.

Unified paradigm hypothesis

The last hypothesis simply assumes the Dutch and English verb paradigms for any given verb to be collapsed as a single paradigm in the mental lexicon of a bilingual participant. The predictions the hypothesis makes stem from calculating the inflectional entropy as such. The data suggests that this generally results in higher inflectional entropy[3]. As Van Ewijk (2013) states, "intuitively, a greater number of members in a paradigm will lead to a greater entropy of the paradigm" (p. 33). If this hypothesis were to be borne out it would certainly provide researchers with a very useful predictive measure for the behaviour of bilinguals. This hypothesis also predicts that bilinguals would be universally slower in production, which seems to be the case Gollan et al. (2005) study.

Neurobiologically speaking, this hypothesis implies that the networks pertaining to the two languages are strongly and closely interconnected, to the point where lemmas contain the lexemes of both languages. This might seem problematic, as it would essentially mean the language 'overlap' in the brain. However, as a system it would seem most efficient to group lexemes together and encode a marker in each lexeme (in addition to the information about tense, person, etc.) to indicate to which language it belongs. The alternative would be to have a fork in the road, mapping each concept to two separate semantic representations. Furthermore, children at a very young age already seem to be experts in telling apart their native language(s). 7-month-old bilingual infants have been found to be able to tell apart their native languages based on prosodic cues (Gervain & Werker, 2013), and monolingual

[3]
Do: $H_D = 2.228$, $H_E = 2.199$, $H_U = 2.804$
Wash: $H_D = 1.606$, $H_E = 1.566$, $H_U = 2.334$
Think: $H_D = 1.475$, $H_E = 1.796$, $H_U = 2.063$
Walk: $H_D = 2.338$, $H_E = 1.620$, $H_U = 2.666$
D = Dutch, *E* = English *U* = unified

children of merely a few hours old were able to distinguish vowels that did or did not belong to their native language based on prenatal learning (Moon, Lagercrantz & Kuhl, 2013). In short, it appears that languages in general are very wealthy in cues that can be used to tell them apart. Assuming that the lexicons of two languages are unified in the brain the hypothesis might, as such, not be dramatically far-fetched or problematic.

Table 2 below provides an overview of the expected outcomes per condition. For the first hypothesis, the expected outcome for the bilinguals is also provided for the sake of completion despite that they are expected to be slower. For every hypothesis, the expected outcomes for the monolinguals remains the same as they are only faced with a single verb paradigm.

Table 2. *Expected outcomes per hypothesis. Per condition the group expected to display faster reaction times is along with a plus or minus to indicate differences in speed within the hypothesis*

Hypothesis	H–H		L–H	
Increase workload	Monolingual	+	Monolingual	-
	(*Bilingual*	*-)*	(*Bilingual*	*--)*
Additional support	Bilingual	+	Bilingual	-
Unified paradigm	Bilingual	+	Bilingual	+

2.4. A final note on entropy and frequency

Though the previously discussed study by Tabak et al. (2006;2010) already showed that, though both frequency and entropy affected reaction times, they have separate effects, I still deem it necessary to justify the choice of entropy over frequency for this study.

First of all, form frequency (i.e., the frequency for a singular inflected form that is part of a larger paradigm) would not be as indicative as entropy as it looks only at a small part of the whole. This would result in wildly different expectations of the results depending on which part of the paradigm we choose to study.

Lemma frequency (i.e., the overall frequency of all forms of a paradigm put together) would then be a better choice. We would expect that items with high lemma frequency indicate stronger connections in the brain and hence would result in faster reaction times. However, this still has a distinct short coming. For example, suppose the existence of the following two paradigms: Paradigm A (frequency item 1: 100; frequency item 2; 100) and Paradigm B (frequency item 1: 350; frequency item 2: 50). A frequency based approach would predict faster reaction times for Paradigm B by virtue of the frequency for all items together is twice as high as that of Paradigm A. An entropy based approach, on the other hand, would predict faster reaction times for Paradigm A, by virtue of the distribution being perfectly equal, whereas this is very unequal for B. This then leads to a fork in the road rather than a conclusion; both measures predict different things. Based on previous work entropy seems to be more representative for neural networks than frequency, but we will justify the choice further still.

As section 3 will explain, the frequency data needed to calculate entropy are obtained from corpora. As this study will look at the processing of two languages, two corpora needed to be consulted. This is where using lemma frequency becomes troublesome: as the size of corpora differs greatly, and based on the source material might over or underrepresent certain verbs, it would be difficult to obtain a relative value representative for verbs in Dutch and English. Especially for the unified paradigm hypothesis, which involves the merger of two paradigms, this would present a problem. If the corpus for Dutch is twenty times as large as the corpus for English, adding the frequency of the English verbs to that of the Dutch verbs

will make for only a minor difference. If then, our hypothesis would be that higher lemma frequency would correspond to faster reaction times, the frequency data would predict a minimal difference, whereas this difference might in fact be much larger. Entropy has the advantage of being a relative value that is calculated based on the encapsulated differences between verb forms *within* the paradigm of the verb itself. In other words, absolute differences in frequency between verb paradigms will not have an impact on the entropy values.

We take the above to be sufficient in justifying the choice of employing entropy for the current study.

3. Methodology

3.1. Design

The design of the experiment was already briefly mentioned at the end of section 2.3 and will be repeated here for the sake of clarity and completeness.

