UNIVERSITEIT GENT

FACULTEIT PSYCHOLOGIE EN
PEDAGOGISCHE WETENSCHAPPEN

Reading in a Second Language: A Corpus Study

Uschi Cop

Promotor: Prof. Dr. Wouter Duyck
Co-Promotor: Prof. Dr. Denis Drieghe

Proefschrift ingediend tot het behalen van de academische graad
van Doctor in de Psychologie

2015

CONTENTS

ACKNOWLEDGMENTS

Vier jaar. Een lange tijd van onderzoeken, kritisch evalueren, en herkalibreren met als resultaat een proefschrift waar ik erg trots op ben. Maar deze eer komt niet alleen mij toe. Dit proefschrift was niet mogelijk geweest zonder de inbreng van meerdere personen, die ik graag hiervoor erken.

Ik bedank het Fonds voor Wetenschappelijk Onderzoek, zij hebben dit doctoraat financieel mogelijk gemaakt. Ook wil ik Wouter Duyck bedanken om mijn promotor te zijn en om altijd zelfstandigheid en het communiceren van resultaten aan te moedigen. Denis Drieghe, je was mijn copromotor en hebt me bijgestaan met wetenschappelijke raad. Ik wil je extra bedanken voor de gastvrijheid en de gezellige avonden in Southampton. I would also like to thank Simon Liversedge for sharing his eye-tracking expertise and for his critical scientific input during the doctoral committee meetings.

Mijn co-auteurs Eva, Emmanuel en Nicolas hebben de hoofdstukken van mijn doctoraat met hun respectievelijke expertise naar een hoger niveau getild. Eva, voor je interesse in mijn onderzoek wil ik je bedanken. Vaak heb ik jouw boeken en papers geraadpleegd. Emmanuel bedankt voor de server services, de interessante frequente gesprekken en de roadtrips. Nicolas, jouw inzet en inzicht zijn bewonderenswaardig. Het was een plezier om met jullie samen te werken.

Ook wil ik de volledige vakgroep Experimentele Psychologie bedanken, inclusief het ondersteunende personeel. Rob, bedankt om de vakgroep te leiden, in mijn begeleidingscommissie te zetelen en voor de vele positieve gesprekken die ik met je heb gehad. Michael, jouw statistische expertise heeft menige doctoraatstudent en postdoc geholpen. Bedankt voor je beschikbaarheid en snelle hulp.

Ik wil mijn gezin oprecht bedanken om de constante te zijn doorheen de laatste vier, soms stormige, jaren:

Ilke, ik kan geen woorden vinden voor de positieve invloed die jij op mijn leven hebt. Bedankt Illie, voor de steun en de bereidheid en openheid om altijd (maar dan ook altijd) mee te denken met mij. Of het nu vanuit een ander land, een andere stad of slecht 1 km verderop is, jouw liefde voel ik en neem ik mee in al mijn ondernemingen.

Mama, ik kwam altijd op de eerste plaats in je gedachten en zorgen. De onvoorwaardelijke trots die je voelt voor mij en het enthousiasme voor alles wat ik doe, hebben me deze vier jaar door moeilijke momenten geholpen en me de nodige moed gegeven vol te houden.

Louis, bedankt voor je onuitputtelijke intellectuele nieuwsgierigheid, die mij meermaals inspireerde om het allemaal langs een andere kant te bekijken en voor de traditionele binnenkomstknuffel die me steeds terug met beide voeten in het hier en nu brengt.

Papa, ook al ben je ver weg, ons wekelijks facetime-gesprek zorgt voor een verkleining van de afstand en weet dat dit erg heeft uitgemaakt de laatste vier jaren.

Er zijn nog een paar vrienden die voor mij een belangrijke rol hebben gespeeld tijdens het afwerken van mijn doctoraat.

Hanne, bobje, mijn maatje door dik en dun, mede-zwartkijker en enorm waardevolle persoon. Niet alleen de laatste vier jaar, maar al zeven jaar, hebben wij alles samen meegemaakt en daarin ben je voor mij een enorme steun geweest.

Hanna, we hebben tweeëneenhalf jaar samen gewoond terwijl we zwoegden aan ons doctoraat. Dit verbond, de kameraadschap, het pintelieren en de feestjes in ‘ons’ huis zal ik nooit vergeten. Je bent een fantastische vrouw.

Louisa, mijn thesis-, bureau-, en promotor-genoot. Ik heb jou meerdere keren als voorbeeld in mijn hoofd gehouden en zo inspireerde je me vaak om door te zetten.

Aster en Eowyn. Jullie zijn een eilandje van zon en strand voor mij. Blijf jullie zelf doorheen de volgende drie, misschien soms zware, jaren. Zo is het goed.

Ik wil afsluiten met iemand die als geen ander weet wat dit doctoraat voor mij betekent. Jonne, je hebt een fantastisch brein en gaat complexiteit en uitdaging niet uit de weg. Het was een intellectueel plezier mijn ideeën aan jou af te toetsen. Belangrijker nog, drie jaar lang heb jij met mij meegeleefd en me vaak geholpen de moeilijke momenten te relativeren. Mijn liefste, bedankt.

Uschi

Juni 2015

CHAPTER 1
Introduction to Bilingual Reading

The limits of my language mean the limits of my world.

Ludwig Wittgenstein, Tractatus Logico-Philosophicus

Multilingualism is not rare. Contrary to popular belief, it is actually the norm in most modern societies: More than half of the world population has knowledge of more than one language (Grosjean, 2010). Only a small number of languages, such as English, Mandarin Chinese, Arabic, and Spanish are known by a very large group of people and are used as languages of communication around the world. Indeed, research shows that these languages are often second or even third languages (e.g., Edwards, 1994). One of the reasons why the omnipresence of bilingualism is often overlooked, is that people assume that bi- or multilingualism requires equal, native-like proficiency in all of the known languages and that second and third languages should be acquired in childhood. In fact, this is not how bilingualism is usually defined in politics or science (for examples: Eurobarometer, 2006; Grosjean, 1992). Another possible reason for the underestimation is that governments have the tendency to overstate the normative nature of monolingualism due to conservative language policies (Crystal, 2003). This idea is supported by the fact that only 25% of the European countries are officially multilingual, while the European commission endorses the explicit goal for each European citizen to speak at least two languages, in addition to their mother tongue (Eurobarometer, 2006). A final reason is that monolingualism is often assumed to be the norm in linguistic and language acquisition theories (Ellis, 2006), whereas in effect this is not the case.

This last misunderstanding is one of the things that make bilingual language research so exciting. In our world, language is central in any concept of human thinking. In fact, some writers even claim that language and cognition are inseparable and one cannot exist without the other (e.g. (Jackendoff, 1992; Wittgenstein, 1929). Nowadays, most researchers agree that language is at least one of the mechanisms through which humans construct their reality (Swoyer, 2011; Whorf & Carroll, 1956). Due to this central role of language in the human mind, cognitive psychologists have studied language extensively, whilst incorporating findings from other disciplines, such as cognitive neuroscience. With the evolution from monolingual to largely bilingual societies, concepts of language use change as well and fields of research change with them. This opens up an endless amount of research questions inherently tied to the cognitive, linguistic and non-linguistic, functioning of a bilingual person: How do people manage to acquire fluency in a second language? Is it possible to gain expert fluency in two (or more) languages? How do people keep track of the languages they are using? What mechanisms are responsible for language control? Do bilingual people have a disadvantage in reading or speaking compared to monolingual people? These questions do not only invite us, researchers, to think about language in a very different way, they also, when answered, advance the knowledge about an enormous challenge that the majority of the world population tackles every day: the fluent use of multiple languages.

In this introduction we will first focus on a task central to our culture and education system, reading. What cognitive processes are involved in this task? How does a person go from letters on a page to a word that has form and meaning? The second part of the introduction will guide us through the challenges that a bilingual person necessarily has to face when reading. Here we zoom in on the focus of this dissertation, namely how bilinguals perform word recognition in sentence context, a seemingly basic part of reading. Finally, we will provide an overview of the goals and the chapters of this dissertation.

READING

Humans are born with innate capacities of language development (e.g. Lucy, 1996, Haun, 2007); babies learn to babble and to understand speech without any formal instruction. Although our current society is littered with written language, literacy has to be attained by formal instruction and education (Pinker, 1997). This is because literacy skills are, evolutionary speaking, relatively new compared to language production abilities (Deacon, 1998): Reading and writing are cultural inventions of only a few thousand years old. The first known symbolic scripts, such as Sumerian and Egyptian hieroglyphs emerged at around 3300 B.C., and the first alphabetic script emerged at 750 B.C. in Greece (Gelb, 1963). In alphabetic scripts, the basis unit represented by a grapheme is essentially a phoneme, although the transparency of this correspondence can vary.

In acquiring the skill to use this alphabetic script, the human brain, possessing no evolutionary developed hard-wired areas for this skill, has to make connections between regions in the brain devoted to more basic processes, such as visual object recognition and spoken language. Indeed, there are no brain regions that are uniquely devoted to reading (Dehaene, Le Clec'H, Poline, Le Bihan, & Cohen, 2002). A large meta-analysis of imaging studies during reading (Bolger, Perfetti, & Schneider, 2005) showed that all four lobes of the brain are recruited in reading. The brain's capacity to make new connections between brain regions, to develop extremely specialized pattern recognition areas and the ability to recruit different areas from the brain in an automatic way (Hebb, 1949), is what makes it possible for a child to learn how to read (Dehaene et al., 2002). For a more detailed description of which brain regions are involved in the different cognitive aspects of reading, see Box 1. In other words, it is the sheer plasticity of the brain (Neville & Bavelier, 2000) and the human capacity to obtain cognitive fluency through practice that enables a child to become a proficient adult reader. In the next sections, we will explain how readers recognize words and how the recording of eye movements can provide a way in which to look at this process in a sentence context.

BOX 1: THE READING BRAIN

During the entire reading process, executive processes such as attention and memory are needed. Attention must be disengaged, moved and focused. These attentional processes activate multiple regions in the brain, such as the back of the parietal lobe and the frontal part of the cingulate gyrus (Posner, Walker, Friedrich, & Rafal, 1984). This last structure is also involved in working memory, which is crucial for holding on to information about what we read (Baddeley, 1992).

The visual cortex is responsible for recognizing the lines on the paper as letters. There is some evidence for the existence of a visual word form area in the occipital-temporal region, sensitive to whether letter strings form words or not (Dehaene et al., 2002).

Another huge part of reading is the control of the movements of the eyes. This process activates a huge network in the brain, often referred to as the oculomotor loop (Alexander, DeLong, & Strick, 1986). The frontal eye fields send activation to the caudate nucleus. This structure projects via direct and indirect pathways to the midbrain, more specifically to the superior colliculi. These structures support the integration of cognitive and sensory information and send motor commands to the brain stem. Other structures, such as the vestibular cerebellum and the thalamus, are essential in monitoring the eye movements that are made.

In the later stages of reading, the temporal lobe is both active in semantic processing and phonological processing (Demonet et al., 1992), for example during the detection of regularity in the language. This is the area of the brain responsible for the well-known N400 ERP-component elicited when a word evokes integration difficulties (e.g., Holcomb, 1988). Also, syntactic parsing information of the sentence is processed in Broca's area (Zurif, Swinney, Prather, Solomon, & Bushell, 1993) and the left temporal areas.

It becomes clear that during reading all of these brain regions must interact with one another to make reading an efficient enterprise. Even more remarkable, all of these processes take only 500ms for an expert reader (Posner & McCandliss, 1999).

HOW DO WE READ?

'And so to completely analyze what we do when we read would almost be the acme of a psychologist's achievement, for it would be to describe very many of the intricate workings of the human mind, ...'

The Psychology and Pedagogy of reading, Huey (1908)

When we turn our attention towards the act of proficient reading, we can see what a tremendous feat of cognitive effort it is. Readers have to recruit multiple mental processes, such as attention and memory, as well as visual and linguistic processes to succeed in their task at hand. It would be immensely complex to model all of these aspects of reading. It is therefore that most models of reading focus on only a small portion of the reading process (for an overview of the most important ones we refer to Harley, 2013). In this dissertation we will focus on how a reader gains access to the correct lexical representation, which contains form information about the words that we know. All of these lexical representations together constitute a person's lexicon (or mental *dictionary*).

VISUAL WORD RECOGNITION

The reading task consists ultimately of the transformation of a collection of symbols to letters, the combination of these letters to words that contain meaning and finally the integration of that word within the sentence and larger context at hand. The first two steps of this task are addressed in the field of visual word recognition. A printed word must activate an orthographic word form stored in the mental lexicon and somehow gain access to the knowledge that person has about this word form. Although there have been other models (the autonomous serial search model by Forster (1976) and the logogen model by Morton (1969) are the most important ones), the Interactive Interaction Model (IA model) of McClelland and Rumelhart (1981) has had by far the most influence on how researchers view visual word recognition (see Figure 1). It is one of the earliest connectionist models and consists of three levels of processing units. The lowest level contains visual feature units, coding the visual input, i.e. the lines on the paper. The next level consists of letter units. The last is the word unit level. There are both excitatory and inhibitory connections between the different levels. The letter and word units have inhibitory connections amongst themselves, introducing competition in the visual word recognition system. This means that when a visual feature is activated, it sends activation along to all of the letter units it is connected to, either increasing or decreasing the activation of the letter unit. These activated letter units feed forward activation to compatible word units. These in turn can send feedback to letter units. This feedback explains the word superiority effect, which is the observation that a letter is recognized easier in the context of a word than in isolation (Reicher, 1969; Wheeler, 1970).

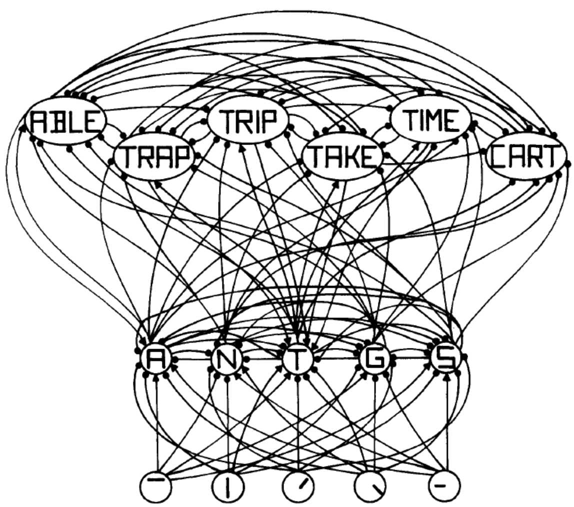


Figure 1. An example of the possibly activated network during the recognition of the letter 'T' according to the architecture of the Interactive Activation model. Arrows indicate excitatory connections; circular ends of the connections represent inhibitory connections. Figure taken from McClelland and Rumelhart (1981), page 830.

The lexical representation that is the first to pass a certain threshold of activation is selected. After this, all of the information stored together with that representation, such as syntactical and morphological specifications and semantic information should become available to the reader. The details of these later processes are not described within the IA model. Although influential, the IA model is thus limited in scope, because semantics and phonology are largely ignored. Other models of visual word recognition have attempted to broaden the scope by including phonology and semantics in their architecture. These models focus largely on how written words are transformed into sound. Interesting ones are the dual route-cascaded model (Coltheart & Rastle, 1994) and the PMSP model of reading (Plaut, McClelland, Seidenberg, & Patterson, 1996), which was derived from the connectionist triangle model of Seidenberg and McClelland (1989), of which the orthographic part shows great overlap with the IA model.

Of course, reading entails a lot more than gaining lexical access to separate words and being able to retrieve the meaning and the sound of the correct word. Readers have to parse the syntactic structure of a sentence, often dealing with ambiguities. Also, selecting the correct word meaning is not as straightforward as we would think it to be. One orthographic word form can have multiple meanings and the sentence context helps to decide which meaning is the correct one. Swinney (1979) discovered that the brain does not retrieve one meaning for a word, but instead activates a rich semantic network of associated words and knowledge. This implies that our language background largely defines our activation network during reading. Readers also make use of large quantities of non-linguistic knowledge to truly *understand* text. Given the cognitive effort all of these processes need, it is fair to say that the reading of complete novels, seemingly without effort, is a staggering human achievement. As said, we will focus on the first step in the reading process, namely the mechanisms behind lexical access to a word. Eye movements while reading could provide an interesting way to study this process.

EYE MOVEMENTS IN READING

The process of visual word recognition, described in detail in the previous section, only takes about 200ms per word. A large part of our apparent ease of reading text comes from the automaticity of the eye movements that we make when we read. During reading, our eyes seem to effortlessly glide across words, sentences, and pages. This smooth movement is an illusion. The eye's trajectory consists of fast movements, called saccades, and brief moments of stand-still, called fixations. Visual information is only obtained during these fixations and vision is suppressed while the eyes are moving (Wolverton & Zola, 1983). These blanks of visual input are masked by the brain and unnoticeable to the reader. The foveal region is the area of vision that is in focus and subtends about 2 degrees of visual angle (mostly 3-6 letters) around the fixation point. The parafoveal region extends about 10 degrees of visual angle (mostly 24 letters) around the fixation. Here, shapes are distinguished, but individual letters are not (Just & Carpenter, 1987). Research has indicated that although the parafoveal region is symmetric, the perceptual span, i.e. the area of effective vision during reading, is asymmetric: it is larger to the right of the fixation than to the left of it for languages that are read from left to right. In English, this region extends about 16 letters to the right. This means that we have a preview of what will come next when reading, enhancing the speed of the processing of the next word. The typical saccade moves our eyes forward about 8 letters at a time and an average fixation lasts about 200-250 ms (Rayner, 2009). A counter-intuitive finding is that, instead of looking at each word in turn, the eyes skip about 25% of all words in a text (Rayner, 1997). This does not happen at random. Function words (these are words that indicate the structural relationship between words such as 'why' and 'the') are skipped more often than content words, short words are skipped more often than long words, and words that are highly predictable from sentence context are more likely to be skipped than less predictable words. The eye also moves backwards in the text to reread certain words about 10-15% of the time. This basic description of eye movements while reading already shows that eyes are not merely driven by visual features, but are heavily influenced by cognitive processing of the text.

THE LINK BETWEEN THE EYE AND THE MIND

A large amount of studies (Just & Carpenter, 1980; Reichle, Pollatsek, Fisher, & Rayner, 1998) suggest that there is a close connection between the eye and the mind. An example is the tight link between word frequency and fixation duration; low frequent words are fixated longer than high frequent words (Just & Carpenter, 1980; Kennedy, Pynte, Murray, & Paul, 2013;

Rayner, Liversedge, White, & Vergilino-Perez, 2003). This effect is still found in disappearing text paradigms, where the text disappears after an interval of 50-60ms (Blythe, Liversedge, Joseph, White, & Rayner, 2009; Liversedge et al., 2004; Rayner, Yang, Castelhano, & Liversedge, 2011). This indicates that even though the word is no longer there, its frequency determines how long the eyes stay at the same position. Another example is the fact that words that are highly predictable given the sentence context are read faster than words that are unpredictable (Balota, Pollatsek, & Rayner, 1985; Morris, 1994). These effects are mostly found on the fixation on the target word, not on the fixation after this word, meaning that there is no discernable eye-to-mind lag. This means that eye movements during reading are perfectly suited to investigate cognitive processes in reading (Rayner & Liversedge, 2011). Rayner goes so far as to say that ‘we have learned as much about reading from eye-movement studies as from any other source of data.’ (Rayner, 1999).

The recording of these eye movements is called eye tracking. The method of eye tracking provides the researcher with an incredibly rich data record: the output is the complete sequential pattern of fixations and saccades for each reader and each portion of text with an extremely high temporal (the location of the eye can be recorded up to twice every millisecond) and spatial resolution (even indicating which part of a letter is fixated). To make the eye movement record interpretable, it is often summarized by concatenating fixations across a certain region of interest (often words, (e.g. Rayner, Sereno, & Raney, 1996) or sentences (Blythe et al., 2009; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989)) or by calculating the probability of skipping a word or regressing backwards into the text. Box 2 provides the reader with an overview of the first studies using the method of eye tracking for reading research.

BOX 2: HISTORY OF THE EYE TRACKING METHODOLOGY IN READING

The use of eye-tracking to study reading dates back to Huey (1908). Although Huey and his contemporaries used rather crude invasive recording devices, they laid the foundations of what is now known about eye movements during reading. Huey's findings that the eyes regress a small percentage of the time and that readers only fixate about 50 percent of the words they encounter still remain valid today.

About 14 years later, Judd and Buswell, (1922) and Buswell (1922) were the first to conduct non-invasive eye tracking experiments with the use of photographs. They also drew attention to the fact that the sentence context is crucial for determining the meaning of a word. Tinker (1936) was the first researcher that actively asked the question of the ecological validity of the eye tracking methodology. His conclusion was that eye movement registration does reveal authentic reading behavior. This opened the door to the extension of experimental findings to natural situations. In his last review article, Tinker (1958) stated that everything that could be learned from eye tracking had been learned.

Maybe because of Tinkers statement, and maybe because of the behiorist doctrine of those years (Rayner & Pollatsek, 1989), it wasn't until the late 1960's that eye tracking was picked up again as a valuable tool for studying reading. One of the reasons for this resurgence was the development of more precise eye trackers that could be interfaced with computers enabling data to be gathered and analyzed on a much larger scale. Another reason is that with the expansion of cognitive psychology, eye movements were needed for the testing of more intricate detailed hypotheses derived from complex theories. Keith Rayner's research is iconic for this era of eye movement research (e.g. McConkie & Rayner, 1975; Rayner, 1975a, 1975b, 1978, 1979; Rayner & McConkie, 1976).

After the 1970's, eye tracking has been used in all fields of reading research, from language acquisition to lexical access. A fruitful approach has been the gathering of massive eye tracking corpora; eye movements recorded over a large amount of participants or large amount of material or both, although only very few of such corpora exist. Advantages of these corpus studies are that they usually consist of a more representative set of stimuli than classical factorial experimental designs do and that these corpora provide researchers with enough data to study a broad range of phenomena without collecting new data. An example is the Potsdam Corpus (Kliegl, Nuthmann, & Engbert, 2006), which presented constructed isolated German sentences to 222 subjects. The SWIFT model of eye movements was developed using the data of this corpus (Engbert, Nuthmann, Richter, & Kliegl, 2005). A rare example of an eye tracking corpus which presented text in paragraphs is the Dundee corpus (Kennedy & Pynte, 2005). This corpus has been successfully used by different research groups to develop, among others, accounts of parafoveal on foveal processing (Kennedy & Pynte, 2005; Pynte & Kennedy, 2006) and models of syntactic processing (Demberg & Keller, 2008; Fossum & Levy, 2012; Frank & Bod, 2011).

A MODEL OF EYE MOVEMENTS IN READING: THE E-Z READER MODEL

The most influential model of eye movements during reading is the E-Z reader model (Pollatsek, Reichle, & Rayner, 2006; Rayner, Ashby, Pollatsek, & Reichle, 2004; Reichle et al., 1998; Reichle, Rayner, & Pollatsek, 1999; Reichle, Warren, & McConnell, 2009). This model was inspired by many of the findings of basic eye movement studies. An important observation is that words are processed on more than one fixation (Rayner & Pollatsek, 1989). Also the fact that word frequency and word length both contribute independently to eye movements (Inhoff & Rayner, 1986; Rayner & Fischer, 1996) has influenced the development of this model.

In the E-Z reader model, the programming of saccades is decoupled from the shifting of attention, which is allocated serially to only a single word at a time (Reichle et al., 1998). The completion of the *familiarity check* process, prior to lexical access, of the word *n*, is the trigger that prompts the oculomotor system to program a saccade directed towards the next word *n+1*. The subsequent *completion of lexical access* of word *n,* causes attention to shift from word *n* to word *n+1*. Because attention shifts are faster than the programming of a saccade (Rayner, Reichle, Stroud, Williams, & Pollatsek, 2006), the lexical processing of word *n+1* usually begins when the eyes are still fixated on word *n*. This feature of the model allows parafoveal processing of upcoming words. Following similar reasoning, the model predicts that parafoveal words, which are processed fast enough, might be skipped. The model assumes that word length, frequency and predictability are important lexical variables that have a large effect on the eye movements, because these variables define the duration of the familiarity check and the completion of lexical access. Consequently, they determine fixation duration, fixation count, rightward saccade length, skipping and regression rates.

BILINGUALISM

It is fascinating to think that people speak 7 102 different languages worldwide, 286 of which are spoken on the European continent (Lewis, Simons, & Fennig, 2015). Seeing that there are 47 countries in Europe, this means that multiple languages are spoken within the same political region. In Europe though, not that many countries are officially bilingual or multilingual, but many offer an official status of some kind to minority languages. Belgium, Luxembourg, and Switzerland are good examples of countries that have institutionalized multilingualism. The Netherlands and Spain have important minority languages, Frisian and Catalan/Basque respectively. On top of this language diversity, English (among other languages) is rapidly taking its place as the global language of choice (Crystal, 2012), being taught in more than 100 countries, making it the most widely taught foreign language. As stated in the beginning of this introduction, most early theories of language have always considered the monolingual person as the default language user. This underlines the importance of bilingual researchers investigating aspects of language processing inherently tied to bilingual language use.

In order to investigate a concept, one must first grasp the concept at hand. This is what we will attempt to do in the next section. After this, bilingual research is summarized with a special focus on effects of word frequency and language exposure and cross-lingual interactions. A separate section is dedicated to bilingual eye tracking research. Finally, a model of bilingual word recognition is presented.

BILINGUALISM AS A CONCEPT

In the case of bilingualism, this first question of what actually defines a *bilingual* person, already proves to be a difficult one (Hoffmann, 2014). What is clear is that next to knowledge of a native language, often referred to as L1, a person must have knowledge of a second language, often referred to as L2. Earlier definitions of bilingualism tend to be narrower than current ideas. For example Bloomfield (1933) stated that being bilingual should entail equal, native-like, proficiency in both languages. Later, Haugen (1953) proposed a broader definition of bilingualism, namely that when a person can utter complete and meaningful utterances in another language than their mother tongue, they can be called bilingual. However, Grosjean's (1982) definition of what is means to be bilingual, is used most commonly in scientific research. He proposed that some knowledge and practical use of a second language is sufficient to qualify as a bilingual (Grosjean, 1982). The reason for the popularity of his definition is of course partly of a practical nature; it's use does not require testing detailed proficiency of participants before qualifying as a bilingual. A more important reason for its use is that this definition is supported by experimental data. There are for example effects of very short (15 minutes) second language learning on performance of non-language cognitive-tasks (e.g. Boroditsky, 2001), as well as multiple artificial language learning experiments, that show that after short learning sessions, participants can work with, and do operations on an artificial language recently acquired (Folia, Uddén, De Vries, Forkstam, & Petersson, 2010).

Next to differences in language combinations, in language dominance (which of the two languages can be considered as more proficient than the other), in the general levels of proficiency, and in the ways of acquiring a second language, there are multiple modalities of language: reading, speaking, writing, and listening. We will attempt to illustrate this diversity with an example: Two native Flemish people that are living in Brussels and are defined as a bilingual person according to Grosjean's (1982) definition, might have completely different levels of skills in the different modalities of language: Whereas the first might read French novels but is unable to order a cup of tea in a French café, the other one might have the opposite problem. He/She might work in a French environment, using French everyday in communication with others, but might falter in reading Proust. On top of the differences across modalities, the second person might have knowledge of a third language, English. Or these two persons might have acquired French in a different time in their lives, referred to as the age of acquisition of a language. When looking at this example and the diversity in bilingualism, it becomes clear that one satisfactory well-defined definition of bilingualism might be impossible and perhaps even undesirable.

This brings us to another exciting characteristic of bilingual research: Researchers are faced with the challenge of establishing their own boundary conditions for defining bilingualism and second language proficiency as their research needs dictate. This offers a lot of intellectual freedom, but can also make replication of findings across research groups rather challenging, certainly in comparison with other, more uniform cognitive concepts. Because bilingualism has so many different manifestations and is hard to define objectively, there is a long history of the use of self-report measures in bilingual research (e.g. Marian, Blumenfeld, & Kaushanskaya, 2007). This holds the implicit assumption that people can correctly estimate the level of their own language skills. It has been shown however that people perform rather poorly on this front (Gollan, Weissberger, Runnqvist, Montoya, & Cera, 2012; Sheng, Lu, & Gollan, 2014). A solution to this problem is documenting this variation in skills, age of acquisition, language dominance and language history in the modality under investigation, to be able to relate these variations to variations in the studied language behavior. Lately, more and more researchers are looking for objective ways in establishing second language proficiency (e.g., Lemhöfer & Broersma, 2012). Throughout this dissertation, we consistently favor the more objective way of measuring second, but also first language proficiency.

RESEARCH INTO BILINGUALISM

On a theoretical level, bilingual research can advance our understanding of language processing in general. On a practical level, seeing that bilingualism in some form or another is spread worldwide, it is important to investigate the implications of bilingualism on language and other cognitive processes. Findings in bilingual research can eventually advance knowledge on the optimal conditions and time point of learning and teaching a second language.

A large research tradition has attempted to clarify the organization of the bilingual lexicon and how the properties of this bilingual lexicon influence the mechanisms used for lexical access. Once a vocabulary in a second language is acquired, how are bilinguals able to identify the correct representation in their mental lexicon? Most of this research, as the research is this dissertation will as well, has focused on three, interdependent questions: a) When bilinguals have more lexical representations than monolinguals do, are they disadvantaged in the process of gaining lexical access to target words? b) Is there an integrated bilingual lexicon or one separate lexicon for each language? c) Is lexical access to this lexicon language non-selective or language selective?

Seeing that most research on the bilingual lexicon implies that there in one integrated lexicon (for a review: Dijkstra & van Heuven, 2002), another large field of research emerged. This field has dedicated itself to studying how the bilingual brain uses control mechanisms in suppressing activation of the non-target language, while using a target language. The inhibitory control model of Green (1998) proposes that bilinguals are able to use both of their languages by means of dynamic regulation. The general executive control system is responsible for this inhibition and activation (Green, 1998), not some language specific mechanism. Keeping this idea in mind, researchers have found evidence supporting effects of bilingualism on cognitive control in non-linguistic tasks (for a recent overview see Bialystok, 2009). One of the most surprising findings is that being a bilingual can even delay the onset of Alzheimer dementia (Craik, Bialystok, & Freedman, 2010; Woumans et al., in press).

In the field of language acquisition research, bilingualism is used as a tool for studying both first as well as second language acquisition. In children, developing a language for the first time, the processes of language and cognitive development are confounded, whereas in late bilinguals these two can be disentangled (Cook, 1981), because their cognitive development is already progressed. Also, it is a useful tool for studying whether there is a critical period in language learning, which refers to the hypothesis that the first few years in life are crucial for language development (Lenneberg, 1967; Singleton, 2005), and whether native-like proficiency can be acquired after a certain age (White & Genesee, 1996).

We will now return to the research questions concerning the organization of the bilingual lexicon.

LEXICON SIZE, LANGUAGE EXPOSURE AND WORD FREQUENCY IN BILINGUAL WORD RECOGNITION

Bilinguals are different from monolinguals: They have more lexical representations in their mental lexicon than monolinguals do. It is possible that this vocabulary growth causes a disadvantage when a bilingual person attempts to access one of his/her lexical representations. In the field of language production, there is a small consistent amount of evidence for production disadvantages for bilinguals compared to monolinguals. Bilinguals are slower in naming pictures than monolinguals are (Gollan, Montoya, Fennema-Notestine, & Morris, 2005; Ivanova & Costa, 2008) and they are more likely than monolinguals to fail in the naming of a picture, but have the feeling that lexical access is imminent. This is called a *Tip of The Tongue* experience (Gollan & Acenas, 2004; Gollan & Silverberg, 2001). The *weaker links account* or *frequency lag hypothesis* (Gollan & Acenas,

2004) explains these bilingual disadvantages in the context of the larger size of the bilingual's lexicon and the bilingual's language exposure. In this view, according to the weaker links account, it is not the search process in this larger lexicon, but the fact that all of the lexical items in the integrated lexicon have had fewer exposures or practice, which causes bilinguals to perform worse in production tasks. In short, bilinguals necessarily divide the frequency of use of their words between languages (Gollan & Acenas, 2004), making the representations and the connections to these representations *weaker* in the bilingual language system than in the monolingual one, leading to slower lexical access for bilinguals (Gollan, Montoya, Cera, & Sandoval, 2008).

Because the authors assume weaker connections in the bilingual language system, they expect these bilingual disadvantages to also arise in reading (Gollan et al., 2011). The authors do predict that the bilingual disadvantages and frequency effects will be smaller in comprehension because the processes needed for language production are less practiced, more difficult and involve more levels of processing where frequency is important (Gollan et al., 2011). Indeed, Ransdell and Fischler (1987) and Lehtonen et al. (2012) found that bilinguals' reaction times were slower than those of monolinguals in an L1 lexical decision task. Both authors attribute this effect to the lower exposure of L1 words for bilinguals compared to monolinguals. From this description, it is clear that the value of the weaker links hypothesis is dependent on what kind of bilingual groups are studies, because the authors of the weaker links account assume that total language exposure is the same across groups (Gollan & Acenas, 2004; Gollan et al., 2008), leading to a lower L1 exposure for bilinguals compared to monolinguals.