For the experiment, verbs were gathered that belong either, a) a high entropy verb paradigm in English and a high entropy verb paradigm in Dutch, or b) a low entropy verb paradigm in English and a high entropy verb paradigm in Dutch. In an auditory lexical decision task, English monolinguals and English-Dutch bilinguals were told to tell apart existing and non-existing English verbs as quickly as possible. In the presence of dual lexical activation, bilinguals were expected to perform differently than the monolinguals by virtue of the added verb paradigm. As this additional verb paradigm was not present for the monolinguals, this group was simply expected to respond faster for high entropy verbs and slower for low entropy verbs. Due to the low availability of English-Dutch participants, English native speakers who were proficient in Dutch were also included in the bilingual group. Before the experiment, all participants were asked to report the languages they speak along with a self-evaluated score of their proficiency.

3.2. Participants

In total, 10 participants were tested. As many English-Dutch bilinguals as possible were recruited, after which an equal number of English monolinguals were recruited to counter-balance the experiment. As very few English-Dutch bilinguals were available, English speakers with varying proficiency levels in Dutch were also included in the bilingual group. Note that, though this group then consists of bilinguals and native speakers of English with Dutch as a second language, the group will be referred to as the ‘bilingual group’ throughout the rest of the paper.

During the evaluation of their language proficiencies, 2 participants revealed to speak more languages in addition to English and/or Dutch. These participants were excluded from the bilingual group during the analysis of the results. Hence, 8 participants remained; 4 in each group. The information on age, language proficiency and occupation can be found below in table 3.

Table 3. Overview participants (age, bilingual, and other language proficiencies).

Participant number	Bilingual	Age	Dutch proficiency	Other language & proficiency	Other language & proficiency	Occupation
pp01	Yes	24	5	n/a	n/a	Student
pp09	Yes	26	3	n/a	n/a	Student
pp10	Yes	20	10	n/a	n/a	Student
pp04	Yes	30	2	n/a	n/a	Student
pp02	No	23	n/a	n/a	n/a	Student
pp03	No	30	n/a	n/a	n/a	Student
pp07	No	25	n/a	n/a	n/a	Student
pp08	No	23	n/a	n/a	n/a	Student
pp05	Other	24	4	Spanish; 7	Italian; 4	Student
pp06	Other	25	4	German; 8	French; 8	Student

The mean age of the bilingual group is 25. The mean age of the monolingual group is 25.3. The excluded multilingual participants had a mean age of 24.5.

3.3. Material

As previously mentioned, two types of verbs were identified and used as stimuli: those belong to a high entropy English verb paradigm and a high entropy Dutch verb paradigm (H-H) and those belong to a low entropy English verb paradigm and a high entropy Dutch paradigm (L-H). Adding the other two possible combinations (i.e., H-L and L-L) would have

broadened the range of conclusions that we could draw from the data. However, it appeared that generally Dutch verb paradigms tend to be higher in entropy than English verb paradigms. This made it impossible to gather sufficient stimuli that fit the other configurations.

As the experiment is an auditory lexical decision task, rather than a visual one, the stimuli were gathered from spoken corpora. The Dutch stimuli were adopted from Van Ewijk (2013). The English stimuli were gathered from the spoken part of the British National Corpus (BNC) (2007).

From the data from Van Ewijk, the Dutch stimuli that featured the fewest synonyms and translation equivalents (i.e. unambiguous verbs that have direct translation equivalents) were selected and matched with the English verbs. The mean entropy of the Dutch verbs were taken as the deciding middle point of what constitutes high or low entropy. After the entropy was calculated for the matched English verbs, the 40 verbs with the highest lemma frequency were selected for each condition. As the spoken part of the BNC is relatively small, as far as corpora go, the lemma frequencies for some stimuli were very low. The ones with the highest lemma frequencies were thus considered to be more reliably representative.

The English verbs served as the target stimuli for the experiment. From the verb paradigm, the simple past form was chosen. The motivation behind this is that the simple past form is most easily recognisable as a verb, as (7) below illustrates.

(7)	a.	is/I like	walking/thinking/sleeping/talking
	b.	I need (a)	walk/think/sleep/talk
	c.	the	walks/thinks*/sleeps*/talks
	d.	the	walked*/thought/slept*/talked*

The continuous form (7a) would have been a very poor choice as this can always be used as a noun in the gerund form. For the same reason, the singular present (7b) would also be a poor choice; though (7) is far from exhaustive, it shows that many of these forms occur also as nouns. Lastly, (7c) and (7d) illustrate that the possibility of 3rd person singular verbs and past tense verbs occurring as nouns are about equal. However, in the current list of material (see Appendix A), the verbs of form (7c) that can be used as nouns are roughly 50% more frequent than those of form (7d).

In total there were 80 stimuli, 40 in each condition (a list of the English stimuli, their Dutch translation equivalent and their English and Dutch entropy values can be found in appendix A). To serve as controls, 80 non-verbs were invented. The non-verbs were all phonologically legal pseudo-words that featured common English regular and irregular simple past inflections. All verbs were recorded by an adult, male native speaker of English. The auditory stimuli were edited so that they did not feature a pause before or after the word. In total, all participants were presented 160 stimuli.

3.3.1. Gathering the English stimuli: the EC4000

One of the largest challenges in applying an information theoretical approach in language research is that inflectional entropy cannot be guessed; despite it having a measurable, subconscious effect on lexical retrieval, one cannot sense what the approximate inflection entropy of a verb paradigm will be. This makes gathering stimuli a rather lengthy process. Seeing especially how in the current study the entropy values had to conform to certain requirements to fit the predetermined conditions (H-H and L-H), over 200 verbs needed to be checked before finding 40 appropriate ones for each condition. Performing this process manually would imply obtaining frequency data from the BNC, performing the calculations

on an online entropy calculator and documenting the results. This process could easily take up to 10 minutes per verb, making the entire process of gathering stimuli last over 33 hours.

Instead, I chose to develop a programme that searched through the BNC automatically, extracted the frequency data and calculated the entropy. The Entropy Calculator 4000 (EC4000) was based on an earlier version (the EC3000) that required manual gathering of the frequency data. The newer version extracts the Dutch verb from a pre-existing .xlsx document and prompts the user to enter the lemma of the translation equivalent. It then searches through a downloaded version of the BNC (which was, by the use of another script, trimmed down to contain only verbs in order to increase the processing speed) and extracts the frequency data for each verb form. As the corpus does not contain information on the number of functions of a verb form, the user is alerted by an audio fragment[4] upon the completion of collecting the frequency data and presented with the verb forms and frequencies the programme gathered. Here, the user has the option to exclude certain verb forms from the calculations (e.g. ungrammatical single instances such as ‘thinked’) before entering the number of functions for each verb. Next, the entropy is calculated using the module from the old EC3000 which employed the calculations in (1) and (6). The entropy value is saved in an new .xlsx document which includes the Dutch lemma and the entropy of the respective verb paradigm, the English lemma and the entropy of the respective verb paradigm and the lemma frequency. The process is then repeated for the next verb on the list.

To ensure that the calculations were correct, frequency data for each verb form of a Dutch verb, obtained from Van Ewijk (2013), were fed into the EC4000. It reproduced the entropy value reported in the original paper exactly. Furthermore, the lemma frequencies the EC4000 reported for a few randomly selected verbs were compared to those obtained from

[4] From the modern classic cinematic master piece The Room (2003).

the web-based version of the BNC. No difference was found. The full code of the EC4000 can be found in Appendix B.

3.4. Procedure

The participants were tested in a soundproof booth in the UiL OTS lab. The booth contained two speakers, set to a comfortable but clear volume, a chair and a monitor and button box placed on a desk at a comfortable distance from the seated participant.

After filling out the language proficiency form, the informed consent form and reading the information letter, they were provided with instructions from the experimenter. All participants received instructions in English to prevent priming effects. The participants were told that their task is to decide, as quickly as possible, whether the verb they heard does or does not exist in the English language, and that they should do so using the button box in front of them. The buttons on the box were labelled *NO* and *YES*. The participants were explicitly informed that the verbs all featured simple past tense.

Before the actual experiment, the participants had to complete a trial phase featuring 2 words that were not included in the test phase. During this trial phase, the experimenter was present to oversee if all went well and to be available for any further questions after the completion of the trial phase. During the test phase, the experimenter left the booth. After completion of the experiment, participants were compensated 5 euros. The total duration of the experiment (including the signing of forms) was around 10 minutes.

The experiment itself was made using ZEP (Veenker, T.J.G., 2016). It was based on the pre-existing template for auditory lexical decision experiments. The script first shows a fixation dot on the monitor to assure participants' attention before playing an auditory stimulus. The stimuli were automatically randomised per participant to prevent any effect the order of the stimuli might have. After the stimulus was played, participants had 2000 ms to

judge the stimulus using the button box. After 2000 ms, the fixation dot was displayed again and the experiment moved on to the next stimulus.

The experimental script automatically recorded the reaction time of the participants for each item, along with their choice and the target answer. If a participant answered correctly within 2000 ms, their reaction time was recorded as well as a ‘1’ in the column ‘correct’ to indicate the quality of their answer. If participants answered incorrectly the reaction time was still reported as well as a ‘0’ was in the ‘correct’ column. In the case a participant did not respond on time, the reaction time was recorded as -9999 and their answer showed in the column ‘correct’ as -1.

3.5. Coding of results

The results were exported to SPSS (SPSS Inc., Chicago IL) where the participant information (bilingual/monolingual/other and Dutch proficiency) were added, as well as the stimulus information (Dutch inflection entropy, English inflection entropy and condition). Furthermore the reaction times for the control items (i.e. the non-verbs) were discarded. The new dataset was saved and further analysed in R studio (RStudio Team, 2015).

In R, the participants who were proficient in more languages beyond Dutch and English were excluded from the bilingual group and the overall comparative analyses between the two groups.

Furthermore, the data for incorrect responses and lack of responses were excluded. For the remaining data, the log of the reaction times were taken in order to ensure a normal distribution. With Bates, Maechler and Bolker’s lme4 package (2013), the data was analysed using linear mixed effects models. Participants and items were added as random factors.

4. Results

Table 4 reports the mean score for correctness and reaction times (RT) of the monolinguals, the bilinguals and the bilinguals who spoke more languages in addition to English and Dutch.