Word frequency is a central concept in understanding the relationship between a possible disadvantage in bilingual lexical access and language exposure. Word frequency attempts to capture how many times a person has encountered a word and it is one of the most studied lexical variables in visual word recognition research (Brysbaert et al., 2011). It is clear that we cannot measure the precise word frequency for every individual. This is why word frequency is usually approximated by the frequency of a word in a specific set of written sources (e.g., Subtlex databse by Keuleers, Brysbaert, & New, 2010). Although this variable accounts for a very large portion of variance in visual word recognition, specifically for a bilingual's second language word frequencies acquired by corpus counts are rather inaccurate. Kuperman and Van Dyke (2013) showed that for monolinguals with smaller vocabularies there is an overestimation of the frequency of low frequent words. Indeed, low frequent words are more sensitive to differences in exposure: additional exposure to a word a person was exposed to only a few times, affects processing time in a major way, whereas the same additional exposure to a high frequent word has a much smaller effect on processing time (Monsell, 1991; Morton, 1970). This means that, given a lower exposure to L2, frequency effects would be larger for bilinguals in L2 than in L1. Also, when assuming a difference in L1 exposure for bilinguals compared to monolinguals (Gollan et al., 2008), the frequency effect would be larger for bilinguals than for monolinguals. And indeed, Duyck, Vanderelst, Desmet, and Hartsuiker (2008) and Whitford and Titone (2012) reported larger frequency effects in L2 word recognition than in L1. When L1 proficiency was matched between bilinguals and monolinguals a similar L1 frequency effect was found for both groups (Duyck et al., 2008; Gollan et al., 2011). When L1 proficiency differed, the frequency effect was larger for bilinguals than for monolinguals (Lehtonen et al., 2012; Lemhöfer et al., 2008).

CROSS-LINGUAL ACTIVATION IN BILINGUAL WORD RECOGNITION

ISOLATED WORD RECOGNITION

The majority of studies of bilingual lexical access have focused on detecting cross-lingual activation. Most of these studies have used experimental tasks that presented words in isolation to the participants, such as lexical decision tasks, word naming tasks or perceptual identification tasks. In the word-naming task, participants read printed words aloud and reaction times are recorded. In the lexical decision task, printed words and non-words are presented. The participants must decide whether the word exists or not with 'yes' or 'no' button press. Usually reaction times and accuracies are recorded. In the perceptual identification tasks, words are either degraded or presented too briefly to identify them with certainty. Participants are asked to repeat the words or to press a button once they think they have identified the word.

Critical to the notion of language independent lexical access is that representations from the non-target language become activated when the target language is being processed. By comparing responses to control words with responses to words that show some overlap across languages, researchers can test this hypothesis.

Cross-lingual homographs are orthographic word forms that occur in both languages, but have two distinct meanings in the bilingual's different languages. An example is the word 'coin', meaning *corner* in French. Beauvillain and Grainger (1987) were the first to test cross-language activation using homographs. In a priming study they showed that both meanings of the homograph were activated in a mixed language context. Because of this mixed language context, bilinguals might be in a multilingual mode (Grosjean, 1998) and they may actively use both languages, making the cross-lingual activation of representations hardly surprising. More compelling evidence of activation of non-target language representations comes from studies presenting words in a unilingual context (e.g., de Groot, Delmaar, & Lupker, 2000; Dijkstra, Grainger, & van Heuven, 1999). Dijkstra et al. (1999) found faster response to homographs without phonological overlap in a lexical decision task, while De Groot et al. (2000) found slower responses for homographs without controlling for phonological overlap.

Another kind of words that is used frequently to study cross-lingual activation is orthographic neighbor words. Orthographic neighbors of a target word are usually defined by changing one letter, while preserving the position of the other letters (Coltheart, Davelaar, Jonasson, & Besner, 1977). The few cross-lingual orthographic neighborhood studies have shown that full orthographic overlap is not needed to find evidence for cross-lingual activation (Bijeljac-Babic, Biardeau, & Grainger, 1997; van Heuven, Dijkstra, & Grainger, 1998). These studies find longer reaction times to cross-lingual neighbors in a lexical decision task.

Both cross-lingual neighbors and homographs share (almost) the same orthography across languages. Another category of words has been used extensively in the search for cross-lingual activation: cognates. Cognates are translation equivalent words that have at least some degree of orthographic overlap between the two languages. A good example of an identical cognate is the word ‘rat’ in English and Dutch; an example of a non-identical cognate is the word ‘apple’ in English and the word ‘appel’ in Dutch. The majority of these studies have shown faster reaction times to cognates in lexical decision tasks (Caramazza & Brones, 1979; Duyck, Van Assche, Drieghe, & Hartsuiker, 2007; Gollan, Forster, & Frost, 1997; Lemhöfer, Dijkstra, & Michel, 2004; Van Hell & Dijkstra, 2002), progressive demasking tasks (Dijkstra et al., 1999), and semantic categorization tasks (Sanchez-Casas, Davis, & García-Albea, 1992).

All of the discussed word-isolation research points towards the conclusion that lexical access is non-selective. However, it is possible that lexical access to a word presented in a sentence context might be restricted to the target language. The language of the sentence might serve as a kind of language cue to the reader, signaling that there is no need to activate the other known language, thus eliminating the cross-lingual effects found in isolation research.

BILINGUAL WORD RECOGNITION IN SENTENCE CONTEXT

To be able to extend the evidence for language independent lexical access during contextual language use, Elston-Güttler, Gunter, and Kotz (2005) presented homographs as the final words in an L2 sentence context. After this a target word was presented for lexical decision. They showed that context might indeed restrict lexical access to one language, because there was no difference in reaction times for homographs compared to control words. Schwartz and Kroll (2006) embedded homographs and cognates in either high constraining or low constraining contexts. There was no effect of homographs or cognates embedded in high constraining contexts on naming latencies, but there was cognate facilitation in the low constraining sentence context. This showed that the semantic nature of the sentence context could help restrict lexical access to one language (for similar results see van Hell & de Groot, 2008), but independent lexical access is still possible in sentence context.

Grainger and Jacobs (1996) pointed out that tasks such as lexical decision, word naming and word identification, all tap into task-specific processes, such as decision-making, the sounding out of words and the resolving of degraded visual input, next to the process of lexical access. This means that the absence of cross-lingual interactions in high constraining sentences might be due to processes after lexical access has taken place. A more sensitive and ecologically valid methodology, like eye tracking, might be more suited. This method does not expect participants to make decisions or produce any utterances.

BILINGUAL EYE TRACKING RESEARCH INTO LEXICAL ACCESS

In contrast to the monolingual research domain, studies that track eye movements during bilingual reading in order to study lexical access are rather rare. So far, only two bilingual studies have used eye tracking to investigate word frequency effects. The first one by Gollan et al. (2011) recorded the eye movements of 36 English monolinguals, 36 balanced Spanish-English bilinguals and 27 unbalanced Dutch-English bilinguals while reading target words embedded in high or low constraint sentences. They found no modulation of the frequency effect for the different groups. The second study by Whitford and Titone (2012) recorded eye movements from unbalanced French-English and English-French bilinguals who read English and French paragraphs. Here, frequency effects were found to be larger when reading in L2. Also L2 proficiency modulated the size of the frequency effect: bilinguals with a lower L2 proficiency, showed larger frequency effects in an L2 context. This is the first study showing that a graded difference in proficiency co-varies with the size of frequency effects. Later this was confirmed by the results of an analysis of word identification data by Diependaele, Lemhöfer, and Brysbaert (2013). However, the effect of graded L1 proficiency on frequency effects in monolingual and bilingual reading has never been studied.

The first study using eye tracking to investigate bilingual cross-lingual activation was conducted by Duyck et al. (2007). They embedded cognates and control words in low constraint L2 sentence contexts. The eye movement pattern showed that identical cognates were processed faster than control words. Hereafter, multiple eye tracking studies were conducted to investigate cognate effects in bilingual L2 reading of sentences and paragraphs (Balling, 2013; Bultena, Dijkstra, & van Hell, 2014; Libben & Titone, 2009; Pivneva, Mercier, & Titone, 2014; Van Assche, Drieghe, Duyck, Welvaert, & Hartsuiker, 2011) and even in L1 reading (Titone, Libben, Mercier, Whitford, & Pivneva, 2011; Van Assche, Duyck, Hartsuiker, & Diependaele, 2009). Van Assche et al. (2009) found cognate facilitation for Dutch-English bilinguals reading low constraint sentences in L1. Titone et al. (2011) recorded eye movements from English-French bilinguals reading cognates or interlingual homographs embedded within low and high constraint sentences. They found no cognate facilitation in high constraint sentences. They detected a cognate facilitation effect only in low constraint sentences and only for bilinguals who started learning their L2 from a very early age. They concluded that bilingual L1 reading is influenced by L2, only under certain conditions. To summarize, cognate facilitation has never been detected in a highly semantically constrained L1 context, implying that lexical access might be restricted to L1 when reading more predictable naturalistic text.

Almost all of these eye-tracking studies have used isolated, constructed sentences as materials. Only Balling (2013) and Whitford and Titone (2012) have used paragraphs in their experiments. This might be an important issue when one is trying to draw conclusions about bilingual reading as it happens day tot day. First, experimentally constructed sentences are unlikely to be representative of sentences in naturalistic texts. Second, most daily reading happens in a larger context than just one isolated sentence. It has been shown that reading passages might bring about changes in the eye movement pattern. Radach, Huestegge, and Reilly (2008) found that the total processing time of words is longer, but that the first fixations are much shorter for reading passages than when reading isolated sentences. Radach et al. explain this by suggesting that readers, when going through larger passages of text, perform a fast first 'pass' across the text followed by a rereading of the passage. It would be very interesting to see whether for a bilingual reading in one language, the non-target language is still engaged in this more strategic form of reading.

Given the small amount of eye tracking studies directed at the investigation of bilingual lexical access, it is not surprising that there is no large data set of bilingual eye movements available to researchers. We strongly believe that the availability of such an eye tracking corpus would dramatically improve the opportunities of testing hypotheses concerned with bilingual reading in a natural context. We have already explained that in the monolingual literature these kinds of corpora have been used for language model development and evaluation. A similar bilingual data set could do the same in the bilingual field.

MODELING EFFORTS OF BILINGUAL LEXICAL ACCESS IN CONTEXT

Most evidence points towards the fact that a bilingual person has one integrated lexicon, to which he has access to via a language independent mechanisms (at least under most conditions). The Bilingual Interactive Activation (BIA) model (Dijkstra & Van Heuven, 1998) was designed as a bilingual extension of the Interactive Activation (IA) model (McClelland & Rumelhart, 1981). Critical to the BIA model is that it assumes parallel language independent access to one integrated lexicon that includes all of the lexical representations of both languages. The BIA model assumes that these lexical representations send activation to language node representations. These language nodes could also be affected by factors outside the lexicon such as instructions. They send top-down activation towards the lexical representations, influencing the selectivity of lexical access. The most recent adaptation of this model, the BIA+ model (Dijkstra & van Heuven, 2002), eliminates the ability of language nodes to send top-down feedback (see Figure 2).

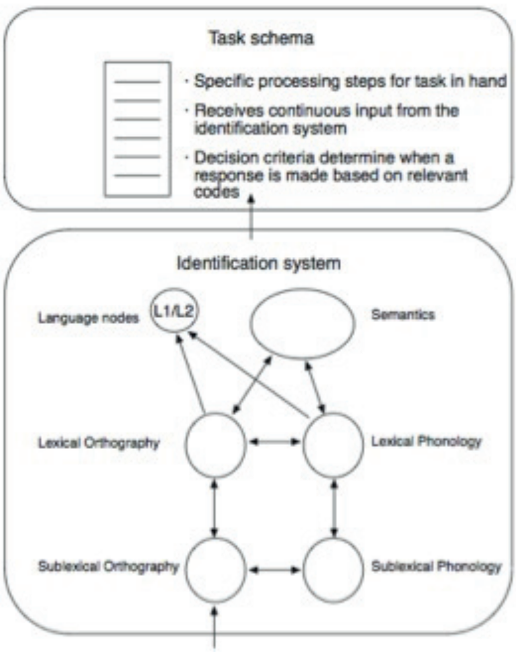


Figure 2. The architecture of the BIA+ model. Figure taken from Dijkstra and van Heuven (2002), page 183.

Also, language nodes are now only influenced by activation within the lexicon. Another change is that in the BIA+ model a strict distinction is made between the lexicon and a task/decision system. Effects of instruction and individual expectations do not impact lexical activation but occur after lexical access is complete. Linguistic context, such as sentence context however, does impact activation of representation within the lexicon directly. How this works exactly is not made explicit (Dijkstra & van Heuven, 2002).

There have been other attempts of modeling bilingual lexical access in sentence contexts, but these efforts were mostly focused on how lexical access is attained for words that are semantically ambiguous across languages, such as cross-lingual homographs. An example is the three-factor framework of Degani and Tokowicz (2010). They added language context as a third factor to the monolingual reordered-access model (Duffy, Morris, & Rayner, 1988), which uses semantic /syntactic sentence context and frequency of the meaning of the target word in ambiguity resolution of homonyms and homographs.

The most cited model of monolingual eye movements, the E-Z reader model, discussed in the previous section, has been used to explain reading patterns of other groups of language users, such as older readers (Rayner et al., 2006), younger readers (Rayner et al., 2006; Reichle et al., 2013) and even non-alphabetic languages (Rayner, Li, & Pollatsek, 2007). Interestingly, it has never been used to explain the differences between bilingual and monolingual eye movement behavior. The previous successful applications of the E-Z reader model, illustrate that this framework could be useful in future modeling efforts concerned with bilingual eye movement patterns, and we will therefore align our analyses of bilingual reading behavior with the core assumptions and variables of this model.

In conclusion, the BIA+ model is the only implemented computational model of bilingual lexical access with an explanatory goal both in isolation and in sentence context. In this dissertation we will therefore use the BIA+ model and its assumptions of an integrated bilingual lexicon with language independent lexical access as a framework for our hypotheses generation and our interpretation of the results.

CURRENT DISSERTATION

GOALS

This dissertation strives to achieve two complementary goals. The first important one is of a methodological nature. We aim to develop a large data set of reading behavior in a natural context, until now absent in the field of bilingualism. In this dissertation, we gathered a large corpus of eye movements of Dutch-English bilinguals and English monolinguals reading a coherent novel. The methodological framework and the motivation to develop this corpus are discussed extensively in CHAPTER 2. We aim to disclose this corpus for other researchers to use freely.

The second goal of this dissertation is to use this eye-tracking corpus to investigate bilingual lexical access in L1 and L2 sentence processing in a natural reading context. In this dissertation we will focus both on sentence level (CHAPTER 3) and word level processing (CHAPTER 4-6) to shed light on the mechanisms underlying bilingual lexical access. Finally, in CHAPTER 7, we discuss the theoretical and practical implications of the results reported in CHAPTERS 2-6, as well as some intriguing conceptual and methodological issues one has to consider when conducting language (bilingual) research. Given that we focus in this dissertation on the first processes of reading, lexical access, it is important to point out that other linguistic processes, such as the parsing of the syntactic structure of a sentence, selecting the correct meaning of a word given the sentence context and using non-linguistic knowledge to comprehend text, can also be studied using this eye tracking corpus. Therefore we also present some pointers for future uses of the eye-tracking corpus in CHAPTER 7.

METHODOLOGICAL APPROACH: CORPUS OF EYE TRACKING

To construct the eye-tracking corpus, we record eye movements from bilinguals reading half of the novel in L1, the other half in L2, and from monolinguals reading the entire novel in their mother tongue. In CHAPTER 2, we discuss in depth the reasons for gathering a large eye-tracking corpus of bilingual naturalistic reading and the possible applications of such a corpus. Other (monolingual) corpora of eye-movements and their applications are discussed in the introduction of this chapter. We show the distribution of the most important eye movement duration variables in the corpus for L1 and L2 bilingual reading and monolingual reading and report objective measures of skewness and deviation from the normal distribution. In this chapter, the reader can find links to in-depth participant information, material information and the eye-tracking data itself.

BASIC SENTENCE-LEVEL PARAMETERS OF BILINGUAL READING

The first empirical aim of the dissertation is to assess global differences between bilingual L1 and bilingual L2, on the one hand, and bilingual L1 and monolingual eye movement patterns, on the other hand, when reading natural text. In CHAPTER 3, we discuss how most bilingual eye-tracking research has focused on analyzing eye movements at the word-level, for example fixation durations on target words embedded in sentences. In this chapter, we provide the reader with a detailed sentence-level description of both bilingual and monolingual eye movement behavior over a set of translation equivalent sentences, extracted from the corpus. Additionally, we conduct two comparisons across languages. The first one is a within-subjects comparison between the eye movement pattern of bilingual readers reading in L1 (Dutch) and L2 (English). The second one is a between-subjects comparison between the eye movement pattern of the same bilingual readers reading in L1 (Dutch) and monolinguals reading in their first and only language (English). This last comparison is especially interesting in regard to the weaker links theory, which expects lexical access to be slower for bilinguals than for monolinguals in L1 reading (Gollan et al., 2011). In this chapter we thus pay close attention to language proficiency as a factor in the possible disadvantages. Because the comparisons are conducted across languages, inherent differences might influence the result patterns. This is why we selected only those sentences that were matched pairwise with the translation equivalent sentence on number of words per sentence, average word length and average content word frequency.

The eye-movement parameters of sentence reading and the results of these comparisons can be found in CHAPTER 3.

EFFECTS OF WORD FREQUENCY AND LANGUAGE EXPOSURE

For our first word level analysis, we start by investigating the most important, and most influential, lexical variable in word recognition research, word frequency, in CHAPTER 4. We examine these word frequency effects in detail and with attention for individual language proficiency differences. Word frequency explains a large part of the variance in reaction times during visual word recognition (Brysbaert et al., 2011). The hypothesis has been put forward that bilingual frequency-effects should be larger than monolingual frequency effects (Gollan et al., 2008), for reasons of lexical exposure and the nature of the word frequency distribution. We provide a critical viewpoint on this idea in CHAPTER 4 and focus on differences in language exposure and accuracy of the corpus frequencies for different groups of participants. We investigate the size of the word frequency effect in the eye-movements observed on all of the content words in the novel for bilingual and monolingual reading. This is the first investigation of effects of L1 and L2 proficiency on frequency effects in natural reading. The results are presented in CHAPTER 4.

CROSS-LINGUAL INTERACTIONS

In CHAPTER 5 and CHAPTER 6, we turn our focus to effects of cross-lingual overlap between lexical representations. Our eye-tracking corpus is especially interesting in this respect because of the chances it provides to solidify the large evidence database for language independent lexical access, expanding it to include strictly unilingual naturalistic contexts. The silent reading of an entire novel requires little intervention from the researcher, providing a window into natural language processing. In the bilingual research domain, a lot of data (e.g. (Dijkstra, Timmermans, & Schriefers, 2000; Lemhöfer et al., 2008; van Heuven & Dijkstra, 1998) has supported the existence of parallel and language independent lexical access during bilingual visual word recognition (BIA+ model by (Dijkstra & van Heuven, 2002). This means that during reading in the target language, a bilingual person activates not only lexical candidates from the target language but also of the non-target language. Most studies have used cognates to investigate this issue. Research in a L1 sentence context however, is rather rare (Titone et al., 2011; Van Assche et al., 2009) and is lacking in an L1 paragraph context. In CHAPTER 5 we investigate cognate facilitation for the nouns in the corpus in both early and late eye movements in both L1 and L2 context. We discuss the results and the implications for cognate representation in this chapter. Another, some would argue more direct, way to investigate cross-lingual activation of lexical candidates is to look at processing time of cross-lingual orthographic neighbors. Until now, only one study (van Heuven & Dijkstra, 1998) investigates this in a series of isolated-word experiments in isolated and mixed language contexts. In CHAPTER 6 we examine these cross-lingual neighborhood effects both in a generalized lexical decision task and in our eye tracking corpus of L1 and L2 reading.

REFERENCES

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annual Review of Neuroscience, 9, 357–381. http://doi.org/10.1146/annurev.ne.09.030186.002041

Balling, L. W. (2013). Reading authentic texts: What counts as cognate? Bilingualism: Language and Cognition, 16(03), 637–653. http://doi.org/10.1017/S1366728911000733

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. Cognitive Psychology, 17(3), 364–390. http://doi.org/10.1016/0010-0285(85)90013-1

Beauvillain, C., & Grainger, J. (1987). Accessing Interlexical Homographs: Some Limitations of a Language-Selective Access. Journal of Memory and Language, 26, 658–672.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. Bilingualism: Language and Cognition, 12(01), 3–11. http://doi.org/10.1017/S1366728908003477

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

Bloomfield, L. (1933). Language. University of Chicago Press.

Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2009). Visual information capture during fixations in reading for children and adults. Vision Research, 49(12), 1583–1591. http://doi.org/10.1016/j.visres.2009.03.015

Bolger, D. J., Perfetti, C. A., & Schneider, W. (2005). Cross-cultural effect on the brain revisited: universal structures plus writing system variation. Human Brain Mapping, 25(1), 92–104. http://doi.org/10.1002/hbm.20124

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: a review of recent developments and implications for the choice of frequency estimates in German. Experimental Psychology, 58(5), 412–424. http://doi.org/10.1027/1618-3169/a000123

Bultena, S., Dijkstra, T., & van Hell, J. G. (2014). Cognate effects in sentence context depend on word class, L2 proficiency, and task. The Quarterly Journal of Experimental Psychology, 67(6), 1214–1241. http://doi.org/10.1080/17470218.2013.853090

Buswell, G. T. (1922). Fundamental reading habits: a study of their development. University of California Libraries.

Caramazza, A., & Brones, I. (1979). Lexical Access in Bilinguals. Bulletin of the Psychonomic Society, 13(4), 212–214.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), Attention and Performance VI (pp. 535–555). London: Academic Press.

Coltheart, M., & Rastle, K. (1994). Serial Processing in Reading Aloud - Evidence for Dual-Route Models of Reading. Journal of Experimental Psychology-Human Perception and Performance, 20(6), 1197–1211. http://doi.org/10.1037/0096-1523.20.6.1197

Cook, V. J. (1981). Some uses for second language learning research. Annals of the New York Academy of Sciences, 379, 251–258.

Craik, F. I. M., Bialystok, E., & Freedman, M. (2010). Delaying the onset of Alzheimer disease. Neurology, 75(19), 1726–1729. http://doi.org/10.1212/WNL.0b013e3181fc2a1c

Crystal, D. (2003). The Cambridge Encyclopedia of the English Language (2 edition). Cambridge, U.K. ; New York: Cambridge University Press.

Crystal, D. (2012). English as a Global Language (2 edition). Cambridge, UK ; New York: Cambridge University Press.

Deacon, T. W. (1998). The Symbolic Species: The Co-evolution of Language and the Brain. W. W. Norton & Company.

Degani, T., & Tokowicz, N. (2010). Semantic ambiguity within and across languages: An integrative review. The Quarterly Journal of Experimental Psychology, 63(7), 1266–1303. http://doi.org/10.1080/17470210903377372

De Groot, A. M., Delmaar, P., & Lupker, S. J. (2000). The processing of interlexical homographs in translation recognition and lexical decision: support for non-selective access to bilingual memory. The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology, 53(2), 397–428. http://doi.org/10.1080/713755891

Dehaene, S., Le Clec'H, G., Poline, J.-B., Le Bihan, D., & Cohen, L. (2002). The visual word form area: a prelexical representation of visual words in the fusiform gyrus. Neuroreport, 13(3), 321–325.

Demberg, V., & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. Cognition, 109(2), 193–210. http://doi.org/10.1016/j.cognition.2008.07.008

Demonet, J., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J., Wise, R., … Frackowiak, R. (1992). The anatomy of phonological and semantic processing in normal subjects. Brain: A Journal of Neurology, 115(6), 1753–1768.

Diependaele, K., Lemhöfer, K., & Brysbaert, M. (2013). The word frequency effect in first-and second-language word recognition: a lexical entrenchment account. The Quarterly Journal of Experimental Psychology, 66(5), 843–863.

Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of Cognates and Interlingual Homographs: The Neglected Role of Phonology. Journal of Memory and Language, 41(4), 496–518.

Dijkstra, T., Timmermans, M., & Schriefers, H. (2000). Cross-language effects on bilingual homograph recognition. Journal of Memory and Language, 42, 445–464.

Dijkstra, T., & Van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. Jacobs (Eds.), Localist Connectionist Approaches to Human Cognition (pp. 189–225). New York: Mahwah.

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. Bilingualism: Language and Cognition, 5(03). http://doi.org/10.1017/S1366728902003012

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. Journal of Memory and Language, 27(4), 429–446. http://doi.org/10.1016/0749-596X(88)90066-6

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. Journal of Experimental Psychology: Learning, Memory, and Cognition, 33(4), 663–679. http://doi.org/10.1037/0278-7393.33.4.663

Duyck, W., Vanderelst, D., Desmet, T., & Hartsuiker, R. J. (2008). The frequency effect in second-language visual word recognition. Psychonomic Bulletin & Review, 15(4), 850–855. http://doi.org/10.3758/PBR.15.4.850

Edwards, J. (1994). Multilingualism (1 edition). London ; New York: Routledge.

Ellis, E. (2006). Monolingualism: The unmarked case. Estudios de Sociolinguistica, 7(2), 173–196.

Elston-Güttler, K. E., Gunter, T. C., & Kotz, S. A. (2005). Zooming into L2: global language context and adjustment affect processing of interlingual homographs in sentences. Brain Research. Cognitive Brain Research, 25(1), 57–70. http://doi.org/10.1016/j.cogbrainres.2005.04.007

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A Dynamical Model of Saccade Generation During Reading. Psychological Review, 112(4), 777–813. http://doi.org/10.1037/0033-295X.112.4.777

Eurobarometer, S. (2006). Europeans and their Languages. European Commission. Retrieved from http://pol.gu.se/digitalAssets/759/759844_Europeans_and_their_Languages_-_EC_2006.pdf

Folia, V., Uddén, J., De Vries, M., Forkstam, C., & Petersson, K. M. (2010). Artificial Language Learning in Adults and Children: Artificial Language Learning in Adults and Children. Language Learning, 60, 188–220. http://doi.org/10.1111/j.1467-9922.2010.00606.x

Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. Walker (Eds.), New Approaches to Language Mechanisms (pp. 257–287). Amsterdam.

Fossum, V., & Levy, R. (2012). Sequential vs. hierarchical syntactic models of human incremental sentence processing. In Proceedings of the 3rd Workshop on Cognitive Modeling and Computational Linguistics (pp. 61–69). Association for Computational Linguistics. Retrieved from http://dl.acm.org/citation.cfm?id=2390313

Frank, S. L., & Bod, R. (2011). Insensitivity of the Human Sentence-Processing System to Hierarchical Structure. Psychological Science, 22(6), 829–834. http://doi.org/10.1177/0956797611409589

Gelb, I. J. (1963). A Study of Writing. (2nd Ed.). Chicago: University of Chicago Press. Retrieved from https://balshanut.wordpress.com/2007/08/05/gelb-ij-a-study-of-writing-2nd-ed-chicago-university-of-chicago-press-1963/

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), 246–269. http://doi.org/10.1037/0278-7393.30.1.246

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew–English bilinguals. Journal of Experimental Psychology: Learning, Memory, and Cognition, 23(5), 1122.

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. Journal of Memory and Language, 58(3), 787–814. http://doi.org/10.1016/j.jml.2007.07.001

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). http://doi.org/10.1017/S136672890100013X

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. Journal of Experimental Psychology: General, 140(2), 186–209. http://doi.org/10.1037/a0022256

Gollan, T. H., Weissberger, G. H., Runnqvist, E., Montoya, R. I., & Cera, C. M. (2012). Self-ratings of spoken language dominance: A Multilingual Naming Test (MINT) and preliminary norms for young and aging Spanish–English bilinguals. Bilingualism: Language and Cognition, 15(03), 594–615. http://doi.org/10.1017/S1366728911000332

Grainger, J., & Jacobs, A. M. (1996). Orthographic Processing in Visual Word Recognition: A Multiple Read-out Model. Psychological

Review, 103(3), 518–565. http://doi.org/10.1037/0033-295X.103.3.518
Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. Bilingualism: Language and Cognition, 1(02), 67–81. http://doi.org/10.1017/S1366728998000133
Grosjean, F. (1982). Preface. In Life with two languages: An introduction to Bilingualism. (p. viii). Cambridge, MA: Harvard University Press.
Grosjean, F. (1992). Another view of bilingualism. In Cognitive processing in bilinguals (pp. 51–62). Oxford, England: North-Holland.
Grosjean, F. (1998). Transfer and language mode. Bilingualism: Language and Cognition, 1(03), 175–176. http://doi.org/10.1017/S1366728998000285
Grosjean, F. (2010). Bilingual: Life and Reality. Harvard University Press.
Harley, T. A. (2013). The Psychology of Language: From Data to Theory. Psychology Press.
Haugen, E. (1953). The Norwegian language in America; a study in bilingual behavior. Philadelphia: University of Pennsylvania Press.
Hebb, D. O. (1949). The Organization of Behavior. New York: Wiley.
Hoffmann, C. (2014). Introduction to Bilingualism. Routledge.
Holcomb, P. J. (1988). Automatic and attentional processing: an event-related brain potential analysis of semantic priming. Brain and Language, 35(1), 66–85.
Huey, E. B. (1908). The Psychology And Pedagogy Of Reading.
Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. Perception & Psychophysics, 40(6), 431–439. http://doi.org/10.3758/BF03208203
Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? Acta Psychologica, 127(2), 277–288. http://doi.org/10.1016/j.actpsy.2007.06.003
Jackendoff, R. (1992). Languages of the Mind. Cambridge, MA: MIT Press.
Judd, C. H., & Buswell, G. T. (1922). Silent Reading: A study of the various types. Chicago: University of Chicago Press.
Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. Psychological Review, 87(4), 329.

Just, M. A., & Carpenter, P. A. (1987). The psychology of reading and language comprehension (Vol. x). Needham Heights, MA, US: Allyn & Bacon.

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. Vision Research, 45(2), 153–168. http://doi.org/10.1016/j.visres.2004.07.037

Kennedy, A., Pynte, J., Murray, W. S., & Paul, S.-A. (2013). Frequency and predictability effects in the Dundee Corpus: An eye movement analysis. The Quarterly Journal of Experimental Psychology, 66(3), 601–618. http://doi.org/10.1080/17470218.2012.676054

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. Behavior Research Methods, 42(3), 643–650. http://doi.org/10.3758/BRM.42.3.643

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the Mind During Reading: The Influence of Past, Present, and Future Words on Fixation Durations. Journal of Experimental Psychology: General, 135(1), 12–35. http://doi.org/10.1037/0096-3445.135.1.12

Kuperman, V., & Van Dyke, J. A. (2013). Reassessing word frequency as a determinant of word recognition for skilled and unskilled readers. Journal of Experimental Psychology: Human Perception and Performance, 39(3), 802–823. http://doi.org/10.1037/a0030859

Lehtonen, M., Hultén, A., Rodríguez-Fornells, A., Cunillera, T., Tuomainen, J., & Laine, M. (2012). Differences in word recognition between early bilinguals and monolinguals: behavioral and ERP evidence. Neuropsychologia, 50(7), 1362–1371. http://doi.org/10.1016/j.neuropsychologia.2012.02.021

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. Behavior Research Methods, 44(2), 325–343. http://doi.org/10.3758/s13428-011-0146-0

Lemhöfer, K., Dijkstra, T., & Michel, M. C. (2004). Three Languages, One ECHO: Cognate Effects in Trilingual Word Recognition. Language and Cognitive Processes, 19(5), 585–611.

Lemhöfer, K., Dijkstra, T., Schriefers, H., Baayen, R. H., Grainger, J., & Zwitserlood, P. (2008). Native language influences on word recognition in a second language: A megastudy. Journal of

Experimental Psychology: Learning, Memory, and Cognition, 34(1), 12–31. http://doi.org/10.1037/0278-7393.34.1.12

Lenneberg, E. (1967). Biological Foundations of Language. New York.

Lewis, M. P., Simons, G. F., & Fennig, C. D. (2015). Ethnologue: Languages of the World (18th edition). Dallas, Texas: SIL International. Retrieved from http://www.ethnologue.com

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. Journal of Experimental Psychology: Learning, Memory, and Cognition, 35(2), 381–390. http://doi.org/10.1037/a0014875

Liversedge, S. P., Rayner, K., White, S. J., Vergilino-Perez, D., Findlay, J. M., & Kentridge, R. W. (2004). Eye movements when reading disappearing text: is there a gap effect in reading? Vision Research, 44(10), 1013–1024. http://doi.org/10.1016/j.visres.2003.12.002

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. Journal of Speech, Language, and Hearing Research, 50(4), 940–967.

McClelland, J., & Rumelhart, D. (1981). An interactive activation model of context effects in letter perception. An account of basic findings. Psychological Review, 88, 375–407.

McConkie, G. W., & Rayner, K. (1975). The Span of Effective stimulus during a fixation in reading. Perception & Psychophysics, 17, 578–586.