Table 4. *Mean error rates and mean RT in ms of correct items per group per condition*

Group & Condition		% correct	Mean RT of correct items
Monolingual;	H-H	95%	969
Monolingual;	L-H	97.5%	996
Bilingual;	H-H	95.6%	826
Bilingual;	L-H	99.3%	831
Multilingual;	H-H	96.3%	829
Multilingual;	L-H	97.5%	845

The rate of correctness did not differ significantly between conditions for the monolinguals (β= 0.03, SE = 0.03, t = .08, p = .423) or the bilinguals ((β= 0.04, SE = 0.03, t = 1.37, p = .175). Furthermore, no significant difference was found in error rate between the bilinguals and monolinguals (β= 0.01, SE = 0.01, t = .79, p = .456).

Figure 2 illustrates a boxplot of the reaction times per group per condition for the correctly judged items.

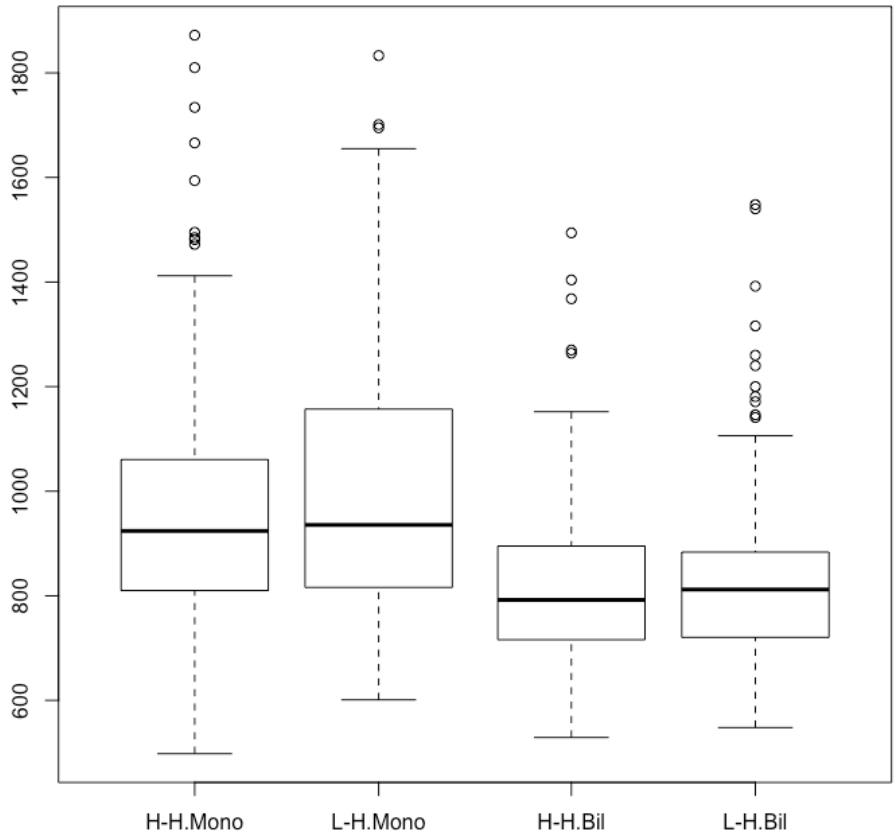

Fig. 2: *Box plot of RTs per group per condition*

The difference between conditions was not significant for either group (monolingual: β = .02, *SE* = 0.03, *t* = .76, *p* = .449; bilingual: β > .01, *SE* = 0.03, *t* = .16, *p* = .878). The difference in RT between groups was significant for the bilinguals compared to monolinguals: (β = -.16, *SE* = 0.06, *t* = -2.86, *p* = .024).

So far, two things have become apparent: the bilingual groups is generally faster than the monolingual group, and condition does not seem to affect either group. The box plot in figure 2 does, however, seem to show a difference between conditions within the monolingual group. We will assume that what we have taken to constitute high and low entropy does not capture the data, and, instead, include the entropy of the English verbs as a main effect. Furthermore, the increased workload hypothesis and additional support

hypothesis would predict that the data of the bilingual group would be captured by using the interaction of the English entropy vales and the Dutch entropy values.

The entropy value of the English verb paradigms had a significant effect for the monolingual group (β = -.11, *SE* = 0.05, *t* = -2.28, *p* = .003), but not for the bilingual group (β = -.06, *SE* = 0.04, *t* = -1.43, *p* = .158). The inclusion of the interaction between the English and Dutch entropy values was not significant either for the bilingual group (β = .29, *SE* = 0.21, *t* = 1.42, *p* = .159).

Lastly, the interaction of the self-indicated Dutch proficiency scores were added to the previous interaction. This, too, had no effect on the bilingual group (β = .04, *SE* = 0.05, *t* = .81, *p* = .419).

5. Discussion & Conclusion

It is important to start this section with a great note of caution: the amount of participants used in the current experiment were, by far, insufficient to draw firm conclusion from the findings. In order for the current study to have enough statistical power, we would have required ten times the current number of participants. However, the following paragraphs will draw conclusions from the current data and discuss the implications of the findings that would apply should the results presented in the previous section be representative of a much larger sample of participants. In short, the conclusions and implications in the paragraphs below should not be taken as empirical truths, but at most as hints at what a larger scale study might find and what the implications of these findings would be.