Monsell, S. (1991). The Nature and Locus of Word Frequency Effects in Reading. In D. Besner & G. W. Humphreys (Eds.), Basic Processes in Reading: Visual Word Recognition (pp. 148–197). Hillsdale, NJ, England: Erlbaum.

Morris, R. K. (1994). Lexical and message-level sentence context effects on fixation times in reading. Journal of Experimental Psychology: Learning, Memory, and Cognition, 20(1), 92–103. http://doi.org/10.1037/0278-7393.20.1.92

Morton, J. (1969). Interaction of information in word recognition. Psychological Review, 76(2), 165–178. http://doi.org/10.1037/h0027366

Morton, J. (1970). A Functional Model for Memory. In D. A. Norman (Ed.), Models of Human Memory (pp. 203–254). Elsevier.

Neville, H. J., & Bavelier. (2000). Specificity and Plasticity in Neurocognitive Development in Humans. In M. S. Gazzaniga (Ed.), The New Cognitive Neurosciences. MIT Press.

Pinker, S. (1997). Foreword. In Why our children can't read. New York: Free Press.

Pivneva, I., Mercier, J., & Titone, D. (2014). Executive Control Modulates Cross-Language Lexical Activation During L2 Reading: Evidence From Eye Movements. Journal of Experimental Psychology-Learning Memory and Cognition, 40(3), 787–796. http://doi.org/10.1037/a0035583

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. E. (1996). Parallel Distributed Processing and Cognitive Neuroscience. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.57.2570&rep=rep1&type=pdf

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. Cognitive Psychology, 52(1), 1–56. http://doi.org/10.1016/j.cogpsych.2005.06.001

Posner, M. I., & McCandliss, B. D. (1999). In R. M. Klein & P. McMullen (Eds.), Converging Methods for Understanding Reading and Dyslexia (pp. 305–338). MIT Press.

Posner, M. I., Walker, J. A., Friedrich, F. J., & Rafal, R. D. (1984). Effects of parietal injury on the covert orienting of visual attention. Journal of Neuroscience, 9, 81–93.

Pynte, J., & Kennedy, A. (2006). An influence over eye movements in reading exerted from beyond the level of the word: Evidence from reading English and French. Vision Research, 46(22), 3786–3801. http://doi.org/10.1016/j.visres.2006.07.004

Radach, R., Huestegge, L., & Reilly, R. (2008). The role of global top-down factors in local eye-movement control in reading. Psychological Research, 72(6), 675–688. http://doi.org/10.1007/s00426-008-0173-3

Rayner, K. (1975a). Parafoveal Identification During a Fixation in Reading. Acta Psychologica, 39(4), 271–282. http://doi.org/10.1016/0001-6918(75)90011-6

Rayner, K. (1975b). Perceptual Span and Peripheral Cues in Reading. Cognitive Psychology, 7(1), 65–81. http://doi.org/10.1016/0010-0285(75)90005-5

Rayner, K. (1978). Eye movements in reading and information processing. Psychological Bulletin, 85, 618–660.

Rayner, K. (1979). Eye Guidance in Reading: Fixation Locations in Words. Perception, 8, 21–30.

Rayner, K. (1997). Understanding Eye Movements in Reading. Scientific Studies of Reading, 1(4), 317–339. http://doi.org/10.1207/s1532799xssr0104_2

Rayner, K. (1999). What have we learned about eye movements during reading? In R. M. Klein & P. McMullen (Eds.), Converging Methods for Understanding Reading and Dyslexia (pp. 23–56). MIT Press.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. The Quarterly Journal of Experimental Psychology, 62(8), 1457–1506. http://doi.org/10.1080/17470210902816461

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The Effects of Frequency and Predictability on Eye Fixations in Reading: Implications for the E-Z Reader Model. Journal of Experimental Psychology: Human Perception and Performance, 30(4), 720–732. http://doi.org/10.1037/0096-1523.30.4.720

Rayner, K., & Fischer, M. H. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. Perception & Psychophysics, 58(5), 734–747. http://doi.org/10.3758/BF03213106

Rayner, K., & Liversedge, S. P. (2011). Linguistic and cognitive influences on eye movements during reading. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.), The Oxford Handbook of Eye movements (pp. 751–766). UK: Oxford University Press.

Rayner, K., Liversedge, S. P., White, S. J., & Vergilino-Perez, D. (2003). Reading Disappearing Text Cognitive Control of Eye Movements. Psychological Science, 14(4), 385–388. http://doi.org/10.1111/1467-9280.24483

Rayner, K., Li, X., & Pollatsek, A. (2007). Extending the E-Z Reader Model of Eye Movement Control to Chinese Readers. Cognitive

Science, 31(6), 1021–1033. http://doi.org/10.1080/03640210701703824

Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements? Vision Research, 16, 829–837.

Rayner, K., & Pollatsek, A. (1989). The Psychology of Reading. Hillsdale, NJ, England: Erlbaum.

Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. Psychology and Aging, 21(3), 448–465. http://doi.org/10.1037/0882-7974.21.3.448

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C. (1989). Eye movements and on-line language comprehension processes. Language and Cognitive Processes, 4(3-4), SI21–SI49. http://doi.org/10.1080/01690968908406362

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: a comparison of two types of models. Journal of Experimental Psychology: Human Perception and Performance, 22(5), 1188.

Rayner, K., Yang, J., Castelhano, M. S., & Liversedge, S. P. (2011). Eye movements of older and younger readers when reading disappearing text. Psychology and Aging, 26(1), 214–223. http://doi.org/10.1037/a0021279

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. Journal of Experimental Psychology, 81(2), 275–280. http://doi.org/10.1037/h0027768

Reichle, E. D., Liversedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2013). Using E-Z Reader to examine the concurrent development of eye-movement control and reading skill. Developmental Review, 33(2), 110–149. http://doi.org/10.1016/j.dr.2013.03.001

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. Psychological Review, 105(1), 125.

Reichle, E. D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: Accounting for initial fixation locations and refixations within the EZ Reader model. Vision Research, 39(26), 4403–4411.

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher level language processing on eye movements during reading. Psychonomic Bulletin & Review, 16(1), 1–21. http://doi.org/10.3758/PBR.16.1.1

Sanchez-Casas, R., Davis, C., & García-Albea, J. E. (1992). Bilingual Lexical Processing -Exploring the cognate non-cognate distinction. European Journal of Cognitive Psychology, 4(4), 293–310.

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. Journal of Memory and Language, 55(2), 197–212. http://doi.org/10.1016/j.jml.2006.03.004

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. Psychological Review, 96(4), 523.

Sheng, L., Lu, Y., & Gollan, T. H. (2014). Assessing language dominance in Mandarin-English bilinguals: Convergence and divergence between subjective and objective measures. Bilingualism (Cambridge, England), 17(2), 364–383. http://doi.org/10.1017/S1366728913000424

Singleton, D. M. (2005). The Critical Period Hypothesis: A coat of many colours. International Review of Applied Linguistics in Language Teaching, 43(4), 269–285.

Swinney, D. A. (1979). Lexical Access during Sentence Comprehension (Re)Consideration of Context Effects. Journal of Verbal Learning and Verbal Behavior, 18, 645–659.

Swoyer, C. (2011). How does language affect thought? In V. J. Cook & B. Bassetti (Eds.), Language and Bilingual Cognition (pp. 23–42). New York: Psychology Press.

Tinker, M. A. (1936). Reliability and validity of eye-movement measures of reading. Journal of Experimental Psychology, 19, 732–746.

Tinker, M. A. (1958). Recent studies of eye movements in reading. Psychological Bulletin, 55, 215–231.

Titone, D., Libben, M., Mercier, J., Whitford, V., & Pivneva, I. (2011). Bilingual lexical access during L1 sentence reading: The effects of L2 knowledge, semantic constraint, and L1–L2 intermixing. Journal of Experimental Psychology: Learning, Memory, and Cognition, 37(6), 1412–1431. http://doi.org/10.1037/a0024492

Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. Journal of Memory and Language, 64(1), 88–107. http://doi.org/10.1016/j.jml.2010.08.006

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. Psychological Science, 20(8), 923–927.

Van Hell, J. G., & de Groot, A. M. B. (2008). Sentence context modulates visual word recognition and translation in bilinguals. Acta Psychologica, 128(3), 431–451. http://doi.org/10.1016/j.actpsy.2008.03.010

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. Psychonomic Bulletin & Review, 9(4), 780–789.

Van Heuven, W. J. B., & Dijkstra, T. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. Journal of Memory and Language, 39, 458–483.

Van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. Journal of Memory and Language, 39(3), 458–483. http://doi.org/10.1006/jmla.1998.2584

Wheeler, D. D. (1970). Processes in word recognition. Cognitive Psychology, 1(1), 59–85. http://doi.org/10.1016/0010-0285(70)90005-8

White, L., & Genesee, F. (1996). How native is near-native? The issue of ultimate attainment in adult second language acquisition. Second Language Research, 12(3), 233–265. http://doi.org/10.1177/026765839601200301

Whitford, V., & Titone, D. (2012). Second-language experience modulates first- and second-language word frequency effects: Evidence from eye movement measures of natural paragraph reading. Psychonomic Bulletin & Review, 19(1), 73–80. http://doi.org/10.3758/s13423-011-0179-5

Whorf, B. L., & Carroll, J. B. (1956). Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf. MIT Press.

Wittgenstein, L. (1929). Tractatus logico-philosophicus. (C. K. Ogden, Trans.).

Wolverton, G. S., & Zola, D. (1983). The temporal characteristics of visual information extraction during reading. In K. Rayner (Ed.), Eye movements in reading: Perceptual and language processes. New York: Academic.

Woumans, E., Santens, P., Sieben, A., Versijpt, J., Stevens, M., & Duyck, W. (in press). Bilingualism delays clinical manifestation of Alzheimer's disease. Bilingualism: Language and Cognition. Retrieved from http://journals.cambridge.org/abstract_S136672891400087X

Zurif, E., Swinney, D., Prather, P., Solomon, J., & Bushell, C. (1993). An on-line analysis of syntactic processing in Broca′ s and Wernicke′ s aphasia. Brain and Language, 45(3), 448–464.

CHAPTER 2
PRESENTING GECO: AN EYE-TRACKING CORPUS OF MONOLINGUAL AND BILINGUAL SENTENCE READING[1][2]

This paper opens up the GECO, the Ghent Eye Tracking Corpus, a monolingual and bilingual corpus of eye-tracking data of participants reading a complete novel. English monolinguals and Dutch-English bilinguals read an entire novel, which was presented on the screen in paragraphs. The bilinguals read half of the novel in their first, the other half in their second language. In this paper we describe the distributions and descriptive statistics of the most important reading time measures for the two groups of participants. This large eye-tracking corpus is perfectly suited for both explorative purposes as well as more directed hypothesis testing, and it can guide the formulation of ideas and theories about naturalistic reading processes in a meaningful context. Most importantly, this corpus has the potential to evaluate the ecological validity and overall soundness of complex monolingual and bilingual language theories and models that cover multiple levels of the reading process.

---

1 This chapter is based on a manuscript co-authored by Denis Drieghe and Wouter Duyck: Cop, U., Drieghe, D., & Duyck, W. (2015). Presenting GECO: An Eye-Tracking Corpus of Monolingual and Bilingual Sentence Reading, submitted for publication.

2 Manuscript submitted for publication in Behavior Research Methods.

INTRODUCTION

Over the years, linguistic data gathered in experimental settings have driven the development of ideas and theories about the cognitive processes involved in language performance. Usually, these experiments are designed to test one or more specific hypotheses and use a meticulously selected and restricted stimulus set, containing one or more, often orthogonal, experimental manipulations. More recently, with the development of larger, more complex, computational reading models that operate on multiple processing levels and/or cover a wide range of phenomena (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Demberg & Keller, 2008; Dilkina, McClelland, & Plaut, 2010; Friederici, 1995; Grainger & Jacobs, 1996; Harm & Seidenberg, 2004), the need for data from a larger, more naturalistic, range of stimuli has become more pressing. This kind of data is necessary to evaluate the generalizability and external validity of these broad language models.

And, indeed, it has now been demonstrated that the collection of large amounts of language behavior can have an important role in development, simulation or confirmation of ideas and theories. The studies that collect these large databases are often referred to as corpus studies or mega studies (e.g., Balota et al., 2007; Seidenberg & Waters, 1989). Because corpus studies gather a large amount of observations from a limited amount of participants, or vice versa, or both, they usually have considerable statistical power and can detect relatively small effects. These studies are often defined by the presentation of a large sample of a wide range of unselected stimuli, in contrast to factorial designs used in traditional experimental settings, where stimuli are selected on the basis of specific characteristics. This constricted range usually includes very high and/or very low values and limits the stimulus set to stimuli that are rather extreme in the critical dimension, which may impede representativeness of processing characteristics and show only a part of possible language behavior. An advantage of the corpus approach is that continuous lexical variables, such as word frequency, can be assessed over their full possible range, instead of a constricted one. Another advantage of large corpora of linguistic data is that it enables researchers to answer multiple hypotheses without the need to design a new experiment and gather new data, which is considerably time consuming.

A good example of an influential psycholinguistic corpus study in the field of visual word recognition is the English Lexicon project (ELP: Balota et al., 2007). Balota et al. (2007) gathered lexical decision latencies of 816 participants for 40 481 different American English words (3 400 responses on average per participants). This project sparked the development of similar databases for French (FLP: Ferrand et al., 2010), Dutch (DLP: (Keuleers, Diependaele, & Brysbaert, 2010), and British English (BLP: Keuleers, Lacey, Rastle, & Brysbaert, 2012). These databases have been used to evaluate psycholinguistic ideas about frequency effects (e.g., Kuperman & Van Dyke, 2013), word length effects (e.g., Yap & Balota, 2009), neighborhood effects (e.g., Whitney, 2011; Yap & Balota, 2009) and the lexical decision task itself (Diependaele, Brysbaert, & Neri, 2012; Kuperman, Drieghe, Keuleers, & Brysbaert, 2013), but have also been used to evaluate complex computational models of word recognition (e.g. Norris & Kinoshita, 2012; Whitney, 2011).

EYE-TRACKING CORPORA

Large databases of language performance in reaction to isolated word stimuli are very useful in evaluating specific hypotheses about word recognition and in simulating models, which are mainly concerned with the process of lexical access to an isolated target word. However, when the goal is to explain how reading occurs on a day to day basis in natural context, the ambition of reading models should be to expand their explicatory power beyond word level processes in order to cover a larger scope of interacting language processes. This means that it should operate on more than one level, including a lexical, semantic and/or syntactic level. It is clear that these kinds of models are rather complex, encompassing large numbers of variables on each of these levels and complex interactions between all of these and hypotheses concerning the time course of processing. It is evident that in order to evaluate such models, more detailed and thus complex natural data are needed.

The technique of eye-tracking enables researchers to record the eye movements of participants during silent reading, with minimal instruction or interference on behalf of the researcher. Also, eye tracking, in contrast to for example lexical decision tasks, captures language performance how it occurs in daily life, without interference of additional decision components or response mechanisms. With modern day eye-tracking equipment, the position of the eye can be determined every millisecond with very high spatial accuracy. This technique results in a very rich and detailed descriptive data set. To study visual word recognition in context, the recording of eye movements during reading has been used often (see Rayner (1998) for an introduction and review of early work and Rayner (2009) for a more recent review). Some models of reading have focused on the influence of surrounding words or sentences on the recognition of target words in reading (e.g., Engbert, Nuthmann, Richter, & Kliegl, 2005; Pynte & Kennedy, 2006; Reichle, Pollatsek, Fisher, & Rayner, 1998), All of these models have relied heavily on experimental findings in eye movement research as a way to understand the cognitive processes of reading, but the most important of these models, the E-Z reader model by Reichle et al. (1998), have put eye movements central in their theorizing. The fact that words are processed on more than one fixation (Rayner & Pollatsek, 1989) and that word frequency and word length both contribute independently to eye movements (Inhoff & Rayner, 1986; Rayner & Fischer, 1996) have influenced the central role that lexical access plays in this model.

We think that collecting eye-tracking data over a large sample of stimuli and/or a large sample of subjects increases the richness of eye movement data even more. This is why corpora of eye movements during naturalistic, contextualized reading of text are invaluable in informing and evaluating language models that go beyond the word level, such as the E-Z reader model. These corpora can be used to examine a large number of variables of different processing levels (e.g. both at word and at sentence level) and the interactions among them simultaneously, as well as the specific time course of these effects. Moreover, the evaluation of broad language models with an eye-tracking corpus of natural reading could provide an idea of the ecological validity of parts or whole of the specific model, especially parts which were inspired by findings obtained in less natural tasks.

On top of that, as already discussed for corpora of isolated word recognition, these eye tracking databases a) are perfectly suited to investigate a very broad scale of phenomena: as long as certain syntactic constructions or words with certain lexical traits occur frequently enough in the corpus, they can be studied, b) have a representative unrestricted set of stimuli, which makes generalizing easier, c) provide researchers with data so there is no need to design new experiments or to collect new data, which requires specific expensive equipment and is a time-intensive process, especially for sentence reading.

A rare example of an existing eye-tracking corpus of natural reading is the Dundee Corpus (Kennedy & Pynte, 2005). Ten native French and ten native English subjects read newspaper articles (50 000 words) that were presented in paragraphs on the screen. Eye movements were recorded with a sampling rate of 1ms and spatial accuracy of 0.25 characters. Initially the authors used this corpus to investigate the effect of parafoveal processing on foveal word processing (Kennedy & Pynte, 2005; Pynte & Kennedy, 2006). Later, the same authors investigated the effect of punctuation (Pynte & Kennedy, 2007), the effect of syntactic and semantic constraints on fixation times (Pynte, New, & Kennedy, 2008, 2009a, 2009b), the effect of violations in reading order (Kennedy & Pynte, 2008) and the interaction between frequency and predictability (Kennedy, Pynte, Murray, & Paul, 2013) using the eye movement data of the Dundee corpus.

Other authors also used this corpus to investigate specific hypotheses. Demberg and Keller (2008), for example, investigated subject/object clause asymmetry with the Dundee corpus data and were inspired by these results to build a model of syntactic processing (Demberg & Keller, 2008). The Dundee data was used to evaluate this model. Mitchell, Lapata, Demberg, and Keller (2010) used the Dundee corpus to investigate prediction in sentence reading. A nice illustration of the power of these kinds of corpora is the fact that these authors only needed ten percent of the data to test their hypothesis. Both Frank and Bod (2011) and Fossum and Levy (2012) used the Dundee corpus to evaluate their language models concerned with the role of hierarchical processing in sentence processing. Kuperman et al. (2013) used both the mega data of the Lexicon Project (Balota et al., 2007) and the Dundee corpus (Kennedy & Pynte, 2005) to correlate lexical decision times with natural reading data. Their results showed very low correlations between these measures, implying that these commonly used methods measure, at least to some extent, different things. Their analysis clearly showed that lexical decision data are deficient as a stand-alone tool to evaluate language models and that natural reading data should be used in as a complementary information source to validate models in a more ecological valid situation.

More recently Frank, Fernandez Monsalve, Thompson, and Vigliocco (2013) gathered eye movements from 43 English monolingual subjects reading 205 sentences. Instead of presenting the sentences in paragraphs, as the Dundee corpus does, Frank et al. selected sentences from natural narrative text and presented these sentences seperately on the screen.

Other interesting databases of eye movements in text reading are the German Potsdam corpus (Kliegl, Nuthmann, & Engbert, 2006) and the Dutch DEMONIC database (Kuperman, Dambacher, Nuthmann, & Kliegl, 2010). In the former 222 subjects read 144 constructed German sentences, in the latter 55 subjects read 224 constructed Dutch sentences. These sentences were presented in isolation and did not form a coherent story in any way. This does not mean that these data cannot be used for model construction (SWIFT-model: Engbert et al., 2005) and/or evaluation (see for example Boston, Hale, Kliegl, Patil, & Vasishth, 2008) or hypothesis testing, but only that we find a unselected natural text even more suited for current evaluation purposes and will focus on this approach.

OUR CORPUS: GECO

As the previous section shows, the building of eye-tracking corpora of natural reading can be very fruitful for the development and evaluation of monolingual models of language processing. Until now however, these databases do not focus on, or even specify, possible differences in language knowledge between participants. All eye-tracking corpora (at least to our knowledge) implicitly assume that their participants have knowledge of only the language they are reading in. As bilingualism is most commonly defined as 'the regular use of two (or more) languages' (Grosjean, 1992), today, in most European countries, due to migration and the fact that foreign languages are a compulsory part of formal education, 54 % of the people are bi-or multi-linguals (European Union & European Comission for Education and Culture, 2012). Even in developing countries such as Cameroon, more than half of the population speaks three or more languages (Bamgbose, 1994). In the United States of America, although foreign language courses are not compulsory, about 20 % of the population has some knowledge of a non-native language (Shin & Kominski, 2007).

This is important because a plethora of evidence shows that bilingualism changes language processes and bilinguals need to allocate resources in a different way than monolinguals do. A major finding for instance is that words of both languages are activated in parallel even in unilingual contexts (for a recent review of the evidence see Kroll, Dussias, Bogulski, & Valdes Kroff, 2012).

So far, there are no mega-data available for participants reading in their first language, while they have confirmed and assessed knowledge of another language, or of participants reading in a second language they have acquired later in life. In short, there is no bilingual eye-tracking corpus available to researchers. In this paper we present the GECO, the Ghent Eye-tracking Corpus. It has the goal to bridge this gap, serving both the bilingual and monolingual reading research domain. We gathered eye movement data from monolingual British English participants and Dutch-English bilinguals while they read an entire novel. The bilinguals read half of the novel in L1 reading or L2 reading. All participants read a total of about 5 000 sentences. The precise language history and proficiency score was gathered for all participants. This is the first bilingual corpus study and also the first corpus of Dutch reading of natural text (i.e. not specifically constructed for the experiment). Information on the participants and the materials of the novel as well as the eye-tracking data are available as online supplementary materials[3]. See Appendix A for a list of the available files and the exact contents of the files.

EXPLOITATION OF THE CURRENT CORPUS

The most dominant model of bilingual word recognition is the bilingual interactive activation plus model (BIA+: Dijkstra & van Heuven, 2002). The authors mentioned that this model concerns the visual word recognition system and is part of a larger 'language user' system, also enveloping sentence parsing and language production. They assume that the linguistic (sentence) context has a direct impact on the word recognition system (Dijkstra & van Heuven, 2002), but how exactly is not specified. Because of the contained nature of their model, it has used no eye movement data obtained from natural reading to inform the architecture or evaluate the system of word recognition they propose. Instead, this model has been adjusted from the BIA model (Dijkstra & Van Heuven, 1998) using the findings of a multitude of experimental studies using lexical decision, progressive demasking and identification tasks (e.g. Bijeljac-Babic, Biardeau, & Grainger, 1997; Dijkstra, Timmermans, & Schriefers, 2000; Van Heuven, Dijkstra, & Grainger, 1998), sometimes but rarely in sentence context (Altarriba, Kroll, Sholl, & Rayner, 1996). We believe that the large corpus of eye movements we present here will not only be able to evaluate the ecological validity of this word recognition model, but it should also be especially helpful to specify the exact nature of the interactions between the

[3] URL: http://expsy.ugent.be/downloads/uschi/
Login: uschi
Password: pp02

sentence context and the word recognition system. In their paper presenting the BIA+ model Dijkstra and van Heuven (2002) said,

“Future studies should focus on disentangling such effects of lexical form features and language membership in sentence processing experiments. They should examine, for instance, to which extent the language itself of preceding words in the sentence can modulate the activation of target word candidates from a non-target language.” (Dijkstra & Van Heuven, 2002, p. 187).

Indeed, the GECO has been exploited in this way already. In our own work, we have used the current corpus to examine cross-lingual orthographic neighborhood and cognate effects on a word-level in L1 and L2 reading (Cop, Dirix, Drieghe, & Duyck, 2015; Cop, Van Assche, Drieghe, & Duyck, 2014). The cross-lingual neighborhood of a target word was defined by all words in the non-target language differing only one letter either by deletion, addition or substitution from the target word. We found both facilitatory and inhibitory effects from cross-lingual neighborhood size (Cop, Dirix, et al., 2015). Cognates are translation equivalent words, overlapping in orthography. We found facilitating effects of cognate status both in L1 and L2 reading (Cop, Van Assche, et al., 2014). By investigating these cross-lingual interactions, we provided evidence for a model of bilingual word recognition with a single lexicon and language independent lexical access (see BIA+, Dijkstra & van Heuven, 2002). We showed that, although linguistic context might affect other parts of the word recognition system, the language of the sentence context does not constrain lexical access to the target language.

We have used the GECO in two other ways. By comparing the basic eye movement measures on sentence level between L1 and L2 reading (Cop, Drieghe, & Duyck, 2014), we provided a database of benchmark parameters of reading with attention for the relation between language history and changes in eye movement behavior. Also, we showed that changes in eye movement patterns from L1 to L2 reading pattern closely with the changes from child to adult reading. We also investigated word frequency effects in the corpus in Cop, Keuleers, Drieghe, and Duyck (2015). We showed that frequency effects are larger in L2 than in L1, but that this change is due to the accuracy of corpus based estimates of exposure frequency. These analyses also showed that qualitative differences between monolingual, L1 and L2 language processing are not necessary to account for the differences in frequency effects. These are all examples of the value of such data to investigate specific research questions without the need to collect new data.

A broader use for this data might be the evaluation and adaptation of the E-Z reader model (Reichle et al., 1998), the most dominant model of eye movements, to bilingual reading. As this model has proven to be successful in accommodating eye movement patterns of older (Rayner, Reichle, Stroud, Williams, & Pollatsek, 2006) and younger (Reichle et al., 2013) readers as well as non-alphabetic languages (Rayner, Li, & Pollatsek, 2007), we have reason to believe that it will perform well as a frame work for bilingual eye movement patterns.

METHOD

A more concise version of this method is present in Cop, Keuleers, et al. (2015), because these data were used to investigate frequency effects.

SUBJECTS

Nineteen unbalanced Dutch (L1) – English (L2) bilingual Ghent University and fourteen English monolingual Southampton University undergraduates participated either for course credit or monetary compensation. Bilingual and monolingual participants were matched on age and education level. The average age was 21.2 years for bilinguals [range: 18-24; sd=2.2] and 21.8 years for monolinguals [range: 18-36, sd=5.6]. All of the participants were enrolled in a bachelor or master program of psychology. In the monolingual group, 6 males and 7 females participated. In the bilingual group, 2 males and 17 females participated. Participants had normal or corrected-to-normal vision. None of the participants reported to have any language and/or reading impairments.

The bilinguals started learning their L2 relatively late: The mean age of acquisition was eleven years [range: 5-14, sd = 2.46]. All participants completed a battery of language proficiency tests. This included a vocabulary test, a spelling test, a lexical decision task and a self-report language questionnaire (for results see Table 1). Vocabulary was tested with the LexTALE (Lexical Test for Advanced Learners of English; Lemhöfer & Broersma, 2012). This is an unspeeded lexical decision task, which is an indicator of language proficiency for intermediate to highly proficient language users validated for English, Dutch and German. Due to the lack of a standardized cross lingual spelling test, we tested the English spelling with the spelling list card of the WRAT 4 (Wilkinson & Robertson, 2006) and the Dutch spelling with the GLETSCHR (De Pessemier & Andries, 2009). A classical speeded lexical decision task was also administered in Dutch and English for the bilinguals, in English for the monolinguals. The self report questionnaire was an adaptation of the LEAP-Q (Marian, Blumenfeld, & Kaushanskaya, 2007). This questionnaire contained questions about language switching frequency/skill, age of L2 acquisition, frequency of L2 use and reading/auditory comprehension/speaking skills in L1 and L2 (for a detailed summary, see Table B.1 and B.2 in Appendix B).

Two bilinguals were classified as lower intermediate L2 language users (50%-60%), ten bilinguals were classified as upper intermediate L2 language users (60%-80%), seven bilinguals scored as advanced L2 language users (80%-100%) according to the LexTALE norms reported by Lemhöfer and Broersma (2012). Most important, the Dutch (L1) proficiency of the bilinguals was matched with the English proficiency of the monolinguals for all but subjective exposure (See Table 1), indicating that both groups were equally proficient in their first language, but bilinguals had less relative exposure to their L1 than the monolinguals. The English (L2) proficiency is clearly lower than the Dutch (L1) proficiency (see Table 1).

Table 1. *Average percentage scores [standard deviations] on the LexTALE, the Spelling test, the accuracy of the Lexical Decision task and Subjective Exposure and the score on the comprehension questions for the bilingual and monolingual group. T-values [degrees of freedom] of t-tests are presented in the last 2 columns.*

	Monolinguals	Bilinguals L1	Bilinguals L2	t-value L1-L2	t-value L1-mono
LexTALE score (%)	91.07 [8.92]	92.43 [6.34]	75.63 [12.87]	7.59 [18] ***	0.49 [18]
Spelling score (%)	80.78 [7.26]	83.16 [7.80]	69.92 [8.74]	8.15 [18] ***	0.99 [18]
Lexical Decision Accuracy (%)	77.89 [12.01]	80.19 [5.41]	56.84 [11.12]	9.93 [17] ***	0.67 [17]
Subjective exposure (%)	100.00 [0]	75.00 [15.24]	25.00 [15.24]	7.10 [18] ***	7.10 [18] ***
Comprehension score (%)	78.27 [9.46]	79.63 [10.96]	78.95 [12.54]	0.40 [18]	0.38 [30]

* p<0.05, ** p<0.01, *** p<0.001

MATERIALS

The participants read the novel "The mysterious affair at Styles" by Agatha Christie (Title in Dutch: "De zaak Styles"). This novel was selected out of a pool of books that were available in a multitude of different languages, for possible future replication in other languages, and which did not have any copyright issues. All of these books were selected from the Gutenberg collection that is freely available on the Internet. We selected the novels that could be read in four hours. The remaining books were inspected for difficulty, indicated by the frequency distribution of the words that the book contained. The Kullback–Leibler divergence (Cover & Thomas, 1991) was used to select the novel whose word frequency distribution was the most similar to the one in natural language use, as observed in the Subtlex database (Brysbaert & New, 2009; Keuleers, Brysbaert, & New, 2010).

The monolinguals read the English version of the novel. These participants read a total of 5 031 sentences. The bilinguals read half of the novel, chapter

1-7, in Dutch, the other half, chapter 8-13, in English. The order was counterbalanced. One of the bilingual participants only read the first half of the novel in English. The 10 participants reading the first part of the novel in Dutch, read 2 754 Dutch sentences and 2 449 English sentences. The 8 participants reading the first part of the novel in English, read 2 852 English sentences and 2 436 Dutch sentences. The participant that only read the first part of the novel in English read 2 852 English sentences. In total we collected eye movements for 59 716 Dutch words (5 575 unique types) and 54 364 English words (5 012 unique types). A summary of the characteristics of the Dutch and English version of the novel is presented in Table 2.

Table 2. *Description of the Dutch and the English version of the novel 'The mysterious case at Styles.' by Agatha Christie.*

	Dutch			English		
Number of words	59 716			54 364		
Number of word types	5 575			5 012		
Number of nouns	7 987			7 639		
Number of noun types	1 777			1 742		
Number of sentences	5 190			5 300		
	M	SD	Range	M	SD	Range
Number of words per sentence	11.64	8.86	[1-60]	10.64	8.20	[1-69]
Word Frequency[a]	4.51	1.39	[0.30-6.24]	4.59	1.37	[0.30-6.33]
Word Length	4.51	2.54	[1-22]	4.18	2.30	[1-17]

[a]Log10 transformed Subtlex frequencies: Subtlex-NL for Dutch words (Keuleers, Brysbaert, et al., 2010), Subtlex-UK for English words (Brysbaert & New, 2009).

APPARATUS

The bilingual eye movement data were recorded with a tower-mounted EyeLink 1000 system (SR-Research, Canada) with a sampling rate of 1 kHz. A chinrest was used to reduce head movements. Monolingual eye movement data were acquired with the same system that was desktop mounted. Reading was always binocular, but eye movements were recorded only from the right eye. Text was presented in black 14 point Courier New font on a light grey background. The lines were triple spaced and 3 characters subtended 1 degree of visual angle or 30 pixels. Text appeared in paragraphs on the screen. A maximum of 145 words was presented on one screen. During the presentation of the novel, the room was dimly illuminated.

PROCEDURE

Each participant read the entire novel in four sessions of an hour and a half. In the first session, every participant read chapter 1 to 4. In the second session chapters 5 to 7, in the third session chapters 8 to 10 and in the fourth session chapter 11 to 13 were read. Every bilingual and monolingual participant completed a number of language proficiency tests. The results of these proficiency measures can be found in Table 1.

The participants were instructed to read the novel silently while the eye tracker recorded their eye movements. It was stressed that they should move their head and body as little as possible while they were reading. The participants were informed that there would be a break after each chapter and that in that pause they would be presented with multiple-choice questions about the contents of the book (Comprehension scores are reported in Table 1). This was done to ensure that participants understood what they were reading and paid attention throughout the session. The number of questions per chapter was relative to the amount of text in that chapter.