First of all, it appeared that the dividing line between high and low inflection entropy defined in this study were not entirely accurate as such as no difference between the condition was found for the monolingual group. However, replacing this measure by the gradual effect of the English entropy values instead of the binary high/low factor proved to be effective in capturing the data of the English monolingual group. The estimate showed a negative correlation between entropy and reaction time, meaning that as entropy increases, reaction time decreases. From this, it can be concluded that the calculated entropy values were reliable to use as an independent variable. The English entropy values did, however, have no effect on the bilingual group, nor did the Dutch entropy values or the interaction between the two entropy values. At this point, it is safe to discard both the increased workload hypothesis and the additional support hypothesis.

The increased workload hypothesis predicted that the addition of an extra verb paradigm would, by definition, require more processing power and, consequently, slow reaction times regardless of condition. However, based on the fact alone that the bilinguals responded significantly faster than the monolinguals, regardless of condition, this hypothesis

does not seem to hold. Furthermore, the hypothesis predicted a difference between conditions within the bilingual group. This difference was not present in the results.

The additional support hypothesis predicted that the activation of an additional verb paradigm would either slow the reaction time or speed it up, depending on whether the additional paradigm was of low or high entropy respectively. With the current material, we would thus expect to find a) the bilinguals to show faster reaction times than the monolinguals and, b) a within-group difference between conditions for the bilinguals. The first prediction was definitely present: the bilinguals were found to be significantly faster than the monolinguals. However, there was no difference found between conditions within the bilingual group. One could argue that, as in both conditions the reaction time was facilitated by the presence of an additional, high entropy verb paradigm, the difference between the two conditions simply became ‘blurred’; i.e. the differences were sufficiently minuscule to become immeasurable. However, this does not correlate with what this hypothesis predicts.

The hypothesis that seems to best account for the data is the unified paradigm hypothesis. This hypothesis predicted that verb paradigms in the bilingual mental lexicon encompass all inflected verb forms of both languages. To predict the performance then of bilingual participants, the entropy value should be calculated on the basis of the information loads of both the English and the Dutch verb forms. As presented in section 2.3, entropy values universally seem to increase as more is added to the paradigm. The hypothesis thus predicts that the bilinguals will be notably faster than the monolinguals. Furthermore, the entropy values of a system A and a system B cannot be taken to be indicative of the entropy value of a system C in which the individual probabilities of systems A and systems B are combined. This was reflected in the results by the fact that neither the Dutch entropy values, the English entropy values or the interaction between the values were able to capture the performance of the bilinguals.

In terms of the mental lexicon, the hypothesis implies that lexemes that carry the same semantic information in both languages are stored as single collection in the same place. Baayen et al. (2006) argued that paradigm that is more complex and inflectionally rich contains more connections to other items in the mental lexicon. According to them, this would result in faster reaction times. This argument pertained to the language internal differences between high and low entropy paradigms, however, it would stand to reason that if the lexemes of two languages are mapped to the same paradigm, it would vastly increase the number of connections to the rest of the mental lexicon. This explanation also accounts for the seemingly universal tendency of English verb paradigms to increase in entropy when the lexemes of the Dutch verb paradigm are added to it.

So far, an important fact about the bilingual participants in this study has been footnoted: namely that the set of participants only included one true English-Dutch bilinguals. The other participants in this group were English native speakers who learned Dutch as a second language with varying degrees of proficiency. If the same results were found on a larger scale experiment, it would imply that even when learning a second language, the L2 words that share a translation equivalent in one's native language are mapped onto the same space at a rather early stage in the acquisition of the language.

Lastly, the account provided here and the relevant entropy measures also seem to account for previous findings. In section 2.2.2, a study by Gollan, Montoya, Fennema-Notestine and Morris (2005) was mentioned that found that bilinguals who had one language that was clearly more dominant than the other were still slower in picture naming tasks than monolinguals. According to the unified paradigm hypothesis, these bilinguals would generally have word paradigms of greater inflection entropy. As was explained in section 2.1.2, this would lead to faster reaction times in comprehension tasks, such as auditory lexical decision tasks, but *slower* reaction times in production tasks such as the one used by Gollan

et al. The unified paradigm hypothesis thus seems to be capable of accounting for the data found in the Gollan et al. study.

In conclusion, the present study set out to discover whether a difference could be found between the mental lexical organisation of bilinguals and monolinguals. The reasoning behind this was that in bilinguals, both languages might be activated at once even when they are performing tasks in only one of their languages. The study lacked sufficient power to make firm, empirical conclusion about this, but the results hint that this is indeed the case. Furthermore, the study applied information theoretical measures to test the hypotheses. A secondary goal was to find out whether these measures could serve as a useful, reliable tool in bilingual studies. Once again, no firm conclusions can be drawn from the current data, but at the very least this study hints at that information theoretical measures can prove to be a powerful approach in bilingual studies and even provided a measure that might be more applicable to bilinguals than calculating the entropy of their verb paradigm separately for each language. Future studies could investigate further if it is indeed more appropriate to treat bilingual verb paradigms as unified by testing whether the combined entropy makes accurate predictions about their reaction time in lexical decision tasks.

References

Alexandrov, A.A., Boricheva, D.O., Pulvermüller, F., & Shtyrov, Y. (2011). Strength of word-specific neural memory traces assessed electrophysiologically. PLoS ONE 6 (8), e22999. doi:10.1371/journal.pone.0022999.

Baayen, R. H., Feldman, L., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language, 53*, 496–512.

Balota, D. A. & Duchek, J. M. (1988). Age-related differences in lexical access spreading activation, and simple pronunciation. *Psychology and Aging, 3*, 84-93.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.0-5.