The text of the novel appeared on the screen in paragraphs. When the participant finished reading the sentences on one screen, they were able to press the appropriate button on a control pad to move to the next part of the novel.

Before starting the practice trials, a nine-point calibration was executed. The participants were presented with three practice trials where the first part of another story was presented on the screen. After these trials, the participants were asked two multiple-choice questions about the content of the practice story. This part was intended to familiarize participants with the reading of text on a screen and the nature and difficulty of the questions. Before the participant started reading the first chapter another nine-point calibration was done. After this, the calibration was done every 10 minutes, or more frequently when the experiment leader deemed necessary.

RESULTS AND DISCUSSION

We will focus on the distribution and descriptive statistics of five word-level reading time measures extracted from the GECO: a) first fixation duration (FFD), the duration of the first fixation landing on the current word, b) single fixation duration (SFD), the duration of the first and only fixation on the current word, c) gaze duration (GD), the sum of all fixations on the current word in the first pass of reading before the eye moves out of the word, d) total reading time (TRT), the sum of all fixation durations on the current word, including regressions, and e) go past time (GPT), the sum of all fixations prior to progressing to the right of the current word, including regressions to previous words originating from the current word.

Fixations that were shorter than 100ms were excluded from the dataset, because these are not likely to reflect language processing. Words that were skipped were excluded in the rest of the description of the data.

DISTRIBUTION OF READING TIMES

Figure 1 and 2 show boxplots of all reading time measures after log transformation and aggregation over subjects. As we can see, the reading time variables are not normally distributed. Due to the exclusion criteria, they all show a minimal value of 100 ms. They also show a large number of reading time observations that are positive outliers.

To correct for these outliers we removed all reading times that differed more than 2.5 standard deviations from the subject mean per language. The quantile-quantile plots of the log-transformed and trimmed reading times are presented in Figure 3. The Lillifiers normality test statistic (L) is included in all panels. The p-value is smaller than 0.001 in all cases.

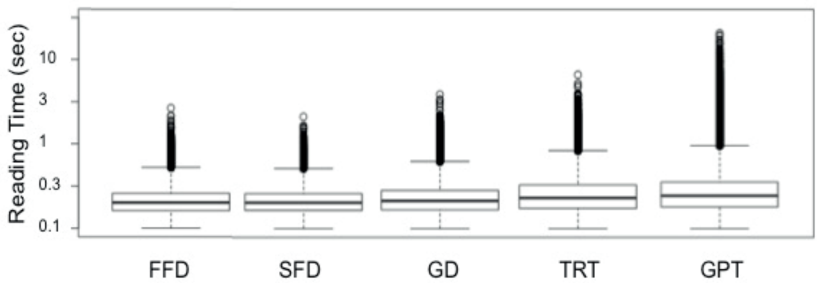


Figure 1. Boxplots of log-transformed reading time data (on the y-axis in seconds) for English monolinguals. Boxes denote the median (thick line), the lower and the upper quartile.

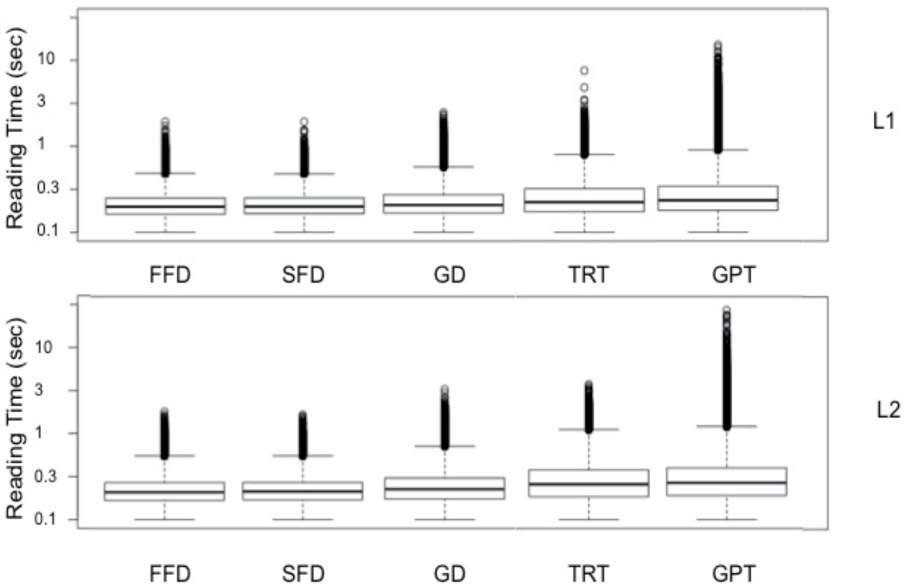
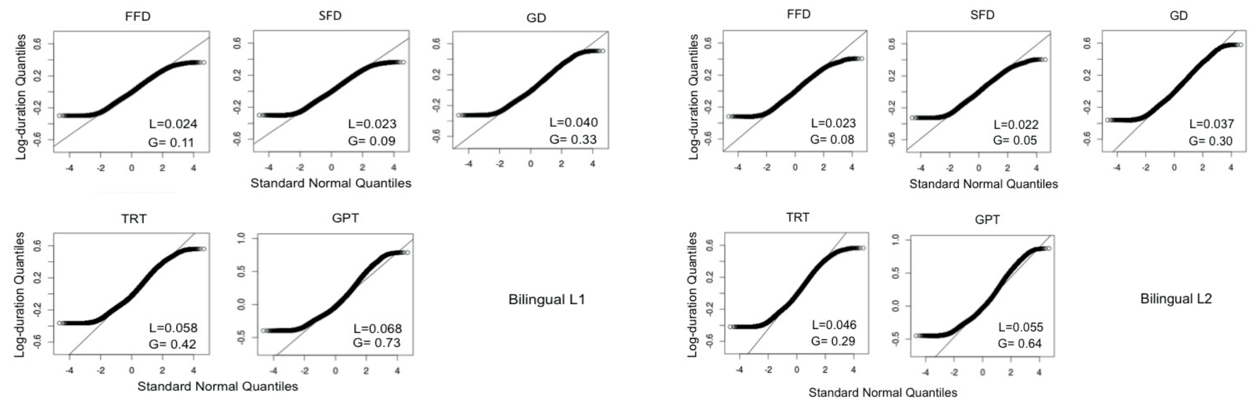


Figure 2. Boxplots of log-transformed reading time data (on the y-axis in seconds) for bilinguals in L1 (upper plot) and L2 (lower plot). Boxes denote the median (thick line), the lower and the upper quartile.

This means that despite of trimming and log-transformation, the reading times were not normally distributed. The Pearson's moment coefficient of skewness (G) is also included in the panels. All G values are positive. This means that the reading times were all positively skewed (skewed to the right). We can see that total reading times and go past times are more skewed than first fixation durations and gaze durations. The variable most approximating normality is the single fixation duration.



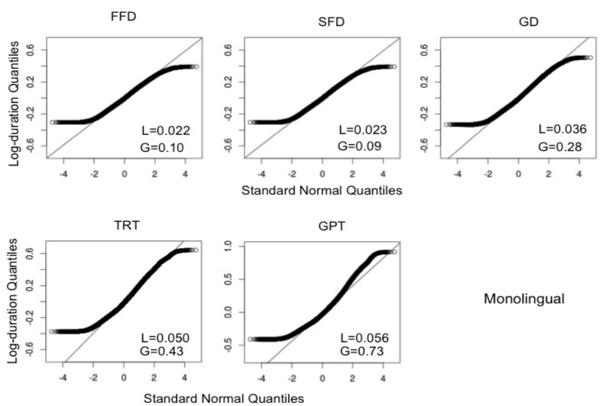


Figure 3. Quantile-quantile plots of standardized log-transformed trimmed reading time durations against a standard normal distribution. Statistic values of the Liliefors test of normality (L) and the Pearson's moment coefficient of skewness (G) are presented on the plots. A larger value for L corresponds to larger deviation from the standard normal distribution. Positive values for G indicate a positive skewness, larger values indicate larger skewness

We refer to Frank et al. (2013) for a similar analysis of the distribution of reading times. Their results also show that despite log-transformation, the reading times gathered by eye-tracking are not normally distributed and are skewed to the right. We must note that although it is the case that our data do not follow a standard normal distribution, this does not pose a problem for analyses because of the large amount of observations.

DESCRIPTION READING TIMES

In Table 3 we present the means of first fixation duration, single fixation duration, gaze duration, total reading time and go past time for monolingual reading and L1 and L2 reading, after trimming. Standard deviations and the range of values are also given. Standard deviations are larger on average for L2 reading. This means that for L2 reading there is more variation in reading times. The larger range in language proficiency for L2 than for L1 might account for this difference in variance. We can see clearly that reading times are longer for L2 reading than for L1 or monolingual reading. We discussed these differences in depth in Cop, Drieghe et al. (2014).

Table 3. *Averages (M), standard deviations (SD) and range of the reading time measures for monolingual, bilingual L1 and bilingual L2 reading.*

	Monolingual (English)			Bilingual L1 (Dutch)			Bilingual L2 (English)		
	M	SD	Range	M	SD	Range	M	SD	Range
First Fixation Duration	214	70	101-502	209	65	101-467	222	74	101-536
Single Fixation Duration	215	69	101-490	210	64	101-464	224	74	101-540
Gaze Duration	232	89	101-695	226	85	101-682	250	105	101-877
Total Reading Time	264	127	101-1060	256	117	101-852	296	194	101-978
Go Past Time	298	187	101-2140	286	168	101-1540	332	218	101-2130

SCOPE OF THE FINDINGS

Of course there are some limitations to the use of a natural eye-tracking corpus. It is much more difficult to control confounding factors in such a setting than in an experimentally controlled design. This deficit can be largely overcome by including possibly confounding factors as covariates in the linear mixed model analyses. However, confounds that are not identified as such and are not included in the model could introduce artifacts in the data. This is why we support an approach of the use of both experimental data and corpus data in conjunction as support for certain theoretical positions. Researchers should focus on both of these approaches in the future development and evaluation of bilingual language models.

CONCLUSION

In this paper, we present the first eye-tracking corpus of natural reading specifically aimed at bilingual application, the GECO, and open it up for free use in future research. Participants were selected on their language history and detailed proficiency measures were gathered. With this corpus, models of bilingual language processing can be evaluated, compared and simulated using one large dataset of bilingual eye movements. This corpus can also be used to test specific hypotheses about differences between L1 and L2 reading or bilingual and monolingual reading. Interesting questions are for example whether bilinguals might use less prediction in reading than monolinguals do or whether specific syntactic constructions are processed differently in L2 than in L1 reading. Another important contribution of these corpora is of a more exploratory nature. The richness in this eye tracking data has potential in inspiring a very wide range of research, yielding new theoretical questions and insights about the time course of reading and specific interactions between multiple levels of a language-user system.

The GECO data is made freely available online for other researchers to analyze and use, provided reference to this paper and corpus is made in resulting writings. The novel that was used is translated in more than 25 languages including Hebrew, Finnish and Japanese. This opens up possibilities for further data collection by other researchers to enable the comparison of natural reading across languages and study bilingualism in different populations and language combinations.

ACKNOWLEDGMENTS

This research was funded by an aspirant scholarship of the Fonds voor Wetenschappelijk Onderzoek Vlaanderen (FWO).

REFERENCES

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition*, *24*(4), 477–492.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., … Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, *39*(3), 445–459.

Bamgbose, A. (1994). Nigeria's Choice. *UNESCO Courier*, pp. 22–31.

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

Boston, M. F., Hale, J., Kliegl, R., Patil, U., & Vasishth, S. (2008). Parsing costs as predictors of reading difficulty: An evaluation using the Potsdam Sentence Corpus. *Journal of Eye Movement Research*, *2*(1), 1.

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. http://doi.org/10.3758/BRM.41.4.977

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256. http://doi.org/10.1037//0033-295X.108.1.204

Cop, U., Dirix, N., Drieghe, D., & Duyck, W. (2015). Cross-lingual Neighborhood Effects in Generalized Leical Decison and Natural Reading. *Manuscript Submitted for Publication*.

Cop, U., Drieghe, D., & Duyck, W. (2014). Eye Movement Patterns in Natural Reading: A Comparison of Monolingual and Bilingual Reading of a Novel. *Manuscript Submitted for Publication*.

Cop, U., Keuleers, E., Drieghe, D., & Duyck, W. (2015). Frequency Effects in Monolingual and Bilingual Natural Reading. *Psychonomic Bulletin and Review*. http://doi.org/10.3758/s13423-015-0819-2 i

Cop, U., Van Assche, E., Drieghe, D., & Duyck, W. (2014). Bilinguals reading a novel: An eye movement study of cognate facilitation in L1 and L2 natural reading. *Manuscript Submitted for Publication*.

Cover, T. M., & Thomas, J. A. (1991). *Elements of information theory*. New York: Wiley.

Demberg, V., & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, *109*(2), 193–210. http://doi.org/10.1016/j.cognition.2008.07.008

De Pessemier, P., & Andries, C. (2009). *GL&SCHR*. Leuven/Apeldoorn: Garant.

Diependaele, K., Brysbaert, M., & Neri, P. (2012). How Noisy is Lexical Decision? *Frontiers in Psychology*, *3*. http://doi.org/10.3389/fpsyg.2012.00348

Dijkstra, T., Timmermans, M., & Schriefers, H. (2000). On being blinded by your other language: effects of task demands on inter-lingual homograph recognition. *Journal of Memory and Language*, *42*, 445–464.

Dijkstra, T., & Van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. Jacobs (Eds.), *Localist Connectionist Approaches to Human Cognition* (pp. 189–225). New York: Mahwah.

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(03). http://doi.org/10.1017/S1366728902003012

Dilkina, K., McClelland, J. L., & Plaut, D. C. (2010). Are there mental lexicons? The role of semantics in lexical decision. *Brain Research*, *1365*, 66–81. http://doi.org/10.1016/j.brainres.2010.09.057

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A Dynamical Model of Saccade Generation During Reading. *Psychological Review*, *112*(4), 777–813. http://doi.org/10.1037/0033-295X.112.4.777

European Union, & European Comission for Education and Culture. (2012). Europeans and their Languages. Retrieved from http://ec.europa.eu/public_opinion/archives

Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., … Pallier, C. (2010). The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods*, *42*(2), 488–496. http://doi.org/10.3758/BRM.42.2.488

Fossum, V., & Levy, R. (2012). Sequential vs. hierarchical syntactic models of human incremental sentence processing. In *Proceedings of the 3rd Workshop on Cognitive Modeling and Computational Linguistics* (pp. 61–69). Association for Computational Linguistics. Retrieved from http://dl.acm.org/citation.cfm?id=2390313

Frank, S. L., & Bod, R. (2011). Insensitivity of the Human Sentence-Processing System to Hierarchical Structure. *Psychological Science*, *22*(6), 829–834. http://doi.org/10.1177/0956797611409589

Friederici, A. D. (1995). The Time Course of Syntactic Activation During Language Processing: A Model Based on Neuropsychological and Neurophysiological Data. *Brain and Language*, *50*(3), 259–281. http://doi.org/10.1006/brln.1995.1048

Grainger, J., & Jacobs, A. M. (1996). Orthographic Processing in Visual Word Recognition: A Multiple Read-out Model. *Psychological Review*, *103*(3), 518–565. http://doi.org/10.1037/0033-295X.103.3.518

Grosjean, F. (1992). Another view of bilingualism. In *Cognitive processing in bilinguals* (pp. 51–62). Oxford, England: North-Holland.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, *111*(3), 662–720. http://doi.org/10.1037/0033-295X.111.3.662

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics*, *40*(6), 431–439. http://doi.org/10.3758/BF03208203

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research*, *45*(2), 153–168. http://doi.org/10.1016/j.visres.2004.07.037

Kennedy, A., & Pynte, J. (2008). The consequences of violations to reading order: An eye movement analysis. *Vision Research*, *48*(21), 2309–2320. http://doi.org/10.1016/j.visres.2008.07.007

Kennedy, A., Pynte, J., Murray, W. S., & Paul, S.-A. (2013). Frequency and predictability effects in the Dundee Corpus: An eye movement analysis. *The Quarterly Journal of Experimental Psychology*, *66*(3), 601–618. http://doi.org/10.1080/17470218.2012.676054

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, *42*(3), 643–650. http://doi.org/10.3758/BRM.42.3.643

Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice Effects in Large-Scale Visual Word Recognition Studies: A Lexical Decision Study on 14,000 Dutch Mono- and Disyllabic Words and Nonwords. *Frontiers in Psychology*, *1*. http://doi.org/10.3389/fpsyg.2010.00174

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods*, *44*(1), 287–304. http://doi.org/10.3758/s13428-011-0118-4

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the Mind During Reading: The Influence of Past, Present, and Future Words on Fixation Durations. *Journal of Experimental Psychology: General*, *135*(1), 12–35. http://doi.org/10.1037/0096-3445.135.1.12

Kroll, J. F., Dussias, P. E., Bogulski, C. A., & Valdes Kroff, J. (2012). Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 56, pp. 229–262).

Kuperman, V., Dambacher, M., Nuthmann, A., & Kliegl, R. (2010). The effect of word position on eye-movements in sentence and paragraph reading. *The Quarterly Journal of Experimental Psychology*, *63*(9), 1838–1857. http://doi.org/10.1080/17470211003602412

Kuperman, V., Drieghe, D., Keuleers, E., & Brysbaert, M. (2013). How strongly do word reading times and lexical decision times correlate? Combining data from eye movement corpora and megastudies. *The Quarterly Journal of Experimental Psychology*, *66*(3), 563–580. http://doi.org/10.1080/17470218.2012.658820

Kuperman, V., & Van Dyke, J. A. (2013). Reassessing word frequency as a determinant of word recognition for skilled and unskilled readers. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 802–823. http://doi.org/10.1037/a0030859

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325–343. http://doi.org/10.3758/s13428-011-0146-0

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967.

Mitchell, J., Lapata, M., Demberg, V., & Keller, F. (2010). Syntactic and semantic factors in processing difficulty: An integrated measure. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics* (pp. 196–206). Association for Computational Linguistics. Retrieved from http://dl.acm.org/citation.cfm?id=1858702

Norris, D., & Kinoshita, S. (2012). Orthographic Processing is Universal; It's What You Do with It That's Different. *Behavioral and Brain Sciences*, *1*(1), 34–35.

Pynte, J., & Kennedy, A. (2006). An influence over eye movements in reading exerted from beyond the level of the word: Evidence from reading English and French. *Vision Research*, *46*(22), 3786–3801. http://doi.org/10.1016/j.visres.2006.07.004

Pynte, J., & Kennedy, A. (2007). The influence of punctuation and word class on distributed processing in normal reading. *Vision Research*, *47*(9), 1215–1227. http://doi.org/10.1016/j.visres.2006.12.006

Pynte, J., New, B., & Kennedy, A. (2008). On-line contextual influences during reading normal text: A multiple-regression analysis. *Vision Research*, *48*(21), 2172–2183. http://doi.org/10.1016/j.visres.2008.02.004

Pynte, J., New, B., & Kennedy, A. (2009a). On-line contextual influences during reading normal text: The role of nouns, verbs and adjectives. *Vision Research*, *49*(5), 544–552. http://doi.org/10.1016/j.visres.2008.12.016

Pynte, J., New, B., & Kennedy, A. (2009b). On-line syntactic and semantic influences in reading revisited. *Journal of Eye Movement Research*, *3*(1), 12.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506. http://doi.org/10.1080/17470210902816461

Rayner, K., & Fischer, M. H. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception & Psychophysics*, *58*(5), 734–747. http://doi.org/10.3758/BF03213106

Rayner, K., Li, X., & Pollatsek, A. (2007). Extending the E-Z Reader Model of Eye Movement Control to Chinese Readers. *Cognitive Science*, *31*(6), 1021–1033. http://doi.org/10.1080/03640210701703824

Rayner, K., & Pollatsek, A. (1989). *The Psychology of Reading*. Hillsdale, NJ, England: Erlbaum.

Rayner, K., Pollatsek, A., Liversedge, S. P., & Reichle, E. D. (2009). Eye movements and non-canonical reading: Comments on. *Vision Research*, *49*(17), 2232–2236. http://doi.org/10.1016/j.visres.2008.10.013

Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and Aging*, *21*(3), 448–465. http://doi.org/10.1037/0882-7974.21.3.448

Reichle, E. D., Liversedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2013). Using E-Z Reader to examine the concurrent development of eye-movement control and reading skill. *Developmental Review*, *33*(2), 110–149. http://doi.org/10.1016/j.dr.2013.03.001

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*(1), 125.

Seidenberg, M., & Waters, G. (1989). Reading Words Aloud - a Mega Study. *Bulletin of the Psychonomic Society*, *27*(6), 489–489.

Shin, H. B., & Kominski, R. A. (2007). Language Use in the United States: 2007. U.S. Census Bureau, Washinton DC.

Van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of*

Memory and Language, *39*(3), 458–483. http://doi.org/10.1006/jmla.1998.2584

Whitney, C. (2011). Location, location, location: How it affects the neighborhood (effect). *Brain and Language*, *118*(3), 90–104. http://doi.org/10.1016/j.bandl.2011.03.001

Wilkinson, G. S., & Robertson, G. . (2006). *Wide Range Achievement Test 4 professional manual.* Lutz, FL: Psychological Assessment Resources.

Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language*, *60*(4), 502–529. http://doi.org/10.1016/j.jml.2009.02.001

APPENDIX A: FILE DESCRIPTIONS

Table A.1. *Description of the file 'SubjectInformation.xlsx'. Column names are in the first column and a description of the content in that column is presented in the second column.*

Column Name	Description
PP_NR	The identification number of the participant.
GROUP	Factor indicating whether the participants belonged to the unbalanced bilingual ("bilingual") or monolingual group ("monolingual")
AGE	Age of the participant in years
SEX	Sex of the participant ("f"=female, "m"=male)
AOA_ENG	Age of Acquisition of the English language or zero when monolingual
%EXP_DUTCH	Percentage of daily language exposure to Dutch
%EXP_ENG	Percentage of daily language exposure to English
LEXTALE_DUTCH	Score on the Dutch LexTALE (Lexical Test for Advanced learners of English; Lemhöfer & Broersma, 2012), NA for monolinguals
LEXTALE_ENG	Score on the English LexTALE (Lexical Test for Advanced learners of English; Lemhöfer & Broersma, 2012)
SPELLING_DUTCH	Percentage score on the Dutch spelling test (GL&SCHR; De Pessemier & Andries ,2009)
SPELLING_ENG	Percentage score on the English spelling test (WRAT4; Dell, Harrold, & Dell, 2008)
COMPR_DUTCH	Percentage score on the multiple-choice questions for the Dutch chapters of the novel
COMPR_ENG	Percentage score on the multiple-choice questions for the English chapters of the novel
LEX_DEC_ACC_DUTCH	Percentage score of accuracy on the Dutch lexical decision task on the word trails, corrected for false positives.
LEX_DEC_ACC_ENG	Percentage score of accuracy on the English lexical decision task on the word trails, corrected for false positives.

Table A.2. *Description of the files 'EnglishMaterials.xlsx' and 'DutchMaterials.xlsx'. Column and sheet names are in the first column and a description of the content in that column or sheet is presented in the second column.*

Sheet Name	Description
ALL	Each word associated with an interest area presented on a separate line.
NOUNS	Each noun of the novel presented on a separate line.
SENTENCE	Each sentence of the novel presented on a separate line.
Column Name	**Description**
IA_ID	Identification number of the interest area. The first number refers to the part of the novel (1,2,3 or 4), the second number refers to the trail number, and the last number refers to the interest area number within the trial.
SENTENCE_ID	Identification number of the sentence. The first number refers to the part of the novel (1,2,3 or 4), the second number refers to the sentence number within the part.
CHRON_ID	Chronological identification number of the current interest area.
WORD	The word contained in the current interest area.
PART_OF_SPEECH	The syntactic function of the current word in the sentence context.
CONTENT_WORD	Factor denoting whether the current word is a content word ("1") or a function word ("0").
SUBTLEX_FREQ_COUNT	The frequency of the current word in the subtlex data base (SUBTLEX-US: Brysbaert & New, 2009)SUBTLEX-NL : Keuleers, Brysbaert, et al., 2010).
LOG_SUBTLEX_FREQ_COUNT	The log transformed value with base 10 of the frequency of the current word in the subtlex data base (SUBTLEX-US: Brysbaert & New, 2009), SUBTLEX-NL: Keuleers, Brysbaert, & New, 2010).

WORD_LENGTH	The number of characters of the current word.
IA_AREA	The size of the current interest area in pixels.
IA_TOP	The top side pixel position of the current interest area.
IA_BOTTOM	The bottom side pixel position of the current interest area.
IA_LEFT	The left side pixel position of the current interest area.
IA_RIGHT	The right side pixel position of the current interest area.
IDENTICAL_COGNATE	Factor denoting whether the current word has an identical cognate in the other language ("1') or not ("0").
CORR_LEVENSHTEIN	The corrected levenshtein distance between the current word and its translation equivalent in the other language.
NEIGHBOR_DENSITY_ENGLISH	The sum of the number of English transposition, addition and deletion neighbors of the current word taken from CLEARPOND (Marian, Bartolotti, Chabal, & Shook, 2012).
NEIGHBOR_DENSITY_DUTCH	The sum of the number of Dutch transposition, addition and deletion neighbors of the current word taken from CLEARPOND (Marian, Bartolotti, Chabal, & Shook, 2012).
SUM_BIGRAM_FREQ	The sum of the frequency of the bigrams of the current word calculated with the WordGen tool (Duyck, Desmet, Verbeke, & Brysbaert, 2004).
SENTENCE	The sentence referred to with the current sentence-ID.
NUMBER_WORDS_SENTENCE	The number of words in the current sentence.

APPENDIX B: RESULTS SELF-REPORT QUESTIONNAIRE

Table B.1. *Count of bilingual participants agreeing and not agreeing on second language skills items.*

Skills	Agree	Don't Agree
Carry on normal conversation in L2	19	0
Watch television shows in L2	19	0
Listen to music in L2	19	0
Read and comprehend questions in L2	19	0
Read books or articles in L2	19	0
No problems in understanding L1 speaker	18	1
Carry on a discussion in L2	17	2
Love speaking L2	16	3
Explain difficult situation in L2	15	4
Answer difficult questions in L2	12	7
Think in L2	11	8
Speak to myself in L2	10	9
Write in L2	8	11
Make no/ almost no mistakes in L2	6	13
Dream in L2	5	14

Table B.2. *Count of bilingual participants agreeing and not agreeing on second language switching items.*

Switching	Agree	Don't Agree
I'm sometimes in a tip of the tongue state	16	3
I sometimes can't get the right word	14	5
I use a different language when I do not remember a word	13	6
I often use different languages intermixed	9	10
I often use different languages intermixed without noticing	5	14
I sometimes speak in a language that my dialogue partner doesn't understand	5	14

CHAPTER 3
A GLOBAL COMPARISON OF MONOLINGUAL AND BILINGUAL READING OF A NOVEL[4][5]

This paper presents a corpus of sentence level eye movement parameters for unbalanced bilingual first language (L1) and second-language (L2) reading and monolingual reading of a complete novel (56 000 words).

We present important sentence-level basic eye movement parameters of both bilingual and monolingual natural reading extracted from this large data corpus.

Bilingual L2 reading patterns show longer sentence reading times (20%), more fixations (21%), shorter saccades (12%) and less word skipping (4.6%), than L1 reading patterns. Regression rates are the same for L1 and L2 reading. These results could indicate, analogous to a previous simulation with the E-Z reader model in the literature, that it is primarily the speeding up of lexical access that drives both L1 and L2 reading development.

Bilingual L1 reading does not differ in any major way from monolingual reading. This contrasts with predictions made by the *weaker links account*, which predicts a bilingual disadvantage in language processing caused by divided exposure between languages.

[4] This chapter is based on a manuscript co-authored by Denis Drieghe and Wouter Duyck: Cop, U., Drieghe, D., & Duyck, W. (2014) Eye Movement Patterns in Natural Reading: A Comparison of Monolingual and Bilingual Reading of a Novel, submitted for publication.

[5] Revised manuscript submitted for publication in PLOS-ONE

INTRODUCTION

By now, psycholinguistics has gained a good understanding of monolingual reading behavior. However, because of the increased globalization of our multicultural society, more and more people acquire, apart from their mother tongue (L1), one or more other languages (L2, L3…). It is now estimated that about half of the world's population has some knowledge of more than one language, and can therefore considered to be bilingual, following the common Grosjean definition: "*bilinguals are those people who need and use two (or more) languages in their everyday lives*" (Grosjean, 1982). In contrast, current models of eye movements during reading still focus exclusively on monolingual reading, so that we do not know in what way L2 sentence reading differs from L1 reading, or whether merely being a bilingual changes L1 reading.

In contrast to the monolingual domain, almost all studies of bilingual reading have focused on the word level. The few studies that do use sentence materials suggest that having a second language available influences the way the first language is processed (Titone, Libben, Mercier, Whitford, & Pivneva, 2011; Van Assche, Duyck, Hartsuiker, & Diependaele, 2009). They do not however consider sentence-level reading parameters, as was done in the monolingual domain (Blythe, Liversedge, Joseph, White, & Rayner, 2009; Joseph, Liversedge, Blythe, White, & Rayner, 2009; Rayner, 1986), but rather focus on the recognition of target words that are embedded in a sentence context (Altarriba, Kroll, Sholl, & Rayner, 1996; Bultena, Dijkstra, & van Hell, 2014; Dussias & Cramer Scaltz, 2008; Dussias & Sagarra, 2007; Duyck, Van Assche, Drieghe, & Hartsuiker, 2007; Gollan et al., 2011; Libben & Titone, 2009; Titone et al., 2011; Van Assche, Drieghe, Duyck, Welvaert, & Hartsuiker, 2011; Van Assche et al., 2009; Winskel, Radach, & Luksaneeyanawin, 2009). The present study aims to address this gap by providing a systematic investigation of eye movements when bilinguals read in their native and second language. These data constitute the necessary constraints to generalize models of eye movement behavior to bilingual readers.

MONOLINGUAL EYE MOVEMENTS WHILE READING

When we read, our eyes move from one position to the next in order to identify and process visual word form information. This entails rapid jerk-like movements (saccades) and short periods of steadiness (fixations). Saccades are necessary to direct the gaze to a new location, bringing new information into the center of the visual field where acuity is best. During these saccades, no meaningful new visual information is gathered. They occur several times per second and typically move the eyes forward about 7-9 character spaces (for reviews: Rayner, 1998, 2009). Psycholinguists assume that eye movements during reading reflect language processing (Liversedge & Findlay, 2000), with fixation durations as a marker of the ease of accessing the meaning of a word and integrating this into the current sentence. Because of the spatially accuracy and high temporal resolution of eye tracking, it allows us to dissociate early from late eye movement measures. In combination with other information, such as word length/frequency, this makes it possible to investigate the time course of the reading process. Additionally, reading processes in eye tracking are not confounded by task-related processes or strategies that other lab tasks (e.g. lexical decision or naming) entail. Hence, this method is considered to be the closest experimental parallel to the natural reading process.

During the last three decades, the development of monolingual theories on visual language comprehension has been heavily influenced by eye tracking research in reading. Rayner's influential review article (Rayner, 1998), now 15 years old, already discusses more than 550 articles investigating this topic (for a more recent review: Rayner, 2009). Also, several corpus studies of eye movements were undertaken, and these data were used to provide an account of (monolingual) reading. The Potsdam Corpus (Kliegl & Engbert, 2005; Kliegl, Grabner, Rolfs, & Engbert, 2004) contained eye movements of 222 subjects reading 144 constructed German sentences (1 138 words). The Dundee corpus (Kennedy & Pynte, 2005), an English and French study in which 10 participants read 50 000 words in paragraphs, was used to investigate effects of parafoveal processing. Clearly, these corpora of eye movements provide a very rich and extended source of information about the mechanisms that underlie language processing in a more natural context and could serve as harvesting grounds for the development of comprehensive language models. For example, the Amherst Sentence Corpus (Reichle, Pollatsek, Fisher, & Rayner, 1998) was used to develop the first version of the SWIFT model of saccade generation (Engbert, Nuthmann, Richter, & Kliegl, 2005).

The E-Z reader model (Pollatsek, Reichle, & Rayner, 2006; Rayner, Ashby, Pollatsek, & Reichle, 2004; Reichle et al., 1998; Reichle, Rayner, & Pollatsek, 1999, 2003; Reichle, Tokowicz, Liu, & Perfetti, 2011; Reichle, Warren, & McConnell, 2009) is the most cited model of monolingual eye movements. It is implicitly limited to native language or even monolingual reading behavior, and it is yet unknown how these mechanisms operate when bilinguals read in a second language, or how knowledge of a second language influences native language reading. However, it is interesting that the original E-Z reader model has been successfully accommodated to account for other reading patterns, such as those of older readers (Rayner, Reichle, Stroud, Williams, & Pollatsek, 2006), children (Reichle et al., 2013), or of non-alphabetic languages (Rayner, Li, & Pollatsek, 2007). This illustrates that this model could be useful and relevant in future modeling efforts concerning bilingual eye movement patterns, and we will therefore align our analyses of bilingual reading behavior with the core assumptions and variables of this model.