Bijeljac-Babic, R., Biardeau, A., & Grainger, J. (1997). Masked orthographic priming in bilingual word recognition. *Memory & Cognition*, *25*(4), 447-457.

De Jong, N. H., Feldman, L. B., Schreuder, R., Patizzo, M. & Baayen, R.H. (2002). The morphological family size effect and morphology. *Language and Cognitive Processes, 15*, 329-365.

Gervain, J., & Werker, J. F. (2013). Prosody cues word order in 7-month-old bilingual infants. *Nature communications*, *4*, 1490.

Gollan, T. H., & Acenas, L. A. R. (2004). What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and tagalog-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(1), 246.

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition*, *33*(7), 1220-1234.

Gollan, T. H., & Silverberg, N. B. (2001). Tip-of-the-tongue states in Hebrew–English bilinguals. *Bilingualism: language and cognition*, *4*(01), 63-83.

Hermans, D., Bongaerts T., de Bot K., Schreuder R. (1998). *Producing words in a foreign language: Can speakers prevent interference from their first language, Bilingualism: Language and Cognition*, *1*, 213-230.

IBM Corp. (2016). IBM SPSS Statistics for Macintosh, Version 24.0. Armonk, NY: IBM Corp.

Jared, D., & Kroll, J. F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words?. *Journal of Memory and Language*, *44*(1), 2-31.

Kostiç, A. (1995). Informational load constraints on processing inflected morphology. In L. B. Feldman (Ed.) Morphological Aspects of Language Processing, Hillsdale, NJ: Erlbaum.

Kuperman, V., Bertram, R., and Baayen, R. H. (2008). Morphological dynamics in compound processing. *Language and Cognitive Processes, 23,* 1089-1132.

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism: Language and Cognition*, *6*(02), 97-115.

Milin, P., Đurđević, D. F., & del Prado Martín, F. M. (2009). The simultaneous effects of inflectional paradigms and classes on lexical recognition: Evidence from Serbian. *Journal of Memory and Language*, *60*(1), 50-64.

Moon, C., Lagercrantz, H., & Kuhl, P. K. (2013). Language experienced in utero affects vowel perception after birth: a two-country study. *Acta Paediatrica*, *102*(2), 156-160.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*, 379–423.

Tabak, W., Schreuder, R., & Baayen, R. H. (2006). *Nonderivational inflection*. Manuscript, Max Planck Institute for Psycholinguistics.

Tabak, W., Schreuder, R. & Baayen, R. H. (2010). Producing inflected verbs: A picture naming study. *The Mental Lexicon 5(1)*, 22-46.

Team, R. (2015). RStudio: integrated development for R. *RStudio, Inc., Boston, MA URL http://www.rstudio.com*.

The British National Corpus, version 3 (BNC XML Edition). 2007. Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: http://www.natcorp.ox.ac.uk/

Van Ewijk, L. (2013). *Word retrieval in acquired and developmental language disorders: A bit more on processing*. Netherlands Graduate School of Linguistics.

Van Ewijk, L. & Avrutin, S. (2016). Lexical access in non-fluent aphasia. *Aphasiology*, 1-19.

Van Heuven, W. J., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of memory and language*, *39*(3), 458-483.

Wiseau, T. (Producer), & Wiseau, T. (Director). (27 June 2003). *The Room* [Motion picture]. USA: Wiseau-Films.

Appendix A: List of stimuli with entropy values

	Verbs	Entropy English	Entropy Dutch		Non-verbs
L1	crept	1,531	2,41	1	Treft
L2	sounded	0,644	1,94	2	Ploarsed
L3	kissed	1,593	2,26	3	Dravelled
L4	scratched	1,31	2,37	4	Oarsed
L5	pinched	1,657	2,15	5	Wrelled
L6	demanded	1,343	2,25	6	Walshed
L7	stole	1,547	1,547	7	Stelt
L8	slept	1,578	2,18	8	Flut
L9	threatened	1,474	2,15	9	Betraved
L10	understood	1,296	2,03	10	Restood
L11	shone	1,461	2,01	11	Blung
L12	cared	1,597	2,16	12	Trowled
L13	hung	1,443	2,26	13	Trung
L14	pushed	1,564	2,24	14	Dwarked
L15	wished	1,572	2,35	15	Brinted
L16	floated	1,086	2,27	16	Earsted
L17	climbed	1,355	2,45	17	Reaved
L18	drove	1,619	2,18	18	Snove
L19	hit	1,534	2,27	19	Jit
L20	meant	1,143	2,39	20	Trunt
L21	cut	1,386	1,89	21	Krit
L22	lied	1,582	2,28	22	Plented
L23	walked	1,438	2,43	23	Sindered
L24	heard	1,635	2,38	24	Dankt
L25	sat	1,42	2,2	25	Strit
L26	let	1,499	2,04	26	Swit
L27	sought	1,433	2,44	27	Jought
L28	ran	1,425	2,19	28	Yint
L29	sang	1,492	2,35	29	Asht
L30	appeared	1,524	1,91	30	Dempled