The E-Z reader model assumes serial lexical processing. The completion of an early stage of lexical processing on word *n,* called the *familiarity check,* is the ‘trigger’ that causes the oculo-motor system to begin the programming of a saccade directed towards the next word *n+1.* The subsequent completion of a second stage of lexical processing on word *n,* called the *completion of lexical access,* causes attention to shift from word *n* to word *n+1.* Thus, the programming of saccades is decoupled from the shifting of attention, which is allocated serially to only a single word at a time (Reichle et al., 1998). Because attention shifts are faster than the programming of a saccade (Rayner et al., 2006), the lexical processing of word *n+1* usually begins when the eyes are still fixated on word *n.* This feature of the model allows parafoveal processing of upcoming words. Following similar reasoning, the model predicts that parafoveal words, which are processed fast enough, might be skipped.

The model assumes that word length and frequency are important lexical variables that have a large effect on the eye movements, because these variables define the duration of the familiarity check (Brysbaert & Vitu, 1998; Nation, 2009). Consequently, they determine fixation duration, fixation count, rightward saccade length, skipping and regression rates. These will also be the core variables that will be assessed in the present paper.

RESEARCH ON BILINGUALISM

Most bilingual language research has focused on the question of how the bilingual lexicon is organized. Do people have separate representational systems for lexical items of different languages or is there one integrated lexicon? Although intuitively the most straightforward option might be to have a separate lexicon for each language, and although bilinguals can use one of their languages without the constant intrusion of the other language (Poulisse & Bongaerts, 1994), the large majority of experimental evidence shows that bilinguals have one integrated lexicon containing representations of all words belonging to both languages and that this lexicon is accessed language independently (Dijkstra & van Heuven, 2002). Evidence for this idea is mainly provided by research on cross-lingual interactions, in which it is typically shown that words with some overlap across languages are processed differently than control words, even during unilingual processing. Most often these overlapping words are cognates presented in isolation (Brenders, van Hell, & Dijkstra, 2011; Bultena, Dijkstra, & van Hell, 2013; Caramazza & Brones, 1979; Costa, Caramazza, & Sebastian-Galles, 2000; Cristoffanini, Kirsner, & Milech, 1986; Davis et al., 2010; Dijkstra, Grainger, & van Heuven, 1999; Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010; Gollan, Forster, & Frost, 1997; Kroll, Dijkstra, Janssens, & Schriefers, 1999; Lemhöfer et al., 2008; Peeters, Dijkstra, & Grainger, 2013; Sanchez-Casas, Davis, & García-Albea, 1992; Van Hell & Dijkstra, 2002).

Cognates are words that are translation equivalents but also show some degree of form overlap (e.g. Dutch-English *appel; apple*). Research shows that bilinguals identify cognates faster than control words in a lexical decision task (e.g., Dijkstra et al., 1999; Van Hell & Dijkstra, 2002), a translation priming task (e.g., Davis et al., 2010; Gollan et al., 1997) and a progressive demasking task (e.g., Lemhöfer et al., 2008). This is the case when participants perform the task in their L2 (e.g. Dijkstra et al., 1999; Lemhöfer et al., 2008) and in their L1 (e.g., Davis et al., 2010), although the effect is usually larger for L2 (Kroll et al., 1999). These cross-lingual interaction effects are also found when a target word is embedded in a sentence context (Gullifer, Kroll, & Dussias, 2013; Schwartz & Kroll, 2006; van Hell & de Groot, 2008). This means that a unilingual sentence context does not restrict lexical access to only the target language. In this way these studies provide evidence for a language non-selective view on bilingual language processing. For an overview of evidence for cross-lingual activation and an integrated bilingual lexicon see Brysbaert and Duyck's (2010) or van Hell and Tanner's (2012) overview.

All of the bilingual research discussed in the previous paragraphs used an alternative method to eye movement recording, such as word naming, categorization tasks or lexical decision tasks to examine lexical processing. Although these tasks have their merits for investigating word recognition in isolation, there also have some limitations, besides those mentioned in the previous section, that make these methods suboptimal for investigating lexical access in natural reading. In natural reading, word processing is influenced by the sentence context and parafoveal stimuli (McConkie & Rayner, 1975). This suggests that words are processed gradually across time and across multiple fixations. Also, during reading of text lexical access takes place while other cognitive processing is going on. Kuperman, Drieghe, Keuleers, and Brysbaert (2013) indeed show that only 5-17% of the variance in gaze durations on target words embedded in sentences is explained by lexical decision times in isolation after partialling out the effects of word frequency and word length. This illustrates that the two approaches are indeed distinguishable and measure, to a large extent, different language processes, making both approaches indispensable to research into language processes. Given that only eye tracking assesses reading behavior as it occurs in natural language processing, it is important not to rely solely on artificial word processing paradigms such as lexical decision tasks for the development of models of reading but to complement them with natural reading tasks.

As mentioned above, monolingual theories on visual word recognition have advanced much through eye tracking studies. In the bilingual domain, most eye tracking studies examined eye movements to detect cross-lingual activation in bilingual reading (Balling, 2013; Bultena et al., 2014; Duyck et al., 2007; Libben & Titone, 2009; Pivneva, Mercier, & Titone, 2014; Titone et al., 2011; Van Assche et al., 2011, 2009). Other eye-tracking studies have focused on syntactic processing (Dussias & Cramer Scaltz, 2008; Dussias & Sagarra, 2007), the effect of semantic constraint (Altarriba et al., 1996), frequency effects (Gollan et al., 2011; Whitford & Titone, 2012) or inter-word spacing effects (Winskel et al., 2009) in bilingual visual word recognition.

Most studies that tracked eye movements in bilinguals examined the fixations directed towards the embedded target words, or some other critical target area, without taking into account changes in global eye movement behavior that L2 reading might entail (Altarriba et al., 1996; Bultena et al., 2014; Dussias & Cramer Scaltz, 2008; Dussias & Sagarra, 2007; Duyck et al., 2007; Gollan et al., 2011; Libben & Titone, 2009; Titone et al., 2011; Van Assche et al., 2011, 2009). Although Titone et al. (2011) and Altarriba et al. (1996) do provide some basic word-level eye movement measures for paragraph reading as a measure of reading proficiency, Whitford and Titone (2012) were the first to analyze bilingual eye movements to all words, not just target words, in bilingual paragraph reading. These data are still presented on a word level. To our knowledge there is only one bilingual eye tracking study by Winskel et al. (2009), that provides sentence level reading measures for bilingual sentence reading. They give the sentence reading time and fixation count for 36 English-Thai bilinguals reading 72 Thai and English spaced and un-spaced sentences. See Van Assche, Duyck, and

Hartsuiker (2012) and Dussias (2010) for an overview of the use of eye movements in bilingual sentence processing research.

THEORIES ABOUT BILINGUAL WORD RECOGNITION

The most cited, and the only implemented, model of bilingual visual word recognition is the Bilingual Interactive Activation plus (BIA+) model (Dijkstra & van Heuven, 2002). This model is an adaptation of the interactive activation model of word recognition (McClelland & Rumelhart, 1981). The main differences are the inclusion of lexical representations of two languages, and a distinction between a word identification system and a task/decision system. The BIA+ states that during bilingual reading there is parallel, language independent activation of lexical representations in an integrated lexicon. Language nodes that represent language membership are included in the model, but they cannot tune word recognition towards a single language via top-down activation. This architecture implies that for every word bilingual readers encounter all lexical candidates from all known languages are activated to some extent. Evidence for this model is generated by studies supporting cross-lingual interactions (see previous paragraph for references).

A limitation of the BIA+ model (Dijkstra & van Heuven, 2002), similar to the monolingual interactive activation model (McClelland & Rumelhart, 1981) is that it is tailored to isolated word recognition, and not to sentence reading. The authors do assume effects of sentence context and non-linguistic information on word recognition but the exact nature of these interactions are not specified. This means that a model of bilingual eye movements, such as the E-Z reader model, is not yet available, as there is also no sentence reading data to base it upon.

The *weaker links account* (Gollan et al., 2011; Gollan, Montoya, Cera, & Sandoval, 2008), sparked by small but consistent production disadvantages exhibited by bilinguals compared to monolinguals (Gollan, Montoya, Fennema-Notestine, & Morris, 2005; Gollan & Silverberg, 2001; Gollan & Acenas, 2004; Gollan et al., 2011; Ivanova & Costa, 2008), has recently gained popularity in the literature. Like the BIA+ model (Dijkstra & van Heuven, 2002), it assumes an integrated bilingual lexicon. According to this *frequency-lag* account, bilinguals will have about double the amount of lexical items in their lexicon as monolinguals and will necessarily divide the frequency of use of these words between languages (Gollan et al., 2008). Considering the lexical quality hypothesis (Perfetti, 1992; Perfetti, Wlotko, & Hart, 2005), which states that increased word practice results in better precision of the corresponding lexical representations, it is plausible that bilingual representations will be of lower precision than those of monolinguals. Indeed, Gollan et al. (2008) predict that weaker links between word form and representations for bilinguals should result in slower lexical access during language comprehension, either while accessing L1 or L2, compared to monolinguals. Effects might be smaller than in production because the processes needed for language production are less practiced, more difficult and involve more levels of processing for which frequency is important (Gollan et al., 2011). In the comprehension domain, it was indeed found that bilinguals show slower L1 lexical decision times than monolinguals do (Lehtonen et al., 2012; Ransdell & Fischler, 1987).

A core assumption at the heart of the weaker links account is that total language exposure is equal for all people. While this maybe the case for bilinguals who are exposed to two languages from birth, it is definitely not true for all groups of bilinguals. The authors that constructed the weaker links account used mostly early Spanish English bilinguals (Gollan et al., 2011, 2008; Gollan & Acenas, 2004). A population of unbalanced bilinguals usually acquires a 2nd language in a classroom context, thus increasing their total vocabulary and language exposure, not per se decreasing their L1 exposure. On top of that, the words of their mother tongue will have been fully lexically entrenched before they start learning their second language. This means that for late learners of an L2, the lexical entrenchment for L1 words might be equally strong as the lexical entrenchment for the words of a monolingual.

THIS STUDY

The current paper provides the first comprehensive description of bilingual (L1 and L2) and monolingual reading on a sentence level by gathering a corpus of eye movement data while participants read an entire novel. Within this single data set a wide range of phenomena can be studied in an ecologically valid context and benchmark parameters of bilingual L1 and L2 natural sentence reading can be extracted. This corpus enables the examination of global changes in eye movement pattern, clarifying localized measures associated with the identification of specific words embedded within a sentence. To be more specific, if our analysis for instance shows that average saccade length is typically reduced in L2 reading compared to L1 reading, this would influence factors that are normally associated with the lexical processing of a specific word (e.g. word skipping, number of fixations) even though these patterns would only reflect global adjustments to reading in L2 and not just the lexical processing of the currently fixated word. Ultimately, these results will promote the development of models and theories on bilingual language processing in L1 and L2.

The aim of this paper is twofold. First, we will compare eye movement patterns of bilinguals reading in L1 and L2 . We will use a within-subjects design. In this way, reading language is not confounded with inter-individual differences such as motivation or intelligence. A direct comparison of individuals' reading performance across languages is rather challenging. We discuss this issue in the section 'Analytic Techniques for Cross-Language Comparison'. Second, we want to investigate whether merely being a bilingual changes native language reading, by comparing bilingual L1 (Dutch) with monolingual L1 (English) reading of cross-lingually matched sentences (between-subjects).

PREDICTIONS L1 VS. L2 READING

As discussed, the weaker links account predicts a disadvantage for the least frequently used language dependent on the relative exposure of L1 and L2, caused by weaker links between L2 word forms and representations (Gollan et al., 2008). Although some of the studies described above, for example

Whitford and Titone (2012), observed longer gaze durations and longer sentence reading times on embedded target words in L2 sentences, no study so far has compared basic sentence parameters for L1 and L2 reading.

We can draw a parallel between the sentence reading pattern of children and the expected sentence reading pattern for unbalanced bilinguals reading in L2. Unbalanced bilinguals are also developing, although for the second time, reading skills. For bilinguals, the first stages of letter recognition should already have been automatized, so on a quantitative level, we expect that the size of the difference between L1 and L2 bilingual reading measures should be somewhat smaller than the size of the difference between adults' and children's reading measures.

As children acquire reading skills and gain language proficiency, sentence reading times and fixation durations get shorter, saccade length gets longer, and fewer fixations, regressions and refixations are made (Blythe, Häikiö, Bertam, Liversedge, & Hyönä, 2011; Blythe et al., 2009; Blythe & Joseph, 2011; Buswell, 1922; Häikiö, Bertram, Hyönä, & Niemi, 2009; Huestegge, Radach, Corbic, & Huestegge, 2009; Joseph et al., 2009; Rayner, 1986; Taylor, 1965). Interestingly, these are strictly quantitative, rather than qualitative differences. This robust evolution is most likely due to a speeding of the lexical identification of the individual words (Blythe et al., 2011) not by oculomotor development (Blythe et al., 2009; Huestegge et al., 2009; Rayner, 1986) So, although children are slower, they do not need more time than adults do to take up the necessary information from the page. Reichle et al. (2013) confirmed this using a simulation of the eye movement data of children using the E-Z reader model (Reichle et al., 1998). The full eye movement pattern of children was simulated by lowering the default rate of lexical processing compared to adults. This supports the fact that the tuning of the oculomotor system is not the main element that drives the development of eye movement behavior in children (Reichle et al., 2013).

Rayner et al. (2006) described a "risky reading strategy" for older readers as a compensation mechanism for slower lexical access. Older people fixate longer on individual words in a sentence and make more regressions in the text, but also that they skip more words and move their eyes with bigger saccades over the text.

In summary, given lower language proficiency for L2, we predict a “child-like” eye movement pattern for bilinguals reading in their L2 vs. their L1. This is compatible with the weaker links hypothesis, which also assumes effects of lower L2 practice. This disadvantage should be more pronounced in readers who score lower on L2 proficiency. We predict more and longer fixations per sentence, a smaller rightward saccade length, a lower skipping rate and a higher regression rate for L2, but we keep in mind that this pattern might be compensated by strategically adjusting the skipping rates and saccade length, as Rayner et al. (2006) observed for older readers.

PREDICTIONS MONOLINGUAL VS. BILINGUAL READING

For bilinguals, reading experience is supposedly spread across two different languages, L1 and L2 (Gollan et al., 2008). This implies lower absolute exposure to each language, which could result in slower lexical access and thus word recognition (Lehtonen et al., 2012; Ransdell & Fischler, 1987) and reading for bilinguals compared to monolinguals. We expect that the weaker links account does not apply to late bilinguals, per se, because these participants might have experienced larger language exposure in general than monolinguals have and because lexical entrenchment of L1 words is in an advanced stage before learning an L2.

Although Gollan et al. 's (2011) eye tracking study does explicitly compare English monolinguals with balanced Spanish-English bilinguals on an English reading task, their bilingual group scored worse on the objective English proficiency measure than their monolingual group did. Bilinguals accordingly showed longer gaze duration and lower skipping rates for the target words than monolinguals did. It is thus unclear whether this difference is a necessary and intrinsic consequence of bilingualism or rather whether it is driven by proficiency.

In our study, we excluded language proficiency as a possible confounding variable by matching our bilingual’s L1 proficiency to our monolingual’s language proficiency. Note, that similar proficiency scores would already imply that the lexical entrenchment of the bilinguals' L1 is on the same level as the lexical entrenchment of the monolinguals.

In conclusion, the weaker links account predicts slower sentence reading times, more and longer fixations per sentence, a smaller saccade length, lower skipping rates and higher regression rates, for bilinguals reading in L1 than for monolinguals. These differences will be subtler than the differences between the bilingual L1 and bilingual L2 reading pattern, because the L1 proficiency is the same for both groups. When we assume similar L1 lexical entrenchment for unbalanced bilinguals, we would expect a similar global eye movement pattern for monolinguals and bilinguals reading in their L1.

METHOD

The ethical committee of the University of Ghent approved the experimental procedure (nr. 2011/44). Participants signed an informed consent form prior to starting the experimental procedure. A summary of this method is included in Cop, Keuleers, Drieghe, and Duyck (2015), because that study presented other analyses of the same eye-tracking corpus data, focusing specifically on word-level frequency effects, rather than the broad sentence-level differences investigated in the present study.

PARTICIPANTS

Nineteen unbalanced Dutch (L1) – English (L2) bilingual Ghent University and fourteen English monolingual Southampton University undergraduates participated either for course credit or monetary compensation. Bilingual and monolingual participants were matched on age and education level. The average age was 21.2 years for bilinguals [range: 18-24; sd=2.2] and 21.8 years for monolinguals [range: 18-36, sd=5.6]. All of the participants were enrolled in a bachelor or master program of psychology. In the monolingual group, 6 males and 7 females participated. In the bilingual group, 2 males and 17 females participated.

Participants had normal or corrected-to-normal vision. None of the participants reported to have any language and/or reading impairments.

The bilinguals had a relatively late age of acquisition for L2: The mean age of acquisition was eleven years [range: 5-14, sd = 2.46]. All participants completed a battery of language proficiency tests, including a spelling test, the LexTALE (Lemhöfer & Broersma, 2012) and a lexical decision task (for results see Table 1). For the bilinguals, a self-report language questionnaire was added. This contained questions about language switching frequency/skill, age of L2 acquisition, frequency of L2 use and reading/auditory comprehension/speaking skills in L1 and L2. All of the bilinguals report that they can carry on a conversation, read and comprehend instructions, sometimes read articles, books, watch TV shows and listen to music in English (their L2). The bilinguals report that they use their L2 on average 3.6 days a week (range: 1-7 days). About half of the bilinguals also report that they sometimes think or talk to themselves in English (for a detailed summary, see Appendix A, Table A.1 and A.2). Due to the lack of a standardized cross lingual spelling test, we tested the English spelling with the spelling list card of the WRAT 4 (Wilkinson & Robertson, 2006) and the Dutch spelling with the GLETSCHR (De Pessemier & Andries, 2009). The LexTALE (Lexical Test for Advanced Learners of English) is an unspeeded lexical decision task, which is an indicator of language proficiency for intermediate to highly proficient language users, validated for English, Dutch and German (Lemhöfer & Broersma, 2012). Two bilinguals were classified as lower intermediate L2 language users (50%-60%), ten bilinguals were classified as upper intermediate L2 language users (60%-80%), seven bilinguals scored as advanced L2 language users (80%-100%) according to the LexTALE norms reported by Lemhöfer and Broersma (2012). A classical speeded lexical decision task was also administered in Dutch and English for the bilinguals, in English for the monolinguals. We calculated a composite proficiency score by averaging the score on the spelling test, the score on the LexTALE and the adjusted score of the L2 lexical decision task. Table 1 shows, mean accuracy for the spelling tests and

LexTALE, lexical decision word accuracy corrected for false alarms, and the composite proficiency score.

Most important, the Dutch (L1) proficiency of the bilinguals was matched with the English proficiency of the monolinguals (See column 5 in Table 1), indicating that both groups were equally proficient in their first language. The English (L2) proficiency is clearly lower than the Dutch (L1) proficiency (see column 4 in Table 1).

Table 1. *Average percentage scores [standard deviations] on the LexTALE, Spelling test and Lexical Decision task for the bilingual and monolingual group. T-values [degrees of freedom] of t-tests in the last 2 columns.*

	Monolinguals	Bilinguals L1	Bilinguals L2	t-value L1-L2	t-value L1-mono
LexTALE score (%)	91.07 [8.92]	92.43 [6.34]	75.63 [12.87]	7.59 [18] ***	0.49 [22.3]
Spelling score (%)	80.78 % [7.26]	83.16 [7.80]	69.92 [8.74]	8.15 [18] ***	0.99 [29.3]
Lexical Decision score (%)	77.89 [12.01]	80.47 [5.45]	56.75 [11.01]	9.87 [18] ***	0.67 [17.1]
Composite Proficiency Score (%)	83.25 [8.30]	85.54 [4.68]	67.81 [9.72]	11.78 [18] ***	0.93 [19.1]

* p<0.05, ** p<0.01, *** p<0.001

MATERIALS

The participants read the novel "The mysterious affair at Styles" by Agatha Christie (Title in Dutch: "De zaak Styles"). This novel was selected out of a pool of books that were available in a multitude of different languages (for possible future replication in other languages) and which did not have any copyright issues. All of these books were selected from the Gutenberg collection that is freely available on the Internet. We selected the novels that could be read in four hours. The remaining books were inspected for difficulty, indicated by the frequency distribution of the words that the book contained. The Kullback–Leibler divergence (Cover & Thomas, 1991) was used to select the novel whose word frequency distribution was the most similar to the one in natural language use (according to the subtlex database). This novel also had one of the lowest number of hapax words (words that occur only once in the subtlex database) of the selected books.

Table 2 shows a summary of the characteristics of the Dutch and English version of the novel. The difference in number of words per sentence and average word length illustrates that English is a denser language than Dutch. Although the differences in absolute values were very small, paired t-tests still yielded significant differences between the two languages concerning number of words per sentence and average word length, because of the extremely big corpus size (*n* = 5 212). The difference between average content word frequencies was not significant.

Table 2. *Summary of the characteristics of the translation equivalent sentences and the restricted set of sentences matched on information density (averages of Word Length, Number of Words per sentence, Number of Characters per sentence, Number of Content words per sentence, Word Frequency and Content word frequency) across languages.*

Descriptive parameters	Translation equivalent sentences			Restricted set of sentences		
	Dutch	English	T value	Dutch	English	T value
Number of Words	55 596	51 594	-	1 628	1 628	-
Number of Sentences	4 804	4 804	-	210	210	-

Number of Words per Sentence	11.52 [8.89]	10.73 [8.10]	5.06***	7.53 [6.65]	7.53 [6.65]	-
Number of Characters per Sentence	51.76 [41.27]	43.28 [34.25]	12.40 ***	32.62 [28.76]	31.46 [27.85]	6.48 ***
Number of Content Words per Sentence	5.87 [4.58]	5.33 [4.06]	6.86 ***	3.68 [3.34]	3.76 [3.46]	-1.19 (p=0.24)
Average Word Frequency	4.49 [0.60]	4.57 [0.59]	-6.86 ***	4.29 [0.88]	4.37 [0.87]	-3.23 **
Average Content Word Frequency	3.84 [0.75]	3.85 [0.76]	-0.40 (p=0.69)	3.89 [0.85]	3.86 [0.84]	0.86 (p=0.39)
Average Word Length	4.52 [1.04]	4.18 [0.97]	16.93***	4.54 [1.42]	4.52 [1.42]	1.33 (p=0.19)

* p<0.05, ** p<0.01, *** p<0.001

APPARATUS

The bilingual eye movement data were recorded with a tower-mounted EyeLink 1000 system (SR-Research, Canada) with a sampling rate of 1 kHz. A chinrest was used to reduce head movements. Monolingual eye movement data were acquired with the same system that was desktop mounted. Reading was always binocular, but eye movements were recorded only from the right eye. For the bilingual participants, sentences were presented on a 22 inch Philips 202P70 CRT-monitor and for the monolingual participants, sentences were presented on a 21 inch g225f view Sonic graphics series monitor. Text was presented in black 14 point Courier New font on a light grey background. The lines were triple spaced and 3 characters subtended 1 degree of visual angle or 30 pixels. Text appeared in paragraphs on the screen. A maximum of 145 words, spread over a maximum of 10 lines, was presented on one screen. During the presentation of the novel, the room was dimly illuminated.

PROCEDURE

Participants read the entire novel in four sessions of an hour and a half. One bilingual participant read only the first half of the novel in English in two sessions. In the first session, every participant read chapter 1 to 4. In the second session chapters 5 to 7, in the third session chapters 8 to 10 and in the fourth session chapter 11 to 13 were read. The bilinguals read half of the novel in Dutch, the other half in English. The order was counterbalanced. The monolinguals read the entire novel in English. Every bilingual and monolingual participant completed a number of language proficiency tests. The results of these proficiency measures can be found in Table 1.

The participants were instructed to read the novel silently while the eye tracker recorded their eye movements. It was stressed that they should move their head and body as little as possible while they were reading. The participants were informed that there would be a break after each chapter and that in that pause they would be presented with multiple-choice questions about the contents of the book. This was done to ensure that participants understood what they were reading and paid attention throughout the session. The number of questions per chapter was relative to the amount of text in that chapter.

The text of the novel appeared on the screen in paragraphs. When the participant finished reading the sentences on one screen, they were able to press the appropriate button on a control pad to move to the next part of the novel.

Before starting the practice trials, a nine-point calibration was executed. The participants were presented with three practice trials where the first part of another story was presented on the screen. After these trials, the participants were asked two multiple-choice questions about the content of the practice story. This part was intended to familiarize participants with the reading of text on a screen and the nature and difficulty of the questions. Before the participant started reading the first chapter another nine-point calibration was done. After this, the calibration was done every 10 minutes, or more frequently when the experiment leader deemed necessary.

RESULTS

As described above, we analyzed the eye movement data at the sentence level. Data collection contained 5 212 data points or sentences per subject. Fixations shorter than 100ms were excluded from analyses. 243 (4.9%) unusual sentences were removed because they contained more than 35 words, had an average word length of more than 7.4 characters or had an average content word frequency lower than 1.56. This left us with 4 969 sentences per subject on average.

The bilinguals scored 81% [sd=13.36] on the L1 multiple-choice questions and 79% [sd=12.54] on the L2 multiple-choice questions. A paired t-test did not yield a significant difference between these two ($t=0.275$, $df=17$, $p=0.787$). The monolinguals scored on average 78% [sd=9.46]. A t-test did not yield a significant difference between the bilingual L1 and the monolingual comprehension scores ($t=0.675$, $df=29.79$, $p=0.505$). See Appendix B for the questions and multiple-choice answers. See Appendix C for a link to all data files, including full subject information, sentence-level materials and eye movements.

ANALYTIC TECHNIQUES FOR CROSS-LANGUAGE COMPARISON

Following our rationale, two comparisons are essential for this paper. The first one is the within-subject comparison of the bilingual L1 and L2 reading data to explore the influence of "Language" (L1 or L2); the second one is the comparison between bilingual L1 and monolingual reading in order to assess the possible effects of being a bilingual. Both comparisons imply by definition the need to directly compare reading behavior across two different languages. There might be inherent differences between languages relating to formal characteristics, information density and difficulty. This necessitates matching for inherent language differences that may influence basic reading characteristics. We tested Dutch-English bilinguals reading a novel in both Dutch and English. Dutch is the closest major language relative to English, so that this language pair is the best-suited combination starting from the dominant language in the reading literature (English).

First, there is a need for matching the materials on semantic content. We manually checked each sentence for translation equivalence. The sentences that did not match this criterion, and thus had slight semantic differences across languages, were excluded from all of the following analyses. 4 764 sentences per subject were retained for analysis (3.99% of Dutch and 3.95% of English sentences were excluded). The sentences were numbered pairwise and this “sentence identity number” will be used in the analysis.

Second, information density is an indication of the amount of syllables needed to convey a certain semantic content (Pellegrino, Coupé, & Marsico, 2011). As we can see in Table 2, there are significant differences between measures of information density (average word length and number of words per sentence) for the two texts in the different languages. By including these factors as fixed effects in our linear mixed model, we made sure that the significance of the other fixed effects in the model is not affected by these differences. To be even more conservative, we created a more restricted data set by matching the sentences pairwise on average word length (threshold = 0.2 characters per sentence) and number of words per sentence (exactly matched) to equalize information density for each translation equivalent English-Dutch sentence pair. After this, text difficulty, as measured by the mean frequency of the content words, was still matched across languages. Only 4.2 % of the sentences were retained in this selected dataset. This selection still contained 210 sentences per subject (for a summary of the lexical variables for the matched material set see Table 2). We report the results for this restricted, optimally matched data set, extracted from the natural reading corpus data.

MODEL FITTING

For analysis, we selected the dependent variables that are well captured by models of reading such as the E-Z reader model. For both comparisons, the dependent measures under investigation are: a) sentence reading time including fixations and re-fixations, b) total number of fixations that landed in one sentence, c) the average fixation duration of the fixations that landed in that sentence, d) the average rightward saccade length per sentence, e) the probability of making an inter-word regression towards or within a certain sentence and f) the probability of first pass skipping.

Our data corpus was analyzed with linear mixed effects models with the lme4 (version 1.1-7) and lmertest (version 2.0-20) package of R (version 3.0.2) (R Core Team, 2014), because a multilevel design is the best way to statistically control for a range of predictors that in this experiment we could not or did not want to manipulate.

For the first within-subject comparison of the bilingual L1 vs. L2 reading data, the same fixed effects model was fitted for every eye movement measure. The fixed factors were language (L1 or L2), number of words per sentence (continuous), average word length per sentence (continuous), average frequency of the content words per sentence (continuous) and L2 proficiency (continuous). This last variable is the composite proficiency score presented in Table 1. Note that this variable represents something different for the two language conditions. For the L2 condition this is the language they are reading in. For the L1 condition it is their proficiency in a second language that they do not use in this condition. For the content word frequency, the subtitle word frequency measures (Brysbaert & New, 2009; Keuleers, Brysbaert, & New, 2010) of the content words in a particular sentence were log transformed to normalize their distribution. All continuous predictors were centered. The absolute value of the maximum correlations among main effects was under 0.51 for all eye movement measures (<0.506 for Saccade length, <0.156 for fixation count, <0.167 for fixation duration, <0.249 for dwell time, <0.386 for regressions, <0.245 for skips).

In a first step, we fitted a "complete" model. The fixed part of the model contained all main effects and interactions (up to 5-way) and the random part contained two random clusters: one for subject (the participant ID-number) and one for sentence (the sentence ID-number). After fitting this first model, we excluded the terms one by one, starting with the factor that contributed the least to the fit. By model comparisons, we decided when we arrived at the best possible fit. Then we added random slopes one by one. When they contributed to the fit, we included the slope in the model. We choose to test addition of every possible random slope, and strive for a maximal random structure (Barr, Levy, Scheepers, & Tily, 2013). We added, in this order, language as a random slope for each sentence and language, word length, word frequency and number of words as random slope per subject. For the count variable and the binomial variables (fixation count, skipping rate, regression rate) we report the p-values for the significant effects. For the continuous variables (sentence reading times, average fixation duration and saccade length), we obtained the p-values by computing the F-Test with Kenward-Roger adjusted degrees of freedom (Kenward & Roger, 1997) for our fixed effects in the final models.

For the second important (between-subject) comparison between the bilingual L1 and monolingual L1 reading, the same model was fitted for every eye movement measure. Here, the fixed factors were bilingualism (Bilingual or Monolingual), number of words (continuous), average word length (continuous), average frequency of the content words (continuous) and L1 proficiency (continuous). This last variable is the composite proficiency score presented in Table 1. Note that for both the bilinguals and the monolinguals this is the language they are reading in. The frequency measure was computed the same way as in the previous comparison. The process of top-down fitting of fixed effects and bottom-up fitting of the random slopes was identical to the process in the first comparison. Again, a maximum random structure was aspired but this time we added, in this order, bilingualism as a random slope for each sentence and word length, word frequency and number of words as random slope per subject. Again, the p-values for the continuous variables were calculated with the F-test with Kenward-Rogers adjusted degrees of freedom (Kenward & Roger, 1997).

BILINGUAL L1 VS. BILINGUAL L2 READING

SENTENCE READING TIME

Sentence reading times that differed more than 3 standard deviations from the general mean reveal unusual distraction and were therefore excluded from the analysis (5.02%). Sentence reading times were log transformed as suggested by the Box-Cox method (Box & Cox, 1964) to obtain a more normal distribution and then analyzed with the linear mixed model described above.

A main effect was found for language ($F=36.43$, $df=24.70$, $p<0.001$): the bilinguals were 17% slower to read a sentence in their L2 than in their L1 (1.52s compared to 1.27s), a rather large effect. This indicates that reading text in a less proficient second language produced an obvious disadvantage. This disadvantage was larger in longer sentences as shown by the interaction between language and number of words ($F=9.92$, $df=207.54$, $p< 0.005$). In other words, an extra word per sentence prolonged the reading time of an L2 reader more than the reading time of an L1 reader (see Figure 1). This was probably caused by the fact that individual fixations were longer when reading in L2. This would accumulate into a longer reading time in longer sentences. Also, longer sentences often entail a higher syntactical complexity, which could come with a cost that is higher in L2 than in L1. When looking at the other dependent variables, it will become clear whether this explanation holds.

A main effect of word length ($F=19$, $df=232.71$, $p<0.001$) and number of words per sentence ($F=80.89$, $df=21.84$, $p<0.001$) was found. Obviously, longer reading times were found with sentences with longer words and more words. The interaction between these two variables was also significant ($F=14.20$, $df= 233.24$, $p<0.001$). They reinforce each other's effect (Figure 2). Apparently long sentences add an additional cost to the reading process when reading long words and do so more for L2 than L1. We did not find a main effect of L2 proficiency on sentence reading time or an interaction of L2 proficiency with language. In our dataset there was no evidence that L2 reading speed was altered by L2 proficiency.