L31	read	1,248	1,96	31	Vung
L32	thought	1,454	2,37	32	Genought
L33	talked	0,742	1,92	33	Drempled
L34	stood	1,562	2,12	34	Goot
L35	played	1,442	2,25	35	Spranded
L36	looked	1,653	2,21	36	Floofed
L37	moved	1,534	2,21	37	Exered
L38	saw	1,625	2,35	38	Trum
L39	put	1,385	1,87	39	Brust
L40	split	1,534	1,85	40	Clarpt
H1	added	1,755	2,32	41	Dreared
H2	died	1,925	1,98	42	Beroaved
H3	served	1,857	2,12	43	Dilaped
H4	sent	1,854	2,08	44	Benst
H5	pulled	1,756	2,39	45	Frankered
H6	won	1,966	1,74	46	Ston
H7	bought	1,761	2,14	47	Endought
H8	wore	1,882	2,34	48	Frore
H9	wrote	1,869	2,04	49	Geft
H10	ate	1,942	1,76	50	Jint
H11	pointed	1,821	2,48	51	Foared
H12	spoke	1,895	2,21	52	Frood
H13	woke	2,277	1,93	53	Doke
H14	found	1,674	2,43	54	Stound
H15	chose	2,006	1,74	55	Fose
H16	asked	1,673	2,22	56	Bittled
H17	threw	1,994	2,24	57	Prew
H18	rang	2,213	2,24	58	Shrang
H19	stopped	1,841	2,37	59	Nellowed
H20	learned	1,674	2,13	60	Curted
H21	helped	1,767	2,18	61	Frittled
H22	followed	1,691	2,21	62	Janked
H23	broke	2,096	1,89	63	Plent

H24	turned	1,934	2,2	64	Tummed
H25	created	1,853	1,91	65	Gangled
H26	showed	2,121	2,2	66	Pinkled
H27	became	1,817	2,2	67	Bood
H28	lead	1,793	2,07	68	Bestrood
H29	stayed	1,739	2,4	69	Klunged
H30	felt	1,848	2,28	70	Sprought
H31	fell	2,153	2,1	71	Zunt
H32	brought	1,931	2,06	72	Jimpt
H33	held	1,783	2,26	73	Frenst
H34	gave	2,078	2,24	74	Hought
H35	took	2,17	2,51	75	Lought
H36	made	1,912	2,16	76	Sut
H37	went	1,768	1,98	77	Splought
H38	came	1,769	2,45	78	Mought
H39	said	1,934	2,09	79	Klought
H40	did	2,227	2,31	80	Fid