None of the 3-way, 4-way or 5-way interactions contributed significantly to the fit of the model (all $\chi^2 < 2.01$).

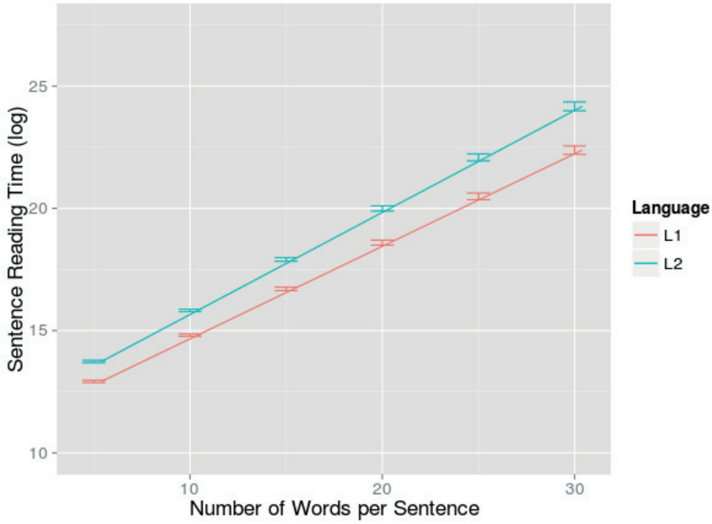


Figure 1. Sentence reading time (log-transformed on the y-axis) in function of number of words (on the x-axis) per sentence for bilinguals reading in L1 and L2. The standard errors are indicated by whiskers on the graph.

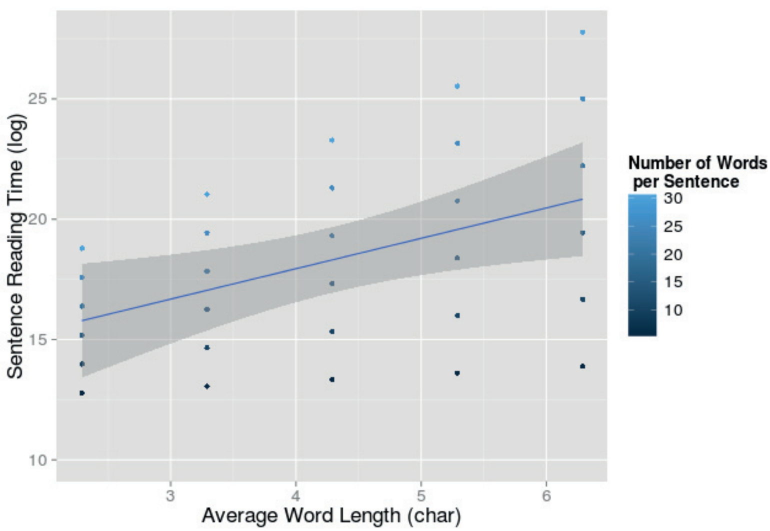


Figure 2. Sentence reading time (log-transformed on the y-axis) in function of average word length per sentence (on the x-axis) and number of words per sentence. The 95% confidence interval for the main effect of word length is indicated in grey.

NUMBER OF FIXATIONS PER SENTENCE

Sentences with fixation counts more than 3 standard deviations from the subject means were excluded (2.15%). The fixation counts per sentence were analyzed with a generalized linear mixed model with a Poisson distribution.

A main effect of language was found (β=0.200, z=6.87, $p<0.001$): bilinguals made 13% more fixations in their L2 than in their L1 (6.75 fixations compared to 5.88 fixations). The E-Z reader model predicts more fixations when words get longer. Indeed, a main effect of word length (β= 0.168; z=3.92, $p<0.001$) was found. A main effect of number of words (β= 0.101; z=28.73, $p<0.001$) was also found, which interacted significantly with word length (β=0.0170; z=3.03, $p<0.005$). Again in longer sentences, the burden put on the reader by longer words increased for reading in L1 and L2. The word length effect was present both in L1 and L2 reading, but behaved in a different way: a significant interaction was found between language and word length (β=-.0555; z=-2.43, $p<0.05$). The effect of word length was smaller for L2 reading and the difference in fixation count for L1 versus L2 was smaller in the sentences with the longer words. This might be explained by the slower lexical processing in L2. When reading in L2, the eyes stayed on a certain word, short or long, for a longer period of time. This might have limited the need for a second fixation to longer words in L2, relative to L1 (See Figure 3). A main effect of L2 proficiency (β= -0.00828; z=-2.21, $p<0.05$) was also found (See Figure 4). As L2 proficiency increased, the number of fixations decreased, also when reading in the mother tongue. This is not surprising because the correlation between the proficiency in L1 and in L2 was 0.76. It is important to note that the interaction between language and proficiency was not significant: even for the bilinguals who are very proficient in their L2, the fixation count was higher in L2 than in L1. The participants scoring 50%-65% on their L2 proficiency fixated on average 6.73 times. The participants scoring above 70%-85% fixated on average 5.79 times. None of the 3-way, 4-way or 5-way interactions contributed significantly to the fit of the model (all $\chi^2<3.24$).

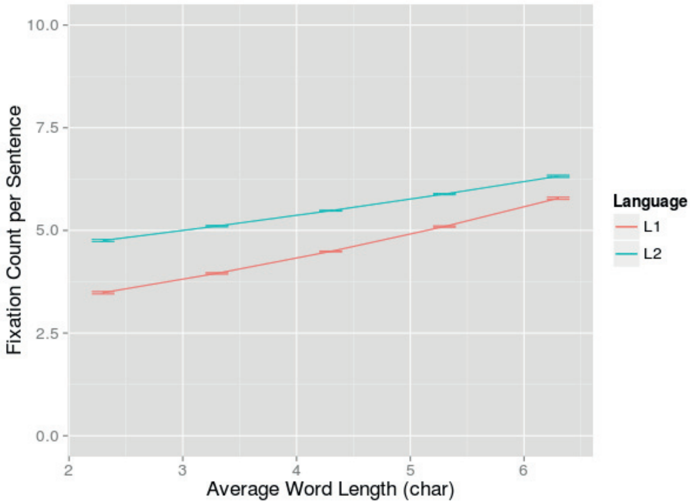


Figure 3. Fixation count per sentence (on the y-axis) in function of average word length per sentence (on the x-axis) for bilinguals reading in L1 and L2. The standard errors are indicated by whiskers on the graph.

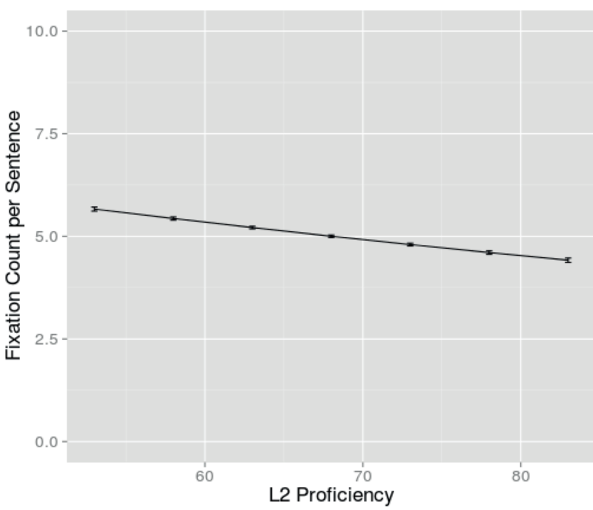


Figure 4. Fixation count per sentence (on the y-axis) dependent on the participant's L2 composite proficiency score (on the x-axis). The 95% confidence interval is indicated by the dotted lines.

AVERAGE FIXATION DURATION

Sentences with an average fixation duration differing more than 3 standard deviations from the general mean were excluded (8.64%).

A main effect of language was found ($F=22.06$, $df=193.61$, $p<0.001$): bilinguals fixated on average 9% or 20ms longer in their L2 than their L1 (238.72ms compared to 218.74ms). This explains the effect that we found when analyzing the Sentence Reading Times: longer sentences prolonged the reading time significantly more in L2 than in L1. For each fixation, extra time was added to the total sentence reading time. Because this additional time was longer for L2, we got a steeper incline in reading time. This finding combined with the higher fixation count in L2 is compatible with a child like reading pattern in L2, caused by a slower second language processing.

A main effect of number of words ($F=7.3$, $df=62.4$, $p<0.01$) was found and this variable interacted with language ($F=14.57$, $df=195.87$, $p<0.001$). This interaction shows us that only in L2, the average fixation durations were longer when the sentences were longer.

The 3-way interaction between language, number of words and frequency ($F=6.41$, $df=201.91$, $p<0.05$) was significant (See Figure 5).

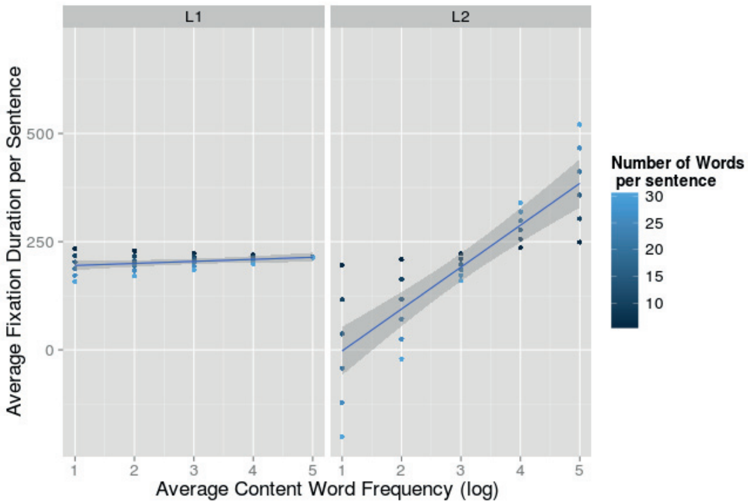


Figure 5. Average fixation duration per sentence (on the y-axis) dependent on average content word frequency per sentence (log-transformed on the x-axis) and number of words per sentence for Bilinguals reading in L1 and L2. The 95% confidence interval of the main effect of content word frequency per language is indicated in grey.

Word frequency is the most frequently investigated determinant of word fixation times. Low frequency words normally yield longer fixation durations, but because we were looking at the average fixation duration including re-fixations and skips, we expected a reversed effect. A high frequent word might receive just a single fixation, while more difficult, less frequent words might receive two or even three fixations. These fixations will be shorter than the single one, but the sum of the two will be longer (Radach, Huestegge, & Reilly, 2008). Indeed, in L2 we found this reversed frequency effect in sentences that contain more than 9 words. When the average content word frequency was low, i.e. sentences with more difficult words, bilinguals fixated shorter on average.

We did not detect this frequency effect in L1, probably because most words received just a single fixation (74.76 % of the fixated words in L1 versus only 65.82 % of the fixated words in L2).

The interaction between language and word length also reached significance and indicated that there was an effect of word length (F=8.18, df=195.87, p<0.01) only when reading in L2, and more specifically that in sentences with longer words the average fixation duration was longer (See Figure 6).

The 3-way interaction between language, number of words and word length (F=6.62, df=195.84, p<0.05) was significant (See Figure 6). In L2, the effect of word length was bigger in sentences with more words. This resulted in inflated fixation durations when long words were positioned in long sentences. In sentences containing very short words, fixation durations were longer in short sentences. In sentences with short words the fixations get shorter in longer sentences, and in sentences with long words the reverse happens. This means that longer words, pose a larger burden on the reading and language processing mechanisms when reading in L2 than in L1. Again L2 proficiency did not influence the average fixation duration of our participants, while reading in L1 or L2. None of the 4-way or 5-way interactions contributed significantly to the fit of the model (all $\chi^2 < 2.65$).

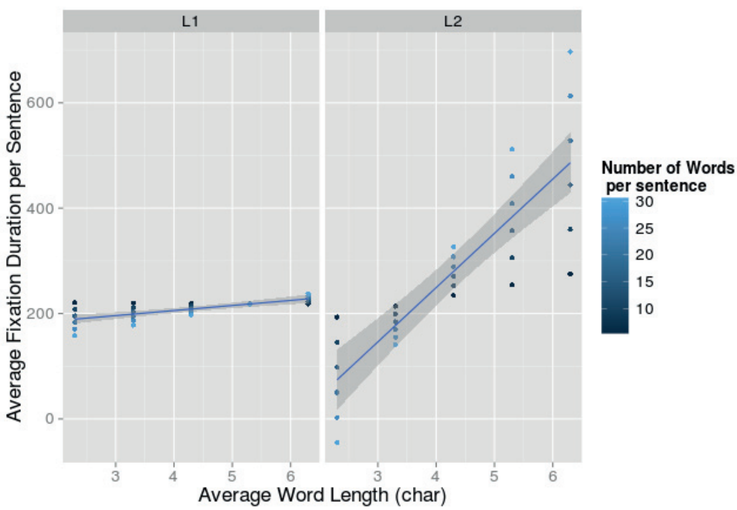


Figure 6. Average fixation duration per sentence (on the y-axis) dependent on average word length per sentence (on the x-axis) and number of words per sentence for bilinguals reading in L1 and L2. The 95% confidence interval for the main effect of word length per language is indicated in grey.

RIGHTWARD SACCADE LENGTH

We analyzed the average saccade length per sentence of the saccades that were directed to the right. The saccades during which the participant blinked and sentences with an average saccade length differing more than 3 standard deviations from the general mean were excluded (1.67%). The Box-Cox method (Box & Cox, 1964) determined that the log transformation of the variable was optimal to achieve a normal distribution. This log of the average saccade length was analyzed.

A main effect of language was found ($F=30.77$, $df=66.56$, $p<0.001$): bilinguals moved their eyes across 12% shorter distances when reading in L2 than in L1 (8.30 compared to 9.35 characters). This result is again in line with our child like reading hypothesis and ties in with the fact that more fixations were made in L2. It has been shown that reading skill influences the size of the perceptual span seeing that beginning readers have smaller perceptual spans than more skilled readers (Häikiö et al., 2009; Rayner, 1986). It is plausible to assume that the same is going on for participants reading in their L2. Because of this smaller perceptual span, less parafoveal processing is possible and people move their eyes more close to their previous fixation. The risky reading strategy that we hypothesized, states that bilinguals might make longer saccades and skip more words in L2. Our bilingual participants did not seem to do that.

A main effect of number of words ($F=17.35$, $df=98.84$, $p<0.001$) was found. Participants moved their eyes further in sentences with more words. Balota, Pollatsek, and Rayner (1985) showed that readers skipped more words when they were predictable in the sentence context. This causes participants to make longer saccades. It is probable that words are more predictable in long sentences because the preceding sentence context is more semantically restrictive, but this requires further investigation.

Where to move the eyes is strongly influenced by low-level variables like word length and space information. Longer words usually lead to longer saccades (Inhoff, Radach, Eiter, & Juhasz, 2003). We did not find an effect of word length. This is due to the fact that we include both intra-word and inter-word saccades in this analysis. This means that for long words, that were often fixated more than once, saccades were shorter. This probably balances out the effect that we would find for the inter-word saccade length, namely that long words would elicit longer saccades.

A significant interaction was found between language and number of words ($F=4.60$, $df=151.58$, $p<0.05$). This suggests a differential number of words effect. In other words, the difference between saccade length in L2 and L1 reading was bigger for sentences with more words (see Figure 7). This could point towards the fact that when reading in L2, participants predicted less of the upcoming words than when reading in L1. None of the 3-way, 4-way or 5-way interactions contributed significantly to the fit of the model (all $\chi^2 < 2.57$).

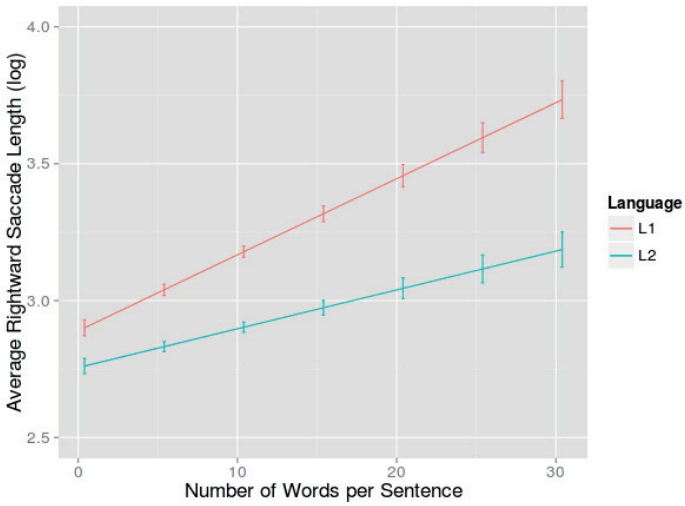


Figure 7. Average saccade length per sentence (on the y-axis) dependent on average number of words per sentence (on the x-axis) for bilinguals reading in L1 and L2. Standard errors are indicated with whiskers on the graph.

SKIPPING RATE

The probability of skipping a word in the first pass was analyzed. We fitted a linear mixed effect model with a binomial distribution.

The main effect of language was significant (β= -0.202; z=-4.180, p<0.001). In line with expectations, participants skip 5% more words when reading in their L1 (52.22%) compared to reading in their L2 (47.62%). Skipped words are thought to be processed on the fixation prior to the skip, when the word was still in the parafovea, and in part after the skip (Rayner, 2009; Reichle et al., 1998). We found that bilinguals skip fewer words when they read in their least proficient language. This result was thus in line with slower language processing in L2, allowing less time for the parafoveal processing of the next word when reading, resulting in less skipping. This does not point towards the possibility that bilinguals might use a risky reading strategy when reading in L2 (Rayner et al., 2006).

Word length has been found to be the most important determinant of word skipping (Brysbaert & Vitu, 1998). Very short words were skipped fairly often, while words of 9 or more characters were almost never skipped. We indeed found an effect of word length on skipping rate (β= -0.120; z=-4.104, p<0.001). More specifically: When sentences contained longer words, the probability of skipping those words was lower. None of the interactions contributed significantly to the fit of the model (all $\chi^2 < 1.73$).

REGRESSIONS RATE

Finally, probabilities of making a regressive eye movement were analyzed. The saccades during which the participant blinked were excluded from the analyses. A saccade was considered a regression when the eye moved from a word further in the sentence to a previous word (intra word regressions were not entered in the analyses). We fitted a linear mixed effect model with a binomial distribution.

The E-Z reader model states that regressions occur when there is difficulty with integrating a certain word in the current sentence context. This means that comprehension difficulties while reading a text can change the eye movement behavior. For example, when participants read garden-path sentences, they make more regressions to earlier parts of the text (Binder, Duffy, & Rayner, 2001). Although we expected that L2 readers would make more regressions, we did not found a higher regression rate when bilinguals read in their L2. No main effect of language was found (bilinguals made a regressive saccade in 22.63% of the cases in L1 and 24.07% of the cases in L2). The only significant effect was the interaction between language and word length ($\beta = -0.208$, $z = -2.039$, $p < 0.05$). In our data L2 readers do regress more than L1 readers, as expected, but only in sentences that contain relatively short words (on average 3.3 characters or less). In the more complex, longer sentences bilinguals made the same amount of regressions when reading in their L1 as in L2. When reading in L1, the longer the words, the more regressions were made (see Figure 8). This could be expected, because these words are usually harder to process, and more integration difficulties are likely to arise. This relationship reversed in L2. This pattern of more regressions towards short words can be explained by the fact that short words were skipped more often. It is thus more likely that such a word was not processed sufficiently and therefore that the reader has to return to that word. Although both patterns are plausible, it is still an open question why we found the former when bilinguals read in L1 and the latter when bilinguals read in L2. This might be because the average fixation duration was longer in L2 than in L1, especially in sentences with longer words. This means that the chance that a long word was not sufficiently processed in a first pass reading was lower in L2 than in L1. None of the 3-way, 4-way or 5-way interactions contributed significantly to the fit of the model (all $\chi^2 < 2.31$).

For a full summary of the averages and standard deviations of the eye movement variables for L1 and L2 reading, see Table 3.

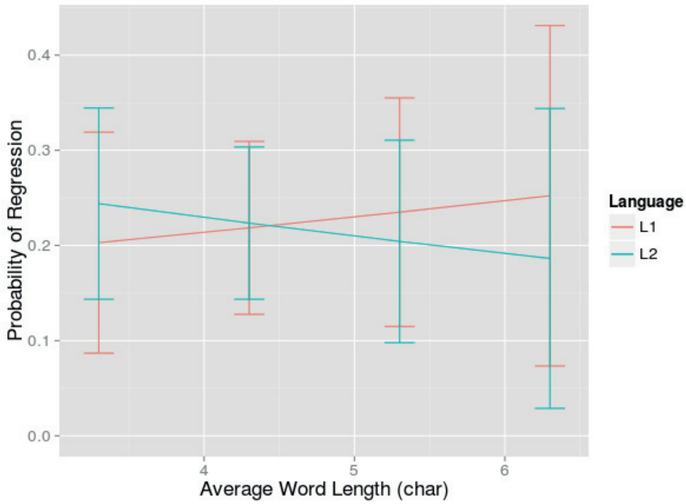


Figure 8. The probability of making a regression (on the y-axis) dependent on the average word length per sentence (on the x-axis) for reading in L1 and L2. The standard errors are indicated by whiskers on the graph.

BILINGUAL L1 READING VS. MONOLINGUAL READING

SENTENCE READING TIME

Sentence reading times that differed more than 3 standard deviations from the general mean reveal unusual distraction and were therefore excluded from the analysis (4.06%). Sentence reading times were log transformed as suggested by the Box-Cox method (Box & Cox, 1964) to obtain a normal distribution and then analyzed with the linear mixed model described above.

We did not find a main effect of bilingualism (F=2.46, df=49.9, p=0.123). Monolinguals read sentences in 1.28s, bilinguals in 1.25s. In order to exclude the possibility that this null effect was due to the use of a restricted (optimally matched on average word length, average word frequency and number of words per sentence) sentence set (n=210), we also analyzed sentence reading times of the translation equivalent sentence set (n=4 804). None of the interactions with the factor of bilingualism reached significance. The main effect of bilingualism was also not significant (F=1.55, df=49, p=0.22). This means that, in this dataset of natural reading, there is no evidence for a slower reading process on a sentence-level for bilinguals in L1 compared to monolinguals in L1. This finding is of great relevance, given that some recent studies in word production and word recognition suggested a considerable speed disadvantage for bilinguals. Gollan et al. (2005) and Ivanova and Costa (2008) found about 33-60ms (5-10%) slower L1 picture naming for bilinguals compared to monolinguals. In the visual word recognition domain, Lehtonen et al. (2012) and Ransdell and Fischler (1987) found 80 to 170ms (13-25%) slower L1 lexical decision times for bilinguals compared to monolinguals. This would correspond to a large difference of 166-320ms in sentence reading times here, which we did not find for natural reading. We found a main effect of number of words ($F=852.29$, $df=166.76$, $p<0.001$), of word length ($F=17.45$, $df=264.1$, $p<0.001$) and a significant interaction between the two ($F=12.86$, $df= 253.07$, $p<0.001$). Again these two variables reinforced each other's effect, so that in longer sentences the length of the words had a larger effect on sentence reading time (see Figure 9).

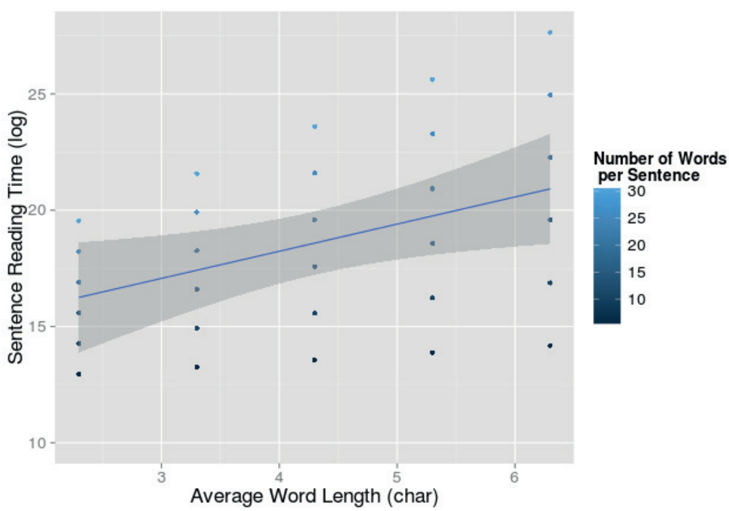


Figure 9. The sentence reading time (log-transformed on the y-axis) dependent on average word length per sentence (on the x-axis) and number of words per sentence for monolinguals and bilinguals reading in L1. The 95% confidence interval for the main effect of word length is indicated in grey.

We found a significant interaction between number of words and frequency ($F=4.05$, $df= 197.36$, $p = 0.045$) indicating that participants read faster when the content words of a sentence were more frequent, but only in longer sentences (See Figure 10). Reading time is a cumulative variable, so the difference between high and low frequency sentences probably only reached significance when there were enough words to be processed. In fact, the (sentence-level) frequency effect was even absent in sentences shorter than 9 words. We have to consider that the frequency measure we used in these models is a very coarse one. Given our focus on sentence-level effects, frequency is averaged over content words, but we do look at the reading time of all the words in the sentence. So this makes the frequency effect hard to detect. Indeed, in a recent paper we showed strong word-level frequency effects for bilinguals and monolinguals in the same eye-tracking corpus (Cop et al., 2015). None of the 3-way, 4-way or 5-way interactions contributed significantly to the fit of the model (all $\chi^2 < 1.37$).

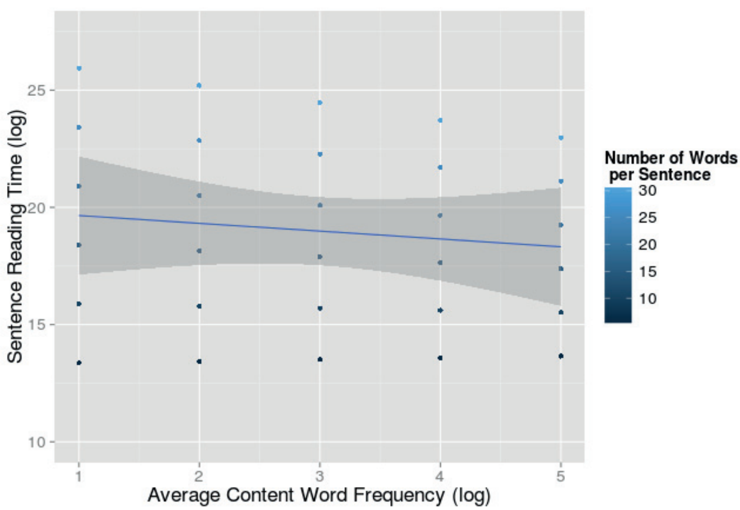


Figure 10. Sentence reading time (log transformed on the y-axis) in function of average content word frequency per sentence (log transformed on the x-axis) and number of words per sentence for monolinguals and bilinguals reading in L1. The 95% confidence interval of the main effect of content word frequency is indicated in grey.

NUMBER OF FIXATIONS PER SENTENCE

Sentences with fixation counts differing more than 3 standard deviations from the subject means were excluded (2.15% for the L1-L2 comparison and 0.4% for the L1-monolingual comparison). The fixation counts per sentence were analyzed with a generalized linear mixed model with a Poisson distribution.

The main effect of bilingualism was not significant. Monolinguals fixated on average 5.63 times, while bilinguals reading in L1 fixated on average 5.59 times, almost exactly the same. Native language reading yielded the same amount of fixations for bilinguals and monolinguals. A main effect of number of words (β= 0.106, z=26.26, $p<0.001$) and word length (β= 0.151, z=3.79, $p<0.001$) was found. Sentences that contain more words or longer words, received more fixations. The interaction between these two variables was also significant (β= 0.0103, z=2.00, $p<0.05$): They strengthened each other's effect. Although the effect of the number of words in a sentence was present for all word lengths, we only found a word length effect in sentences with more than 9 words.

A significant interaction between bilingualism and word length was also found (β=0.0403, z=-2.00, $p<0.05$). Bilingualism also interacted significantly with number of words per sentence (β= -0.00451, z=-2.46, $p<0.05$). In both cases the effects of the latter variable was larger for the bilinguals compared to the monolinguals, although both were reading in their first language (See Figure 11 and 12).

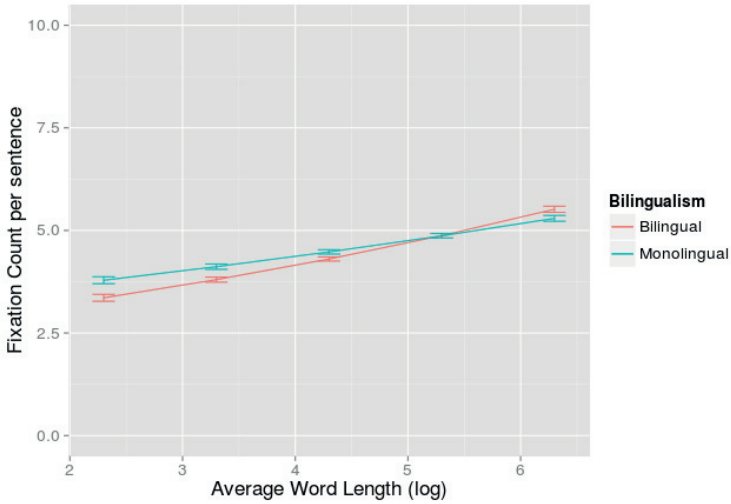


Figure 11. Fixation count per sentence (on the y-axis) in function of average word length per sentence (on the x-axis) for bilinguals reading in L1 and monolinguals (separate regression lines). Standard errors are indicated by whiskers on the graph.

The average word length of the sentences had a larger impact on how many times a participant fixates in a certain sentence when this participant is a bilingual than when he is a monolingual. Sentences with an average word length smaller than 5 characters were fixated less and sentences with an average word length larger than 5 were fixated more by bilinguals than by monolinguals. Also, bilinguals needed to fixate slightly more in long (more than 20 words) sentences compared to monolinguals, but this effect was relatively small.

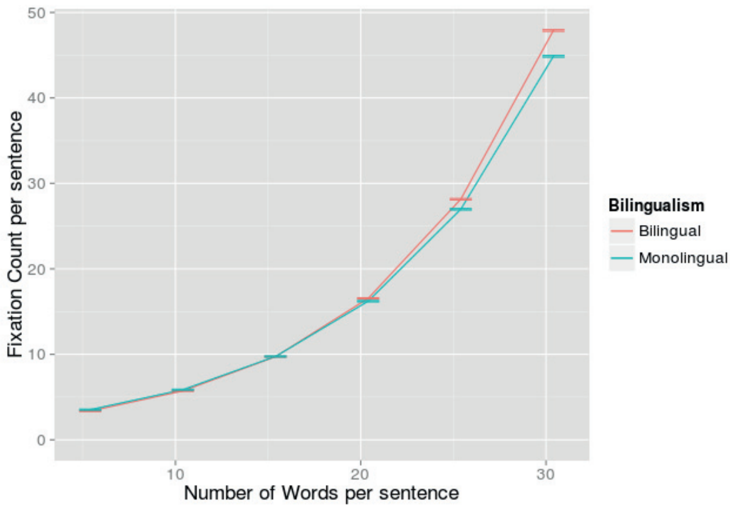


Figure 12. Fixation count per sentence (on the y-axis) in function of number of words per sentence (on the x-axis) for monolinguals and bilinguals reading in L1 (separate regression lines). The standard errors are indicated by whiskers on the graph.

We also found a significant interaction between frequency and number of words (β= -0.0125, z=-2.59, p<0.01): In long sentences we found a frequency effect. This means that there were more fixations in the sentences with a lower average word frequency (See Figure 13). In short sentences this effect was absent. Because of the focus on sentence-level effects, the average content word frequency measure we used is not a sensitive measure and would be even less accurate for shorter sentences. The effect was enlarged because fixation count is a cumulative variable. A more sensitive word level analysis will probably reveal larger and more ubiquitous frequency effects. None of the 3-way, 4-way or 5-way interactions contributed significantly to the fit of the model (All χ^2 <1.3).

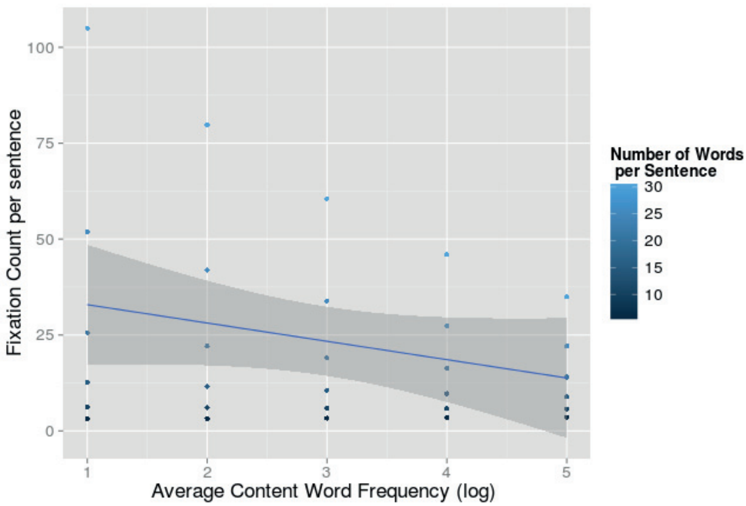


Figure 13. Fixation count per sentence (on the y-axis) in function of average content word frequency per sentence (log-transformed on the x-axis) and average number of words per sentence. The 95% confidence interval of the main effect of content word frequency is indicated in grey.