Appendix B: Code of EC4000

```
import xml.dom.minidom
from xml.dom.minidom import parse, parseString
import os
import collections
import time
import math
import re
import os
import sys
import pygame
from openpyxl import load_workbook
from openpyxl import Workbook
pygame.mixer.init()


def load_wbw():
   global wbw
   global wsw
   try:
      wbw = load_workbook(filename = 'entropy.xlsx')
      wsw = wbw['Sheet']
      print('ADDING TO EXISTING WORKBOOK')
   except FileNotFoundError:
      wbw = Workbook()
      wsw = wbw['Sheet']
      print('CREATING NEW WORKBOOK...')
load_wbw()
wb = load_workbook(filename = 'verbs.xlsx')
sheetnames = wb.sheetnames
ws = wb[sheetnames[0]]
wsw['D1'] = 'verb-english'
wsw['E1'] = 'H'
wsw['F1'] = 'Lemfreq'
wsw['G1'] = 'Comparison'

def ohhi():
   pygame.mixer.music.load("OhHiMark.wav")
   pygame.mixer.music.play()
   pygame.mixer.music.stop
def lisa():
   pygame.mixer.music.load("Lisa.wav")
   pygame.mixer.music.play()
   pygame.mixer.music.stop
def what():
   pygame.mixer.music.load("What.wav")
   pygame.mixer.music.play()
   pygame.mixer.music.stop
def yeah():
   pygame.mixer.music.load("Yeah.wav")
   pygame.mixer.music.play()
   pygame.mixer.music.stop
def okay():
   pygame.mixer.music.load("Okay.wav")
```

```
    pygame.mixer.music.play()
    pygame.mixer.music.stop

def collect_verb(word_english):
    verb = word_english
    thing = []
    total_list = []
    FR = []
    sum_log_FR = []
    P = []
    Plog = []
    Psum = []
    count = 0
    t0 = time.time()


    file_dir = './verb_files'
    for subdirs, dirs, filenames in os.walk(file_dir):

        for file in filenames:
            print(file,'\n')
            count+=1
            print("percentage done \n %s" % ((count/len(filenames))*100),'% \n\n')
            if file.endswith('.xml'):
                datasource = open(os.path.join(subdirs, file), 'r', encoding='UTF-8')
                dom = parse(datasource)
                w=dom.getElementsByTagName('w')
                for node in w:
                    if node.getAttribute('pos') == 'VERB' and node.getAttribute('hw') == verb:
                        thing.append(str.strip(str.lower(node.lastChild.nodeValue)))

    t1= time.time()
    print('finished in: %s seconds' % (t1-t0),0)
    dict_temp = collections.Counter(thing)
    print(dict_temp)
    ohhi()

    def delete():
        yn = input(str("Would you like to delete anything? "))
        if yn == 'y':
                        what1 = input(str("What would you like to delete? " ).lower())
            if what1 in dict_temp:
                del dict_temp[what1]
                                print(dict_temp)
                delete()
            else:
                print("Doesn't exist, boi")
                delete()
        elif yn == 'n':
                        print("Moving on...")
        else:
                        delete()
    delete()

    for x, value in dict_temp.items():
        key = str((x, value))
        def options():
```

```
            usr_input = input("Add number of functions for %s : " % (key))
            try:
                a = int(usr_input)
                dict_temp[x] = [dict_temp[x], a]
            except ValueError:
                print("YOU ARE TEARING ME APART LISA!!!")
                lisa()
                options()
        options()

    dict_verb = dict(dict_temp)
    print(dict_verb)

    #Entropy calculating sub-module

    for item, values in dict_verb.items():
        total_list = []
        total_list.append(values[0])
    global total
    total = sum(total_list)
    print(total)

    for item, values in dict_verb.items():
        print('The frequency counts are: ',values[0])
        print('The number of functions are: ',values[1])
        FR.append((values[0]/total)/values[1])
        sum_log_FR.append(((values[0]/total)/values[1])*math.log2((values[0]/total)/values[1]))

    sum_FR = math.fsum(FR)
    print(FR)

    n = 0
    while n < len(FR):
        P.append(FR[n]/sum_FR)
        Plog.append(math.log2(FR[n]/sum_FR))
        Psum.append(P[n]*Plog[n])
        n+=1

    global H
    H = (-1*round(sum(Psum),3))
    print("H =",-1*round(sum(Psum),3))


def copy_data():

    #function lists
    col = ['A','B','C']

    #function
    for letter in col:
        verb_d=ws[letter]
        for cel in verb_d:
            x = 0
            y = 1
            while x < len(verb_d):
                wsw['%s%s' % (letter,y)] = verb_d[x].value
                x+=1
```

```
            y+=1


def go_through_list():
    copy_data()
    #function variables
    try:
        with open("saved_state.txt") as f:
            saved = f.read().splitlines()
            f.close()
        if os.stat("saved_state.txt").st_size == 0:
            x = 1
            print(str.upper('Starting fresh'))
        else:
            inp=input("CONTINUE FROM SAVED STATE (Y/N)? ")
            if inp.lower() == 'y':
                x = int(saved[0])
                print(str.upper('Resuming from saved state'))
            else:
                x = 1
                print("Starting fresh")
    except FileNotFoundError:
        x=1

    y = x+1
    verb_form = wsw['A']
    verb_form_En = ws['D']
    Hd = wsw['B']
    He = wsw['E']
    empty_string = ""

    for cel in verb_form:
        ##controlled runtime
        #while x < 4:
        ##whole thing
        while x < len(verb_form):
            word_dutch = verb_form[x]
            print(verb_form[x].value)
            print(verb_form_En[x].value)
            #skip regular verbs
            if verb_form_En[x].value == None or verb_form_En[x].value == empty_string:
                x+=1
                y+=1
                print("regular")
            elif He[x].value != None or He[x].value != empty_string:
                x+=1
                y+=1
                print("entropy known")
            else:
                if 1 == 1:
               # if verb_form_En[x].value != None or verb_form_En[x].value != empty_string:
                    word_english = verb_form_En[x].value
                    print(word_english)
                    wsw['D%s' % (y)] = word_english

                    print(wsw['D%s' % (y)])
                    print(wsw['D%s' % (y)].value)
```

```
            wsw['D%s' % (y)] = word_english
            collect_verb(word_english)
            wsw['E%s' % (y)] = H
            wsw['F%s' % (y)] = total
            try:
              a = float(Hd[x].value)
              if a > 1.987 and H > 1.987:
                wsw['G%s' % (y)] = "H - H"
              elif a > 1.987 and H < 1.987:
                wsw['G%s' % (y)] = "H - L"
              elif a < 1.987 and H < 1.987:
                wsw['G%s' % (y)] = "L - L"
              elif a < 1.987 and H > 1.987:
                wsw['G%s' % (y)] = "L - H"
            except ValueError:
              print("YOU ARE TEARING ME APART LISA!!! \n\n (something went wrong in comparing H-
values)")
              lisa()
            wbw.save('entropy.xlsx')
            x+=1
            y+=1
          else:
            #word_english = str(input('INFO: type \'QUIT\' to quit and \'SKIP\' to skip word \n\n Enter
translation for verb %s: ' % (word_dutch.value)))
            word_english = "SKIP"
            print(word_english)
            if word_english == "SKIP":
              x+=1
              y+=1
            elif word_english == "QUIT":
              wbw.save('entropy.xlsx')
              saved=open('saved_state.txt','w')
              saved.write(str(x))
              saved.close()
              quit()
            else:

              print(wsw['D%s' % (y)])
              print(wsw['D%s' % (y)].value)
              wsw['D%s' % (y)] = word_english
              collect_verb(word_english)
              wsw['E%s' % (y)] = H
              wsw['F%s' % (y)] = total
              try:
                a = float(Hd[x].value)
                if a > 1.987 and H > 1.987:
                  wsw['G%s' % (y)] = "H - H"
                elif a > 1.987 and H < 1.987:
                  wsw['G%s' % (y)] = "H - L"
                elif a < 1.987 and H < 1.987:
                  wsw['G%s' % (y)] = "L - L"
                elif a < 1.987 and H > 1.987:
                  wsw['G%s' % (y)] = "L - H"
              except ValueError:
                print("YOU ARE TEARING ME APART LISA!!! \n\n (something went wrong in comparing H-
values)")
                lisa()
```

```
            wbw.save('entropy_2.xlsx')
            x+=1
            y+=1

    wbw.save('entropy.xlsx')
go_through_list()
```