AVERAGE FIXATION DURATION

Sentences with an average fixation duration differing more than 3 standard deviations from the general mean were excluded (6.06%).

No effect of bilingualism was found. Bilinguals fixated on average 213.42ms in their L1 and monolinguals fixated on average for 217.28ms. Being a bilingual did not alter the durations of the fixations. None of the effects contributed significantly to the fit of the model (All $\chi^2 < 2.73$).

RIGHTWARD SACCADE LENGTH

We analyzed the average saccade length per sentence of the saccades that were directed to the right. The saccades during which the participant blinked were excluded from the analyses. Sentences with an average saccade length more than 3 standard deviations from the general mean were excluded (4.69%). The Box-Cox method (Box & Cox, 1964) determined that the log transformation of the variable was optimal to achieve a normal distribution. This log of the average saccade length was analyzed.

The effect of bilingualism was not significant. Bilinguals reading in L1 did not move their eyes further than monolinguals (9.45 characters for bilinguals and 10.09 characters for monolinguals).

There was a significant effect of number of words (F=53.12, df=90.09, p<0.001). In longer sentences, longer saccades were made. Again this might have been due to the end of sentences being more predictable than the beginning, making saccades longer the further you progress in that sentence.

The effect of L1 proficiency was marginally significant (F=3.70,df=25.18, p=0.066). More proficient participants moved their eyes further. This finding clarifies that the knowledge of another language does not change the saccade strategy of the reader. It might however be influenced by the knowledge and proficiency of the language you are reading in. Again, this can be related to the development of children where they develop larger saccades as they augment their language skill. None of the interactions contributed significantly to the fit of the model (All $\chi^2 < 3.45$).

SKIPPING RATE

The probability of skipping a word in the first pass was analyzed. We fitted a linear mixed effect model with a binomial distribution.

We did not find a difference between the skipping probability for monolinguals (51.99%) and bilinguals reading in L1 (52.27%). Again as expected, the word length effect was significant (β=0.202; z=-4.303, p<0.001). In sentences with longer words, the skipping rate was lower. There was also a significant effect of number of words (β= 0.00586; z=2.792, p<0.01). In long sentences, the probability of skipping was higher than in short sentences. The 3-way interaction between number of words, word length and frequency was significant (β= -0.0222; z=-3.258, p<0.005). Sentences with longer words had a lower skipping rate, but this effect reversed in difficult, long sentences (see the two left panels of Figure 14). It seems that words were glossed over more when a sentence in L1, on a whole, became too difficult. When a sentence contained a lot of difficult words, the probability of skipping in those sentences with longer words was higher. None of 4-way or 5-way interactions contributed significantly to the fit of the model (All χ^2<1.74).

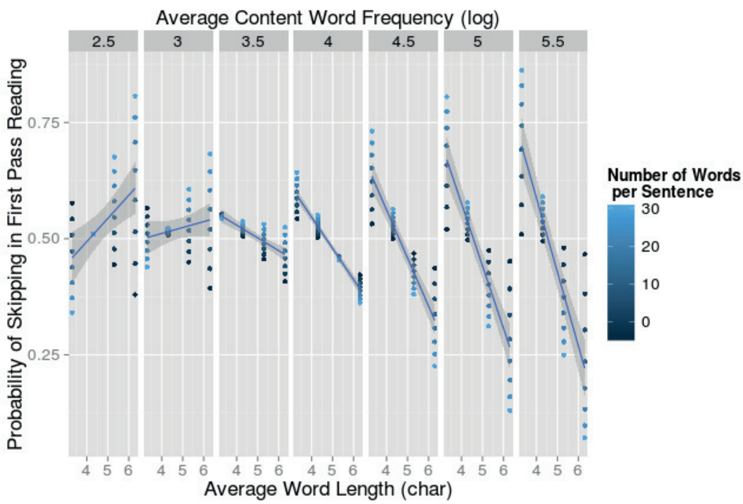


Figure 14. The probability of skipping a word in first pass reading (on the y-axis) dependent on average word length per sentence (on the x-axis), number of words per sentence and average content word frequency per sentence (log-transformed in the separate panels). The 95% confidence intervals for the effects of word length per content word frequency value are indicated in grey.

REGRESSIONS RATE

Finally, probabilities of making a regressive eye movement were analyzed. The saccades during which the participant blinked were excluded from the analyses. A saccade was considered a regression when the eye moved from a word further in the sentence to a previous word (intra word movements were not considered regressions). We fitted a linear mixed effect model with a binomial distribution.

No main effect of bilingualism was found (Regression rates 22.58% for bilingual L1 reading and 25.23% for monolingual reading). No other factors yielded significant effects. None of the interactions contributed significantly to the fit of the model (all χ^2 <2.89).

For a full summary of the averages of the eye movement measures of L1 and monolingual reading see Table 3.

Table 3. *Eye movement variable averages for young and older children and adults from Rayner's (1986) and Blythe et al.'s (2009) study and eye movement variable averages for bilingual L1/ L2 and monolingual reading. Differences between the means are reported in the last two columns [percentage] in each section.*

	Rayner (1986)				
Variable	7-8 year olds	11-12 year olds	adults	young children - adults	older children-adults
Sentence Reading Time (ms)	-	-	-	-	-
Fixation Count per sentence	15	8	6	9 (150%)	2 (33.3%)
Average Fixation Duration (ms)	280	240	235	45 (19.1%)	5 (2.1%)
Saccade length (characters)	2.8	6.4	6.8	-4 (-58.8%)	-0.4 (-5.9%)
Average skipping probability (%)	-	-	-	-	-
Average regression probability (%)	-	-	-	-	-

	Blythe et al. (2009)				
Variable	7-9 year	10 -11 year	adults	young children-adults	older children - adults
Sentence Reading Time (ms)	5473	4666	2965	2508 (84.6%)	1701 (57.4%)
Fixation Count per sentence	16.8	15.6	10.3	6.5 (63.1%)	5.3 (51.5%)
Average Fixation Duration (ms)	285	256	249	36 (14.5%)	7 (2.8%)
Saccade length (characters)	-	-	-	-	-

Average skipping probability (%)	39	44	44	-5 (-11.4%)	0 (0%)
Average regression probability (%)	-	-	-	-	-

	Our data				
Variable	Monolingual	Bilingual L1	Bilingual L2	Mono-L1	L2-L1
Sentence Reading Time (ms)	1279.34 [1030.49]	1254.41 [1073.06]	1522.98 [1293.51]	24.93 (1.9%)	268.6 (17.6%) ***
Fixation Count per sentence	5.63 [4.59]	5.59 [4.83]	6.75 [5.77]	0.04 (0.7%)	1.16 (17.2%) ***
Average Fixation Duration (ms)	217.28 [44.74]	213.42 [42.47]	238.72 [109.74]	3.86 (1.8%)	25.3 (10.6%) ***
Saccade length (characters)	10.09 [3.58]	9.45 [3.24]	8.30 [2.54]	0.64 (6.3%)	-1.15 (-13.9%) ***
Average skipping probability (%)	51.99 [49.96]	52.27 [49.95]	47.62 [49.95]	-0.28 (-0.5%)	-4.65 (-9.8%) ***
Average regression probability (%)	25.23 [43.43]	22.58 [41.81]	24.07 [42.75]	2.65 (10.5%)	1.49 (6.1%)

GENERAL DISCUSSION

We gathered a large comprehensive natural reading corpus of monolingual and bilingual eye movements. The goal of this study was twofold: a) to compare the eye movement pattern of bilinguals reading in L1 vs. reading in L2 and b) to compare the eye movement pattern of bilinguals reading in L1 vs. monolinguals reading in the mother tongue.

BILINGUAL L1 VS. BILINGUAL L2

We found clear sentence-level differences between L1 and L2 reading. In line with our expectations, and in concordance with the hypothesis of more child-like reading, we observed: a) 17.6% longer sentence reading times, b) 17.2% more fixations per sentence, c) 10.6% longer fixation durations, d) 13.9% shorter saccade lengths and e) a 9.8% lower probability of skipping a word in L2 compared to L1 reading (for more details see Table 3). Hence, slower sentence reading times in L2 were due to a higher amount of fixations, which were longer and closer together, and to the fact that fewer words were skipped.

COMPARISON WITH EYE MOVEMENT PATTERN OF CHILDREN

We predicted that the eye movement pattern of bilinguals reading in L2 would resemble the eye movement pattern of another kind of language learners, namely children. We will compare our L1-L2 results with Rayner's (1986) and Blythe et al.'s (2009) results of eye tracking studies in which children read sentences (summary in Table 3). Rayner tested three groups of children (7-8, 9-10 and 11-12 year olds) and adults. Each group read text material taken from textbooks suited for second grade children. Blythe et al. showed 3 groups of participants (adults, 7-9 year old children and 10-11 year old children) the same set of constructed sentences. We must note that the sentences that Blythe et al. presented were between 70-80 characters long and the ones that Rayner presented were 25-37 characters long, while ours were on average 32 characters long (56% shorter than Blythe's). The differences in the absolute size of fixation count and reading time between our data and Blythe et al.'s are probably due to this difference in sentence length. As you can see in Table 3, our L1 sentence reading times and fixation counts are about 55% lower than Blythe et al.'s adult sentence reading time and fixation count, while the adult fixation count in Rayner's study was comparable to our L1 fixation count.

Looking at Table 3, it is clear that the changes that L2 reading causes in the eye movement behavior are similar to, and in the same direction as, the changes that reading as a child, or an L1 learner entails: Sentence reading times, average fixation duration and fixation count increase, while rightward saccade length decreases for children compared to adults and for L2 readers compared to L1 readers. The exception is that Blythe et al. (2009) did not find a significant effect for skipping rates, while we did find less skipping for bilinguals reading in L2. A more recent study by Blythe et al. (2011) and one by Häikiö et al. (2009) did find a decrease in skipping rate of about 55%

for younger children and 20% for older children compared to adults. Another difference is that we did not find a difference between the regression rates for reading in L1 and reading in L2, while the largest part of the studies of children's eye movement studies found a higher regression rate for children (Blythe et al., 2011, 2009; Blythe & Joseph, 2011; Buswell, 1922; Joseph et al., 2009; Rayner, 1986; Taylor, 1965). In our data, the regression rate was only slightly higher in L2 than in L1 and only when the participant reads sentences containing short words. It is known that regression rates indicate integration difficulty. It is possible that because our participants have a relatively high L2 proficiency, they did not have more integration difficulties when reading in L2 compared to reading in L1 (which was confirmed in the text comprehension scores), while children do have more trouble integrating words in a cohesive sentence context. This might arise from the fact that children have less semantic knowledge than adults or from the fact that children have a more limited working memory capacity than adults (Dempster, 1985; Siegel, 1994), given that Just and Carpenter (1992) relate capacity of the working memory to text comprehension and semantic integration.

Looking at the sizes of the differences between L1 and L2 reading (Table 3), these are subtler and smaller than those found in the comparison between children and adults, except for average fixation duration. We explain this by the fact that our participants have already acquired the skills needed for efficient reading of an alphabetic language (their L1), despite the fact that our participants were not balanced bilinguals, and were clearly less proficient in their L2 (see Table 1).

Another similarity between the L2 and the children's eye movement pattern is the fact that in our dataset the effect of word length on average fixation duration only exists in L2. Studies show a larger word length effect on timed eye movement measures for children compared to adults (Blythe et al., 2011; Huestegge et al., 2009; Joseph et al., 2009). This suggests that both children and L2 readers need additional processing time for long words and are thus less efficient at lexical processing (Blythe & Joseph, 2011).

COMPATIBILITY OF RESULTS WITHIN E-Z READER MODEL

We will argue in the following paragraphs that all of the changes discussed above have one and the same underlying cause, which can be easily accounted for by the E-Z reader model.

The first cause of the longer reading times is the rise in the number of fixations when bilinguals read in L2. This is in part due to less skips and more re-fixations of words. Following the rationale of the E-Z reader model, when the eyes land in a word, the programming of an intra-word saccade is immediately initiated. When this programming is faster than the familiarity check of the fixated word, the intra-word fixation is made (Reichle, 2011). The higher fixation count in L2 reading can thus be related to a slower familiarity check, the first phase of lexical access.

The second reason for the slower reading speed is that the average fixation duration is longer for L2 reading compared to L1 reading. This difference is rather considerable (on average ± 20ms) and can also be related to a slower lexical processing for L2 reading. If more time is needed to identify a word in L2, the eyes should rest longer at the same location. This is exactly what we found.

The third one is that skipping of words is more rare when reading in L2. When the familiarity check of a parafoveal word is completed before the saccade programming to that word is completed, the E-Z reader model predicts that this word will be skipped (Reichle, 2011). More words are skipped in L1 than in L2. This probably means that the familiarity check can be completed faster when reading in the mother tongue than when reading in L2. It follows from the differences in skipping rate that when reading in L2, participants made smaller saccades compared to reading in L1 and monolinguals.

The differences between L1 and L2 reading concerning reading time, saccade length and average fixation duration are inflated in long sentences. This indicates that sentences with more words pose an extra burden on L2 language processing. This might be caused by the fact that longer sentences tend to be syntactically more complex and will have more clauses than short sentences. This will cause larger jumps from one part of the sentence to the next and longer fixation durations because of longer semantic integration times.

In conclusion, all of these findings are consistent with a more effortful familiarity check and slower overall lexical processing for bilinguals reading in L2. Considering that the familiarity check is dependent on word frequency, which is off course subjectively lower for L2 (weaker links), and predictability, the bilingual L2 disadvantage in visual language processing might be reduced to a quantitative difference of exposure to the lexical items in the lexicon. Reichle et al. (2013) already showed that the eye movement pattern of children could be modeled by simply reducing the rate of lexical processing. Given that we established a close parallel between patterns of eye movement in children and L2 readers, we hypothesize the same, although smaller, adjustment to the E-Z reader model parameters could possibly also model the L2 reading pattern of unbalanced bilinguals.

BILINGUAL L1 VS. MONOLINGUAL READING

The weaker links account predicts a drop in the strength of the links between all word forms and their representations in the bilingual lexicon because reading practice is divided across more (almost double the amount of) lexical items (Gollan & Acenas, 2004). Therefore, this account predicts slower silent reading for bilinguals. Although some studies (Lehtonen et al., 2012; Ransdell & Fischler, 1987) do report such a bilingual disadvantage for isolated word recognition, this was never investigated for language comprehension in a natural reading context when the target language proficiency was matched across the bilingual and monolingual group.

Contrary to predictions made by the weaker links account, we did not find a clear general disadvantage for bilinguals reading in their mother tongue compared to monolinguals. We did find a small bilingual disadvantage for fixation count per sentence. Bilinguals fixate slightly more often than monolinguals, but only in sentences with more than 23 words. Also, the amount of fixations that bilinguals made is more strongly determined by the average word length of the sentences than it is for monolinguals. Importantly, there is no interaction of word length or number of words with L1 proficiency. This means that these subtle differences are indeed caused by having a second language and not by a possibly reduced L1 language proficiency for bilinguals. Remember that this bilingual disadvantage does not show in the overall sentence reading time, while in production substantial bilingual slowing of reaction times was found (Ivanova & Costa, 2008) . In lexical decision tasks, the evidence is more mixed. Ransdell and Fischler (1987) found a significant disadvantage for bilinguals in their first acquired language compared to monolinguals. Duyck, Vanderelst, Desmet, and Hartsuiker (2008) did not find any difference in reaction times for Dutch-English bilinguals and English monolinguals. These bilinguals were taken from the same population, as the one tested in the current paper.

We want to point out that our design was very sensitive: we were able to detect significant differences of 0.2 fixations per sentence. This adds robustness to the observed null effect for bilingual L1 and monolingual sentence reading times. On top of that the bilinguals actually show a slightly faster sentence reading time than the monolinguals do (see Table 3) although this difference does not reach significance.

Gollan and Acenas (2004) assume a reduced integration between semantic and phonological codes in bilingual language production. It is very unlikely that similar weaker links for bilingual comprehension would not have an impact on lexical access and thus on fixation durations and reading times. From this, we could conclude that the weaker links theory does not provide a full picture of the underlying reason for the more subtle bilingual comprehension disadvantage we found in our unbalanced bilingual population with a late age of acquisition of L2.

Gollan et al. (2011) also predicted that the bilingual disadvantages would be smaller in comprehension than in production because the latter is less practiced, more difficult and involves more levels of processing where frequency is important. One has to consider that comprehension and production processes might be very distinct. In order to speak in one language (production), the speaker has to by definition make a language selection. In a picture-naming task, the picture needs to be named either in L1 or L2, and one of the two lexical representations needs to be inhibited, during each utterance. Such inhibition is not necessary in reading, in which bilinguals may rely on bottom-up information coming from the visual input to the lexical representation. Even if lexical representations from both languages become active, an actual language selection is not needed, and therefore recognition implies less inhibition than production does. Since some have proposed that distinct lexical forms serve comprehension and production (Roelofs, 2003; Zwitserlood, 2003) it is not improbable that being a bilingual would have a different impact on the representational strength of the lexical entities in comprehension than on those in production.

Finally, we want to stress that the participants in this study are different than those who are usually used in the studies reporting an L1 disadvantage for bilinguals (Gollan et al., 2008; Lehtonen et al., 2012). It is therefore completely possible that the weaker links account holds for bilinguals who indeed have less exposure to their L1 due to an increased exposure to L2. These are mostly balanced or early bilingual populations. As described in the introduction, we do not think however that the weaker links hypothesis necessarily holds for all bilingual populations. More specifically, late L2 learners, who have acquired full L1 proficiency before acquiring an L2 are likely to have a larger language exposure overall than monolinguals do, due to an active seeking of extra language exposure. Also, late bilinguals are more likely to have already developed a certain level of lexical entrenchment for words in L1, before acquiring the new L2 words. This makes the 'weakening' of links between L1 semantic representations and word forms rather unlikely.

In conclusion, our results show no evidence that unbalanced late bilinguals read slower in their L1 than monolinguals do. Any possibly subtler differences (e.g. fixation counts or other differences), which may emerge at a word level, are at least compensated elsewhere so that on a whole, unbalanced bilinguals do not show any disadvantages compared to monolinguals when reading in L1. These findings imply that at least for unbalanced bilinguals, no 'weaker' links have to be assumed to understand bilingual language processing. This is compatible with the notion that language comprehension and production might overlap only at the level of meaning (Roelofs, 2003; Zwitserlood, 2003), or at least are not completely shared or aligned (Gollan et al., 2011).

COMPARISON OF L1 DATA WITH META-ANALYSIS OF EYE MOVEMENTS

It is important to compare our L1 results with the other eye movement studies that have already reported sentence level measures, in order to establish to what extent such results are generalizable to different settings. When we compare our average reading parameters with a meta-analysis by Rayner (2009), mostly including findings from earlier sentence-embedded eye movement research, we observe some slight deviations: We found shorter average fixation durations, longer saccade lengths and a higher regression rate (See Table 4).

Table 4. *A comparison of the reading meta-analysis of Rayner (2009), based on sentence reading research, and our natural L1 reading data. In our analysis we use first pass skipping rate (52%), but in the table we report total skipping rate (41.5%).*

	Sentence Reading Rayner (2009)	Book Reading Our L1 data
Avg Fixation duration	225-250ms	215.8ms
Avg Saccade length	7-9 characters	9.9 characters
Regression Rate	10%-15 %	24.2%
Fixations per 100 words	75-118	72
Skipping probability		
Content words	15%	34.2%
Function words	65%	48.8%

These differences indicate that reading a continuous text or story is not the same as reading isolated sentences (Radach et al., 2008). Radach et al. (2008) found that the overall fixation duration of words is longer for reading passages, but the first pass measures are slower than when reading isolated sentences. Radach et al. explain this by suggesting that readers of passages of text perform a fast first pass across the text followed by a rereading of the passage. This is compatible with our findings, illustrated by lower average fixation durations, longer saccade lengths and more regressions compared to results from isolated or sentence embedded research (see Table 4). Additional evidence for this difference in reading strategy comes from analyses of natural reading data, that found regression rates (21% for adults, 36% for 10-11 year olds) and fixation durations (200ms for adults, 243ms for 10-11 year olds) that are similar to ours (McConkie et al., 1991). As inter-word regressions indicate integration difficulties, it is plausible that people reading individual sentences would have less need to move their eyes back in the text. This is compatible with Radach et al., who state that when reading continuous text, people reread the text after a fast first pass.

LIMITED EFFECTS OF PROFICIENCY AND WORD FREQUENCY

Our data suggests that the influence of proficiency on sentence level reading parameters is small. In fact, we only find a significant effect of L2 proficiency in the fixation count analysis of the L1 vs. L2 reading comparison. The sentence-level differences in eye movements between L1 and L2 reading are apparently not very sensitive to the L2 proficiency level of the bilinguals. Our bilingual participants were all L1-dominant, unbalanced bilinguals who nevertheless showed considerable variation in L2 composite proficiency scores [52.5%-86.8%]. Note that this range was large enough to yield an effect of L2 proficiency for fixation count. A 10-point increase in the L2 composite proficiency score yields about a decrease of 1.35 fixations per 100 L2 words. For example, a person scoring 65% on his/her L2 proficiency would fixate 92.9 times per 100 words, a person scoring 75% would then on average make 91.6 fixations per 100 words. When we look at the fitted value of the least L2 proficient bilingual scoring lowest on L2 we observe 94 fixations per 100 words, while the highest L2 proficient person has 90 fixations per sentence. When we use the L2 LexTALE scores instead of the L2 composite proficiency score we observe even smaller effects. Here, a 10-point increase in the L2 LexTALE score yields about a decrease of 0.51 fixations per 100 words. A person scoring 65% on his/her L2 proficiency would fixate 76.8 times per 100 words, and a person scoring 75% would make 76.3 fixations per 100 words. The difference between the highest scoring bilingual on the L2 LexTALE and the lowest scoring bilingual is only 0.3 fixations. Even though these effects are small, they are nevertheless detected, illustrating that proficiency just does not yield big effects on sentence reading measures, rather than that these null interaction effects are caused by a small range of L2 proficiency scores for the tested bilinguals. Our results suggest that the differences in eye movement pattern between L1 and L2 reading are more determined by the fact that the L2 is acquired after the L1, than merely by the L2 language proficiency. The absence of (strong) L2 proficiency interactions effects also supports the generalizability of these findings to other unbalanced bilingual populations with somewhat different L2 proficiency scores. Of course, for balanced bilinguals, a different pattern may emerge.

In our analyses, we find few effects of or interactions with word frequency. This is not surprising, given that word frequency measures affect early measures of language processing, like single fixation durations and first fixation durations (Hyönä & Olson, 1995) and have a smaller effect on natural reading than on reading of isolated words or sentence embedded target words (Kuperman et al., 2013; Radach et al., 2008). The low frequent words would be more easy to process in continuous text because of the context it provides to identify such a word (Kuperman et al., 2013). An additional reason for the absence of an influence of frequency is that the focus of this paper was on sentence reading parameters, and therefore we used an average frequency measure of only the content words in the sentence. This is likely to be a rather insensitive measure, and any frequency effects may be compensated by words on the other end of the scale. This hypothesis is confirmed in a separate study, where we have analyzed frequency effects in word-level eye movements of this corpus (Cop et al., 2015). Here we found clear effects of word frequency in L1, L2 and monolingual reading.

Another issue in our analyses is that we find some 3-way interactions, with rather small effect sizes, that we did not predict or expect. They may offer inspiration for future research that is aimed at more specific questions than the current paper, using smaller, controlled experiments, aimed specifically at that interaction effect.

FURTHER USE OF PARAMETERS/FINDINGS

From our analyses, it is clear that the eye movement behavior of bilinguals in L2 shows some similarities to the reading behavior of children. The most parsimonious explanation is that both patterns have the same underlying cause: slower lexical processing. This suggests that the L2 reading pattern could possibly be modeled in the E-Z reader model by changing the same parameter as Reichle et al. (2013) used in modeling the reading pattern of children. It would be interesting for further research to try and simulate the results of this corpus with the E-Z reader model.

The same pattern of changes from adult to child reading has been consistently found in German, Finnish and English. This is remarkable because English and Finnish are dissimilar languages (Seymour, Aro, & Erskine, 2003). If we draw our parallel even further, we might then assume that the differences that we found between L1 and L2 reading will be universal and consistent across different language pairs, given that the bilingual participants acquired L2 later than their L1 and are less proficient in their L2 than they are in their L1. There is additional evidence that these results will generalize to bilingual populations with other L1 languages than Dutch (Lemhöfer et al., 2008). It has been shown that although cross-language influences, like cognate status of words, exist, word recognition by bilinguals in L2 is mostly determined by within language factors, like frequency and word length (Lemhöfer et al., 2008).

CONCLUSION

In summary, we have analyzed the sentence level eye movement behavior of bilinguals reading in L1 and L2, and of monolinguals reading in their mother tongue. We find large differences between sentence reading in the dominant language (L1) and a later acquired language (L2). All of these differences can be paralleled to the reading pattern found in 7-11 year old children acquiring reading skills, although the differences between L1 and L2 reading are smaller than the differences between adult's and children's reading pattern. These changes are all compatible with the concept of a general slowing of the process of lexical access, in parallel with the modeling effort of Reichle et al. (2013).

We do not find clear disadvantages for bilinguals reading in L1 compared to monolinguals reading in their only language. This shows that the bilingual disadvantage found in language production tasks and comprehension tasks using isolated words as stimuli, is not universally present across all modalities of language use or all bilingual language users. This means that the weaker links account (Gollan et al., 2008), which did a good job accounting for the balanced bilingual disadvantage in production, might not apply for comprehension of continuous text or language processing of unbalanced late bilinguals.

We hope these findings will inspire future research targeted at specific effects reported here, and promote the method of eye tracking of natural language reading, parallel to continued isolated word recognition research.

REFERENCES

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. Memory & Cognition, 24(4), 477–492.

Balling, L. W. (2013). Reading authentic texts: What counts as cognate? Bilingualism: Language and Cognition, 16(03), 637–653. http://doi.org/10.1017/S1366728911000733

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. Cognitive Psychology, 17(3), 364–390. http://doi.org/10.1016/0010-0285(85)90013-1

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language, 68(3), 255–278.

Binder, K. S., Duffy, S. A., & Rayner, K. (2001). The Effects of Thematic Fit and Discourse Context on Syntactic Ambiguity Resolution. Journal of Memory and Language, 44(2), 297–324. http://doi.org/10.1006/jmla.2000.2754

Blythe, H. I., Häikiö, T., Bertam, R., Liversedge, S. P., & Hyönä, J. (2011). Reading disappearing text: Why do children refixate words? Vision Research, 51(1), 84–92. http://doi.org/10.1016/j.visres.2010.10.003

Blythe, H. I., & Joseph, H. S. S. L. (2011). Children's Eye Movements during Reading. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.), Oxford Handbook On Eye Movements (pp. 643–662). Oxford, England: Oxford University Press.

Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2009). Visual information capture during fixations in reading for children and adults. Vision Research, 49(12), 1583–1591. http://doi.org/10.1016/j.visres.2009.03.015

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. Journal of the Royal Statistical Society, 26(2), 211–252.

Brenders, P., van Hell, J. G., & Dijkstra, T. (2011). Word recognition in child second language learners: Evidence from cognates and false friends. Journal of Experimental Child Psychology, 109(4), 383–396. http://doi.org/10.1016/j.jecp.2011.03.012

Brysbaert, M., & Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? Bilingualism: Language and Cognition, 13(03), 359–371. http://doi.org/10.1017/S1366728909990344

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. Behavior Research Methods, 41(4), 977–990. http://doi.org/10.3758/BRM.41.4.977

Brysbaert, M., & Vitu, F. (1998). Word skipping: Implications for theories of eye movement control in reading. Eye Guidance in Reading and Scene Perception, 125–147.

Bultena, S., Dijkstra, T., & van Hell, J. G. (2013). Cognate and word class ambiguity effects in noun and verb processing. Language and Cognitive Processes, 28(9), 1350–1377. http://doi.org/10.1080/01690965.2012.718353

Bultena, S., Dijkstra, T., & van Hell, J. G. (2014). Cognate effects in sentence context depend on word class, L2 proficiency, and task. The Quarterly Journal of Experimental Psychology, 67(6), 1214–1241. http://doi.org/10.1080/17470218.2013.853090

Buswell, G. T. (1922). Fundamental reading habits: a study of their development. University of California Libraries.

Caramazza, A., & Brones, I. (1979). Lexical Access in Bilinguals. Bulletin of the Psychonomic Society, 13(4), 212–214.

Cop, U., Keuleers, E., Drieghe, D., & Duyck, W. (2015). Frequency Effects in Monolingual and Bilingual Natural Reading. Psychonomic Bulletin and Review. http://doi.org/10.3758/s13423-015-0819-2 i

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. Journal of Experimental Psychology: Learning, Memory, and Cognition, 26(5), 1283–1296. http://doi.org/10.1037//0278-7393.26.5.1283

Cover, T. M., & Thomas, J. A. (1991). Elements of information theory. New York: Wiley.

Cristoffanini, P., Kirsner, K., & Milech, D. (1986). Bilingual lexical representation: The status of Spanish-English cognates. The Quarterly Journal of Experimental Psychology, 38, 367–393.

Davis, C., SáNchez-Casas, R., GarcíA-Albea, J. E., Guasch, M., Molero, M., & Ferré, P. (2010). Masked translation priming: Varying language experience and word type with Spanish–English bilinguals. Bilingualism: Language and Cognition, 13(02), 137. http://doi.org/10.1017/S1366728909990393

Dempster, F. N. (1985). Short-Term Memory Development in Childhood and Adolescence. In C. J. Brainerd & M. Pressley (Eds.), Basic Processes in Memory Development (pp. 209–248). Springer New York. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4613-9541-6_5

De Pessemier, P., & Andries, C. (2009). GL&SCHR. Leuven/Apeldoorn: Garant.

Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of Cognates and Interlingual Homographs: The Neglected Role of Phonology. Journal of Memory and Language, 41(4), 496–518.

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. Journal of Memory and Language, 62(3), 284–301. http://doi.org/10.1016/j.jml.2009.12.003

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. Bilingualism: Language and Cognition, 5(03). http://doi.org/10.1017/S1366728902003012

Dussias, P. E. (2010). Uses of Eye-Tracking Data in Second Language Sentence Processing Research. Annual Review of Applied Linguistics, 30, 149–166. http://doi.org/10.1017/S026719051000005X

Dussias, P. E., & Cramer Scaltz, T. R. (2008). Spanish–English L2 speakers' use of subcategorization bias information in the resolution of temporary ambiguity during second language reading. Acta Psychologica, 128(3), 501–513. http://doi.org/10.1016/j.actpsy.2007.09.004

Dussias, P. E., & Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish-English bilinguals. Bilingualism: Language and

Cognition, 10(01), 101–116. http://doi.org/10.1017/S1366728906002847

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. Journal of Experimental Psychology: Learning, Memory, and Cognition, 33(4), 663–679. http://doi.org/10.1037/0278-7393.33.4.663

Duyck, W., Vanderelst, D., Desmet, T., & Hartsuiker, R. J. (2008). The frequency effect in second-language visual word recognition. Psychonomic Bulletin & Review, 15(4), 850–855. http://doi.org/10.3758/PBR.15.4.850

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A Dynamical Model of Saccade Generation During Reading. Psychological Review, 112(4), 777–813. http://doi.org/10.1037/0033-295X.112.4.777

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), 246–269. http://doi.org/10.1037/0278-7393.30.1.246

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew–English bilinguals. Journal of Experimental Psychology: Learning, Memory, and Cognition, 23(5), 1122.

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. Journal of Memory and Language, 58(3), 787–814. http://doi.org/10.1016/j.jml.2007.07.001

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). http://doi.org/10.1017/S136672890100013X

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. Journal of Experimental Psychology: General, 140(2), 186–209. http://doi.org/10.1037/a0022256

Grosjean, F. (1982). Preface. In Life with two languages: An introduction to Bilingualism. (p. viii). Cambridge, MA: Harvard University Press.

Gullifer, J. W., Kroll, J. F., & Dussias, P. E. (2013). When Language Switching has No Apparent Cost: Lexical Access in Sentence Context. Frontiers in Psychology, 4. http://doi.org/10.3389/fpsyg.2013.00278

Häikiö, T., Bertram, R., Hyönä, J., & Niemi, P. (2009). Development of the letter identity span in reading: Evidence from the eye movement moving window paradigm. Journal of Experimental Child Psychology, 102(2), 167–181. http://doi.org/10.1016/j.jecp.2008.04.002

Huestegge, L., Radach, R., Corbic, D., & Huestegge, S. M. (2009). Oculomotor and linguistic determinants of reading development: A longitudinal study. Vision Research, 49(24), 2948–2959. http://doi.org/10.1016/j.visres.2009.09.012

Hyönä, J., & Olson, R. K. (1995). Eye fixation patterns among dyslexic and normal readers: effects of word length and word frequency. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21(6), 1430.

Inhoff, A. W., Radach, R., Eiter, B. M., & Juhasz, B. (2003). Distinct subsystems for the parafoveal processing of spatial and linguistic information during eye fixations in reading. The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology, 56(5), 803–827. http://doi.org/10.1080/02724980244000639

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? Acta Psychologica, 127(2), 277–288. http://doi.org/10.1016/j.actpsy.2007.06.003

Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., & Rayner, K. (2009). Word length and landing position effects during reading in children and adults. Vision Research, 49(16), 2078–2086. http://doi.org/10.1016/j.visres.2009.05.015

Just, M. ., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. Psychological Review, 99(1), 122.

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. Vision Research, 45(2), 153–168. http://doi.org/10.1016/j.visres.2004.07.037

Kenward, M. G., & Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. Biometrics, 53(3), 983–997.

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. Behavior Research Methods, 42(3), 643–650. http://doi.org/10.3758/BRM.42.3.643

Kliegl, R., & Engbert, R. (2005). Fixation durations before word skipping in reading. Psychonomic Bulletin & Review, 12(1), 132–138.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. European Journal of Cognitive Psychology, 16(1-2), 262–284. http://doi.org/10.1080/09541440340000213

Kroll, J. F., Dijkstra, T., Janssens, N., & Schriefers, H. (1999). Cross-language lexical activity during production: Evidence from cued picture naming. In A. Vandierendonck, M. Brysbaert, & K. Van der Goten (Eds.), Proceedings of the 11th congress of the European society for cognitive psychology (p. 92). Ghent: ESCOP/Academic Press.

Kuperman, V., Drieghe, D., Keuleers, E., & Brysbaert, M. (2013). How strongly do word reading times and lexical decision times correlate? Combining data from eye movement corpora and megastudies. The Quarterly Journal of Experimental Psychology, 66(3), 563–580. http://doi.org/10.1080/17470218.2012.658820

Lehtonen, M., Hultén, A., Rodríguez-Fornells, A., Cunillera, T., Tuomainen, J., & Laine, M. (2012). Differences in word recognition between early bilinguals and monolinguals: behavioral and ERP evidence. Neuropsychologia, 50(7), 1362–1371. http://doi.org/10.1016/j.neuropsychologia.2012.02.021

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. Behavior

Research Methods, 44(2), 325–343. http://doi.org/10.3758/s13428-011-0146-0
Lemhöfer, K., Dijkstra, T., Schriefers, H., Baayen, R. H., Grainger, J., & Zwitserlood, P. (2008). Native language influences on word recognition in a second language: A megastudy. Journal of Experimental Psychology: Learning, Memory, and Cognition, 34(1), 12–31. http://doi.org/10.1037/0278-7393.34.1.12
Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. Journal of Experimental Psychology: Learning, Memory, and Cognition, 35(2), 381–390. http://doi.org/10.1037/a0014875
Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. Trends in Cognitive Sciences, 4(1), 6–14.
McClelland, J., & Rumelhart, D. (1981). An interactive activation model of context effects in letter perception. An account of basic findings. Psychological Review, 88, 375–407.
McConkie, G. W., & Rayner, K. (1975). The Span of Effective stimulus during a fixation in reading. Perception & Psychophysics, 17, 578–586.
McConkie, G. W., Zola, D., Grimes, J., Kerr, P. W., Bryant, N. R., & Wolff, P. M. (1991). Children's eye movements during reading. In J. F. Stein (Ed.), Vision and Visual Dyslexia (pp. 251–262). Boston, US: CRC Press.
Nation, K. (2009). Form–meaning links in the development of visual word recognition. Philosophical Transactions of the Royal Society B: Biological Sciences, 364(1536), 3665–3674. http://doi.org/10.1098/rstb.2009.0119
Peeters, D., Dijkstra, T., & Grainger, J. (2013). The representation and processing of identical cognates by late bilinguals: RT and ERP effects. Journal of Memory and Language, 68(4), 315–332. http://doi.org/10.1016/j.jml.2012.12.003
Pellegrino, F., Coupé, C., & Marsico, E. (2011). Across-Language Perspective on Speech Information Rate. Language, 87(3), 539–558. http://doi.org/10.1353/lan.2011.0057
Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), Reading

acquisition (pp. 145–174). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Perfetti, C. A., Wlotko, E. W., & Hart, L. A. (2005). Word Learning and Individual Differences in Word Learning Reflected in Event-Related Potentials. Journal of Experimental Psychology: Learning, Memory, and Cognition, 31(6), 1281–1292. http://doi.org/10.1037/0278-7393.31.6.1281

Pivneva, I., Mercier, J., & Titone, D. (2014). Executive Control Modulates Cross-Language Lexical Activation During L2 Reading: Evidence From Eye Movements. Journal of Experimental Psychology-Learning Memory and Cognition, 40(3), 787–796. http://doi.org/10.1037/a0035583

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. Cognitive Psychology, 52(1), 1–56. http://doi.org/10.1016/j.cogpsych.2005.06.001

Poulisse, N., & Bongaerts, T. (1994). First Language Use in Second Language Production. Applied Linguistics, 15(1), 36–57. http://doi.org/10.1093/applin/15.1.36

Radach, R., Huestegge, L., & Reilly, R. (2008). The role of global top-down factors in local eye-movement control in reading. Psychological Research, 72(6), 675–688. http://doi.org/10.1007/s00426-008-0173-3

Ransdell, S. E., & Fischler, I. (1987). Memory in a monolingual mode: When are bilinguals at a disadvantage? Journal of Memory and Language, 26(4), 392–405.

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. Journal of Experimental Child Psychology, 41(2), 211–236.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. Psychological Bulletin, 124(3), 372–422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. The Quarterly Journal of Experimental Psychology, 62(8), 1457–1506. http://doi.org/10.1080/17470210902816461

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The Effects of Frequency and Predictability on Eye Fixations in Reading:

Implications for the E-Z Reader Model. Journal of Experimental Psychology: Human Perception and Performance, 30(4), 720–732. http://doi.org/10.1037/0096-1523.30.4.720

Rayner, K., Li, X., & Pollatsek, A. (2007). Extending the E-Z Reader Model of Eye Movement Control to Chinese Readers. Cognitive Science, 31(6), 1021–1033. http://doi.org/10.1080/03640210701703824

Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. Psychology and Aging, 21(3), 448–465. http://doi.org/10.1037/0882-7974.21.3.448

R Core Team. (2014). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Reichle, E. D. (2011). Serial Attention Models of Reading. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.), Oxford Handbook On Eye Movements (pp. 767–786). Oxford, England: Oxford University Press.

Reichle, E. D., Liversedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2013). Using E-Z Reader to examine the concurrent development of eye-movement control and reading skill. Developmental Review, 33(2), 110–149. http://doi.org/10.1016/j.dr.2013.03.001

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. Psychological Review, 105(1), 125.

Reichle, E. D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: Accounting for initial fixation locations and refixations within the EZ Reader model. Vision Research, 39(26), 4403–4411.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The EZ Reader model of eye-movement control in reading: Comparisons to other models. Behavioral and Brain Sciences, 26(04), 445–476.

Reichle, E. D., Tokowicz, N., Liu, Y., & Perfetti, C. A. (2011). Testing an assumption of the E-Z Reader model of eye-movement control during reading: Using event-related potentials to examine the familiarity check: E-Z Reader and event-related potentials.

Psychophysiology, 48(7), 993–1003. http://doi.org/10.1111/j.1469-8986.2011.01169.x

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher level language processing on eye movements during reading. Psychonomic Bulletin & Review, 16(1), 1–21. http://doi.org/10.3758/PBR.16.1.1

Roelofs, A. (2003). Modeling the relation between the production and recognition of spoken word forms. Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities, 115–158.

Sanchez-Casas, R., Davis, C., & García-Albea, J. E. (1992). Bilingual Lexical Processing -Exploring the cognate non-cognate distinction. European Journal of Cognitive Psychology, 4(4), 293–310.

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. Journal of Memory and Language, 55(2), 197–212. http://doi.org/10.1016/j.jml.2006.03.004

Seymour, P. H., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. British Journal of Psychology, 94(2), 143–174.

Siegel, L. S. (1994). Working Memory and Reading: A Life-span Perspective. International Journal of Behavioral Development, 17(1), 109–124. http://doi.org/10.1177/016502549401700107

Taylor, S. E. (1965). Eye Movements in Reading: Facts and Fallacies. American Educational Research Journal, 2(4), 187–202. http://doi.org/10.2307/1161646

Titone, D., Libben, M., Mercier, J., Whitford, V., & Pivneva, I. (2011). Bilingual lexical access during L1 sentence reading: The effects of L2 knowledge, semantic constraint, and L1–L2 intermixing. Journal of Experimental Psychology: Learning, Memory, and Cognition, 37(6), 1412–1431. http://doi.org/10.1037/a0024492

Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. Journal of Memory and Language, 64(1), 88–107. http://doi.org/10.1016/j.jml.2010.08.006

Van Assche, E., Duyck, W., & Hartsuiker, R. J. (2012). Bilingual Word Recognition in a Sentence Context. Frontiers in Psychology, 3. http://doi.org/10.3389/fpsyg.2012.00174

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. Psychological Science, 20(8), 923–927.

Van Hell, J. G., & de Groot, A. M. B. (2008). Sentence context modulates visual word recognition and translation in bilinguals. Acta Psychologica, 128(3), 431–451. http://doi.org/10.1016/j.actpsy.2008.03.010

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. Psychonomic Bulletin & Review, 9(4), 780–789.

Van Hell, J. G., & Tanner, D. (2012). Second Language Proficiency and Cross-Language Lexical Activation. Language Learning, 62(s2), 148–171.

Whitford, V., & Titone, D. (2012). Second-language experience modulates first- and second-language word frequency effects: Evidence from eye movement measures of natural paragraph reading. Psychonomic Bulletin & Review, 19(1), 73–80. http://doi.org/10.3758/s13423-011-0179-5

Wilkinson, G. S., & Robertson, G. . (2006). Wide Range Achievement Test 4 professional manual. Lutz, FL: Psychological Assessment Resources.

Winskel, H., Radach, R., & Luksaneeyanawin, S. (2009). Eye movements when reading spaced and unspaced Thai and English: A comparison of Thai–English bilinguals and English monolinguals. Journal of Memory and Language, 61(3), 339–351. http://doi.org/10.1016/j.jml.2009.07.002

Zwitserlood, P. (2003). The Internal Structure of Words: Consequences for listening and speaking. In N. O. Schiller & A. S. Meyer (Eds.), Phonetic and Phonology in Language Comprehension and Production: Differences and Similarities (pp. 79–114). Berlin, Germany: Mouton de Gruyter.

APPENDIX A: ADDITIONAL TABLES

Table A.1. *Count of bilingual participants agreeing and not agreeing on second language skills items.*

Skills	Agree	Don't Agree
Carry on normal conversation in L2	19	0
Watch television shows in L2	19	0
Listen to music in L2	19	0
Read and comprehend questions in L2	19	0
Read books or articles in L2	19	0
No problems in understanding L1 speaker	18	1
Carry on a discussion in L2	17	2
Love speaking L2	16	3
Explain difficult situation in L2	15	4
Answer difficult questions in L2	12	7
Think in L2	11	8
Speak to myself in L2	10	9
Write in L2	8	11
Make no/ almost no mistakes in L2	6	13
Dream in L2	5	14

Table A.2. *Count of bilingual participants agreeing and not agreeing on second language switching items.*

Switching	Agree	Don't Agree
I'm sometimes in a tip of the tongue state	16	3
I sometimes can't get the right word	14	5
I use a different language when I do not remember a word	13	6
I often use different languages intermixed	9	10
I often use different languages intermixed without noticing	5	14
I sometimes speak in a language that my dialogue partner doesn't understand	5	14

APPENDIX B: MULTIPLE-CHOICE QUESTIONS

English multiple-choice questions with possible answers that were used to test text comprehension. The correct answer is indicated in bold.

1. Who does Mr. Hastings meet for the first time in the front garden when Mr. Hastings arrives at Styles Court?
a) John Cavendish
b) Mary Cavendish
c) Evelyne Howard
d) Emily Inglethorpe

2. The book is narrated by Mr. Hastings. Of what secret ambition does he tell Mary Cavendish?
a) Becoming a biographer.
b) Becoming a police officer.
c) Becoming a detective.
d) Becoming a doctor in medicine .

3. After Evelyne Howard leaves Styles Court because of a discussion with Emily Inglethorpe she goes to live in Middlingham. What position does she find there?
a) nanny
b) housekeeper
c) nurse
d) pharmacist
4. What is the nickname of the colleague of Cynthia Murdoch?
a) Nibs
b) Buns
c) Barny
d) Snug

5. Who wakes Mr. Hastings on the night of the murder?
a) Nobody. He wakes up because of the noises.
b) Lawrence Cavendish.
c) John Cavendish.
d) Dorcas.

6. How does Alfred Inglethorpe explain that he did not sleep in his bed on the night of the murder?
a) **He forgot his house key.**
b) He didn't want to wake his wife.
c) Friends of his needed his help.
d) He was drunk and slept with friends.

7. Why is the fact whether Emily Inglethorpe ate on the evening of her death of importance for the investigation?

a) Her appetite says a lot about her emotional state.
b) A big dinner could have slowed down the effect of the poison.
c) A big dinner could have aggravated the effect of the poison.
d) It is not important to the investigation.

8. Which of the following statements is not correct? When Poirot investigates the bedroom of Emily Inglethorp for the first time he

a) notices a coffee cup stamped to pieces.

b) discovers a wet spot on the carpet.

c) takes the documents out of the purple case.

d) discovers the remains of a will in the ashes of the fireplace.

9. The sixth point of importance is

a) a piece of green fabric Poirot finds.
b) That Dorcas overheared the argument between Mrs. and Mr. Inglethorpe
c) the spare key to the purple box was missing
d). An empty box of sleeping powders

10. What did Annie notice was strange about the cocoa?

a) It was not heated yet.

b) there was some kitchen salt on the tray.

c) the tray was displaced by somebody else.

d) the cocoa was brought up later than usual.

11. How many coffee cups does Poirot count in the drawing room?

a) 3
b) 5
c) 6
d) 7

12. What is true about the last intact will of Mrs. Inglethorpe?

a) She left her entire fortune to John Cavendish.
b) She left her entire fortune to Alfred Inglethorp.
c) She left her entire fortune to Evelyne Howard.
d) She left her entire fortune to Lawrence Cavendish.

13. What does Lawrence think is the cause of dead of his stepmother?

a) She was poisoned.
b) Tetanus.
c) Heart failure.
d) Old age.

14. Who admits to not being a good friend to Mr. Hastings?

a) John Cavendish.
b) Hercule Poirot.
c) Lawrence Cavendish.
d) Mary Cavendish.

15. What element does Poirot thinks is implicating Alfred Inglethorpe in the murder?

a) that he was not in the house at the moment of the murder.
b) that the door of Mrs. Inglethorpes bedroom was closed from the inside so she must have opened the door for her husband.
c) that everybody else suspects him.
d) that Alfred Inglethorpe had an argument with Emily Inglethorpe on the day of her murder.

16. Two people are observed to shed tears for the murdered Emily Inglethorpe. Who are they?

a) Evelyne Howard and Hercule Poirot
b) Evelyne Howard and Dorcas
c) Evelyne Howard and Mary Cavendish
d) Evelyne Howard and Cynthia

17. Why is it impossible according to Dr. Wilkes that Emily Inglethorpe was poisoned by her medicine?

a) Because it did not contain any poisonous substances.
b) Because Emily Inglethorpe was too smart to take an overdose.
c) Because the dose of strychnine in the medicine is too small to poison somebody.
d) Because Emily Inglethorpe did not take the medicine that night.

18. How does Alfred Inglethorpe explain the last words of his wife?

a) She was accusing him of her death.
b) She was trying to tell him something.
c) She wrongfully thought that Dr. Bauerstein was Mr. Inglethorpe.
d) She was talking nonsense.

19. What two things are of primary significance to Poirot about the day of the murder?

a) The temperature and the green fabric he found.
b) The temperature and the distinctive appearance of Alfred Inglethorpe.
c) The temperature and the broken cup
d) The broken cup and the distinctive appearance of Inglethorpe

20. Why is Poirot determined that Alfred Inglethorpe shall not be arrested for the murder of Emily Inglethorpe?

a) Alfred will be found innocent.
b) Alfred is innocent.
c) There is not enough evidence to arrest Mr. Inglethorpe.
d) The real murderer will escape.

21. What is the reason that Alfred Inglethorp could not give a believable alibi for the time the strychnine was bought in the pharmacy?

a) He was scared to admit he had been with Mrs. Raikes.
b) He does not have an alibi.
c) He was taking a walk and nobody could confirm this story.
d) He wanted to be arrested.

22. Who inherits ‘Styles Court’?

a) Alfred Inglethorpe.
b) John Cavendish.
c) Lawrence Cavendish.
d) Nobody.

23. Which statement is true?

a) Annie shows Poirot where to find the dressing-up trunk
b) Poirot finds a green dress and a fake beard in the dressing-up trunk.
c) Poirot finds a green dress in the dressing-up trunk.
d) Poirot finds a fake beard when looking for a green dress.

24. Which person does Evelyne Howard’s intuition tells her commited the murder?

a) Cynthia Murdoch
b) Alfred Inglethorp
c) She does not say who.
d) Mr. Hastings

25. Where does Mr. Hastings hear Mary and John Cavendish arguing?

a) the park

b) their bedroom

c) in the village

d) in the drawing room

26. When Cynthia Murdoch confides in Mr. Hastings, she tells him that certain people in the household hate her. Who is she talking about?

a) John and Lawrence Cavendish

b) Emily Howard en Mary Cavendish

c) Lawrence Cavendish and Emily Howard

d) Lawrence and Mary Cavendish

27. The mysterious Dr. Bauerstein, although he does not turn out to be the murderer, nonetheless does turn out to be a criminal. What is his crime?

a) Espionage
b) Burglary
c) Embezzlement
d) Blackmail

28. Which of the following is true of Mary Cavendish?

a) Her first husband died in prison.
b) Her father died under mysterious circumstances.
c) Her mother was Russian.
d) Her father was shot for being a traitor.

29. The fourth letter that Mrs. Inglethorpe had sent on the day before the murder was addressed to…

a) Evelyne Howard.
b) Alfred Inglethorp.
c) A French music publisher.
d) A nurse in Middlingham.

30. Whose fingerprints were found on the bottle of strychnine in the pharmacy?

a) Lawrence Cavendish.
b) Lawrence Cavendish and Cynthia Murdoch.
c) Lawrence Cavendish and John Cavendish.
d) Cynthia Murdoch.

31. In the case against John Cavendish, what contention of the prosecution will Poirot be able to refute?

a) That it was John who bought the strychnine in the pharmacy.
b) That John would benefit from the death of Emily Inglethorpe.
c) That John brought the coffee to Mrs. Inglethorpe's room.
d) **That John was the one who burned the will.**

32. In the court case against John Cavendish, whom else is the prosecution trying to implicate in the murder?

a) Mary

b) Evelyne

c) **Lawrence**

d) Hastings

33. How does the defense explain the bottle of poison in John Cavendish's room?

a) Somebody is trying to frame John Cavendish
b) He used this to poison a stray dog.
c) John hid this bottle for somebody else he is trying to protect.
d) There was no bottle of poison in John's room.

34. Which of the following statements is false?

a) John Cavendish claims that he does not remember the exact words that were used by his step mother in their discussion.
b) The prosecution believes that John bought the strychnine in the pharmacy.
c) There is a consensus in the court room that the handwriting on the poison list in the pharmacy is not Alfred Inglethorp's.
d) Lawrence denies that he touched the bottle of strychnine in the pharmacy.

35. In addition to the strychnine, Emily Inglethorpe (along with Cynthia) was given a mild narcotic. Who drugged Emily Inglethorpe and Cynthia Murdoch?

a) Dr. Bauerstein
b) Mary Cavendish
c) John Cavendish
d) Evelyne Howard

36. How was Emily Inglethorpe poisoned?

a) Bromide was added to her medication.
b) Strychnine was added to her coffee.
c) Strychnine was added to the bromide powders.

d) Strychnine was added to her medication.

37. Which of the following is true of Evelyn Howard and Alfred Inglethorp's relationship?

a) They hate each other intensely.
b) They are trying to frame each other for murder.
c) They suspect each other of committing the murder.
d) They are cousins.

38. Whom does Evelyn Howard attempt to implicate in the murder?

a) Mary Cavendish
b) Cynthia Murdoch
c) Alfred Inglethorpe
d) John Cavendish

39. How does Poirot realize that there is a letter in the vase?

a) There was one vase more on the mantle the first time he entered the room.
b) The vases were empty first.
c) He had to straighten the objects on the mantle twice.
d) The objects on the mantle were straight while they were crooked first.

40. Why did Poirot want John Cavendish to go on trial for the murder of Emily Inglethorpe?

a) He believed John was guilty.
b) John was interfering with the investigation.
c) He believed it would cause the real killer to confess.
d) He thought it would bring him and his wife closer together.

APPENDIX C: LINK TO DATA FILES

url: http://expsy.ugent.be/downloads/uschi2/

login: uschi

password: pp02

CHAPTER 4
FREQUENCY EFFECTS IN MONOLINGUAL AND BILINGUAL NATURAL READING[6]

This paper presents the first systematic examination of the monolingual and bilingual frequency effect (FE) during natural reading. We analyzed single fixations durations on content words for participants reading an entire novel. Unbalanced bilinguals and monolinguals show a similarly sized FE in their mother tongue (L1), but for bilinguals the FE is considerably larger in their second language (L2) than in their L1. The FE in both L1 and L2 reading decreased with increasing L1 proficiency, but it was not affected by L2 proficiency. Our results are consistent with an account of bilingual language processing that assumes an integrated mental lexicon with exposure as the main determiner for lexical entrenchment (Diependaele, Lemhöfer, & Brysbaert, 2013; Gollan et al., 2008). This means that no qualitative difference in language processing between monolingual, bilingual L1 or bilingual L2 is necessary to explain reading behavior. We specify this account and argue that not all groups of bilinguals necessarily have lower L1 exposure than monolinguals do and, in line with Kuperman and Van Dyke (2013), that individual vocabulary size and language exposure change the accuracy of the relative corpus word frequencies and thereby determine the size of the FE's in the same way for all participants.

[6] This chapter is based on the manuscript: Cop, U., Keuleers, E., Drieghe, D., & Duyck, W. (2015) Frequency Effects in Monolingual and Bilingual Natural Reading. *Psychonomic Bulletin and Review, in press*

INTRODUCTION

Although word recognition and production are both very complex processes influenced by a wide range of variables, the frequency of occurrence of a word in a language is by far the most robust predictor of language performance (Brysbaert et al., 2011; Murray & Forster, 2004). In both word identification (e.g. Rubenstein, Garfield, & Millikan, 1970; Scarborough, Cortese, & Scarborough, 1977) and word production tasks (e.g. Forster & Chambers, 1973; Monsell, Doyle, & Haggard, 1989) high frequency words are processed faster than low frequency words. This observation is called the word frequency effect (FE), and it is one of the most investigated phenomena in (monolingual) psycholinguistics.

Multiple language models of comprehension (e.g. Dijkstra & Van Heuven, 2002; McClelland & Rumelhart, 1981; Morton, 1970) explain frequency effects using implicit learning accounts. These state that repeated exposure to a certain lexical item raises this item's baseline activation in proportion to their distance to the activation threshold, so that lexical selection of that particular word is faster during recognition (e.g. Monsell, 1991). The maximal speed of lexical access is limited, so once a word has received a certain amount of exposure, no more facilitation will be expected when there is additional exposure to that particular item (Morton 1970).

In the visual domain, word recognition speed increases with the logarithm of word corpus frequency (Howes & Solomon, 1951). A certain number of additional exposures to a low frequency word will result in a large decrease of its lexical access time, while the same number of additional exposures to a high frequency word will result in a much smaller decrease of its lexical access time. This particular characteristic of the relationship between word frequency and processing time causes the size of the frequency effect to be modulated by language exposure.

Bilinguals offer an interesting opportunity to study the relationship between exposure and lexical access, because of the within-subject difference in language exposure for L1 and L2. We will examine the effect of word frequency in bilingualism on the basis of new natural reading data collected for English monolinguals and Dutch-English bilinguals. We will start by examining the literature on individual differences in the word frequency effect and discuss the relation of these findings to the frequency effect in bilinguals. Following Kuperman and Van Dyke (2013), we will formulate an account of exposure-related differences in the effect of corpus word frequency that originates in the statistical characteristics of word frequency distributions.

INDIVIDUAL DIFFERENCES IN THE FE

The collection and evaluation of frequency norms based on text corpora is central to psycholinguistic research (e.g., Brysbaert & New, 2009; Keuleers, Brysbaert & New, 2010; Van Heuven, Mandera, Keuleers, & Brysbaert, 2014). The number of exposures to a certain word is often operationalized as the count of word occurrences in language corpora like the Subtlex database (Keuleers et al., 2010). Mostly, corpus frequencies are expressed as *relative* values because these can be used independent of corpus size. These *objective* corpus word frequencies are supposed to reflect the average number of exposures to certain words of an experienced reader. While corpus word frequencies are a tremendously useful proxy measure for relative exposure, it should not be forgotten that the relative frequency of a word in a text corpus is not necessarily equal to the relative frequency of exposure to that word for a particular individual.

Solomon and Howes (1951) already emphasized that word counts from text corpora are based on an arbitrary sample of the language and that there may be individual variation in the relative frequency of exposure to specific words. In other words, corpus word frequencies may under- or overestimate subjective word frequencies, which can lead to a difference in the size of the FE when corpus word frequencies are used in analyses. The differences in the FE size would disappear when a measure of actual exposure or subjective frequency (e.g., Connine, Mullennix, Shernoff, & Yelen, 1990; Gernsbacher, 1984) is used. Still, in experiments where words from different semantic domains (for example tools or clothing) are used as stimuli, such differences in relative frequency would in principle not lead to systematic differences in the frequency effect between individuals. This is because differences in subjective frequency in particular semantic categories would be cancelled out by the use of stimuli from multiple domains.

Next to the possibility of individual differences in the relative frequency for *specific* words due to differences in experience with a specific vocabulary, it is possible that individuals, who are at different stages in the language acquisition process, or, more broadly, have a differing amount of total language exposure, may have different relative frequencies for words. For this reason, some studies have used familiarity ratings of words as a more accurate reflection of the actual exposure to certain words for a specific group of readers (e.g. Balota, Pilotti, & Cortese, 2001; Kuperman & Van Dyke, 2013). Balota et al. (2001) observed that these subjective norms explained unique variance above and beyond objective corpus frequencies for lexical decision and naming tasks. Kuperman and Van Dyke (2013) confirm that objective corpus frequencies are particularly poor estimates and systematically overestimate the subjective frequencies for low frequent words for individuals with smaller vocabularies.

BILINGUAL FE'S

Most research on the frequency effect in language processing has focused on monolingual participants, while more than half of the world population, the 'default' person, is bilingual or multi-lingual. Taking into account that bi- or multilingualism is at least as widespread as monolingualism, it is important to assess how exposure to L1 or L2 affects bilingual person language processing. This is not straightforward because there is now a consensus that L1 and L2 constantly interact during visual word recognition (e.g. Duyck, Van Assche, Drieghe, & Hartsuiker, 2007; Van Assche, Duyck & Hartsuiker, 2012). These cross-lingual interactions strongly suggest the existence of a unified bilingual lexicon with parallel activation for all items in that lexicon, with items competing for selection within and across languages (for a more comprehensive overview of the evidence for an integrated bilingual lexicon see Brysbaert & Duyck, 2010 and Dijkstra & Vanheuven, 2002). Not only does L1 knowledge influence L2 lexical access, but the knowledge of an L2 also changes L1 visual word recognition (e.g. Van Assche, Duyck, Hartsuiker & Diependaele, 2009). Because these interactions occur in both directions, it is not only important to assess the differential influence of word exposure on lexical access for L1 and L2 reading, but also the possible differences between the frequency effect for monolinguals and bilinguals in L1.

Although the individual differences in frequency distribution described above are relevant for monolingual research, this is even more the case for bilingual research. The integrated bilingual lexicon will contain on average more lexical items than that of a monolingual. For advanced learners of an L2, who have a lexical entry for almost all concepts, we can assume that they would have almost double the amount of words in their lexicon. Inspired by observations of bilingual disadvantages in production tasks (e.g. Ivanova & Costa, 2008; Gollan, Montoya, Fennema-Notestine & Morris, 2005, Gollan et al., 2011), the weaker links theory (Gollan & Silverberg, 2001;Gollan & Acenas, 2004; Gollan et al. 2008, 2011) was proposed. This theory posits the idea that bilinguals necessarily divide their language use across two languages, resulting in lower exposure to all of the words in their lexicon, including L1 words. The lexical representations of bilinguals in both languages will have accumulated less exposure than the ones in the monolingual lexicon. Over time, this pattern of use would lead to weaker links between semantics and phonology for bilinguals, relative to monolinguals (Gollan et al. 2008).

Diependaele et al. (2013) generalize the weaker links account and assume a decrease in lexical exposure for bilinguals, and suggest that this can result in a reduced lexical entrenchment either by reduced lexical precision of those representations (e.g. Perfetti, 1992, 2007), or by reduced word-word inhibition or weaker integration between phonological and semantic codes (e.g. Gollan et al., 2008, 2011).

In short, the mere knowledge of a second language (and being exposed to its words) will reduce the lexical entrenchment of the first language, because this language will receive less exposure. Gollan et al. (2008) suggest a direct relationship between the weaker links and the frequency effect. They make the explicit hypothesis that bilinguals should have a larger frequency effect than monolinguals because a) bilinguals have used words in each language less often than monolinguals have and b) increased use leads to increased lexical accessibility only until a certain ceiling level of exposure, meaning that low frequency words should be more affected by differences in degree-of-use than high frequency words. From this hypothesis, we can also predict that in the case of unbalanced bilinguals, for whom L2 exposure is lower than the L1 exposure, the L2 FE's will also be larger than the L1 FE's. We support the idea posited by the weaker links account that differential FE's in the bilingual domain can be explained without assuming qualitatively different language processing for bilinguals compared to monolinguals and aim to specify the hypotheses put forward by the weaker links account (Gollan et al., 2008).

WORD FREQUENCY DISTRIBUTION

Because of the logarithmic relationship between corpus word frequency and lexical access time, it is customary to use logarithmically transformed corpus word frequencies in any analysis where word frequency is a variable in the model. This transformation changes the functional relationship between corpus word frequency and lexical access time from a logarithmic one to a linear one (See the upper and middle panel of Figure A.1 in Appendix A for an illustration).

When detecting changes in the size of the FE related to language exposure, it is important to note that when these transformed corpus word frequencies are used, the *size* of the word frequency effect is not affected by absolute exposure. In other words, while a participant who has more exposure to a certain language will be faster to process words in that language than a participant who has little exposure to that language, an analysis based solely on transformed corpus word frequency would predict that the difference in processing times for high frequency and low frequency words, in other words the FE, is the same for both participants. Still another way of putting it is that when x and y are untransformed relative corpus word frequencies (for instance x=100 per million and y=1 per million), then for a participant who has been exposed to 100 million words the difference in absolute exposure between x and y is 9,900 (10,000-100) while for a participant who has been exposed to 10 million words, the difference is 990 (1000-10), which would lead to larger frequency effect for the participant with more exposure. When logarithmically transformed frequencies are used, for the participant with exposure to 100 million words the difference between x and y is 2 (log10 (10,000) – log10 (100) = 4 - 2 = 2), while for the participant with exposure to 10 million words, the difference between x and y is also 2 (log10 (1000) - log10 (10) = 3 – 1 = 2).

Another element to consider is that word frequency distributions are fundamentally different from normal distributions, which psychologists are used to working with. For instance, a typical characteristic of normal distributions is that the mean of a sample is an estimate that could be higher or lower than the population average and that gets more and more precise as the sample size grows. This characteristic is not shared with word frequency distributions. Instead, one of the characteristics of word frequency distributions is that the mean predictably increases as the sample, or the corpus, grows (Baayen, 2001). Importantly, Kuperman and Van Dyke (2013) show that relative word frequency is also related to the corpus size. They demonstrate that as corpus size grows, the relative frequency of low frequency words increases while the relative frequency of high frequency words stays almost constant (See Table 1). By dividing words in ten frequency bands, they show that words in the lowest frequency band (1) have an estimate of relative frequency that is twice as large in a corpus of 50 million words than in a corpus of 5 million words (ratio: 2.234); relative frequency estimates for words in the highest frequency band (10), on the other hand, were nearly equivalent (ratio: 1.003).

Table 1. *The ratio of a word's relative frequency in the 50-million token SUBTLEX corpus to its relative frequency in a sample of 5 million tokens (Relative frequencies averaged over 1000 samples). Taken from Kuperman & Van Dyke (2013).*

Frequency class	1	2	3	4	5	6	7	8	9	10
Between-sample ratio	2.234	2.083	1.672	1.344	1.020	1.020	0.996	0.998	1.012	1.003

Note. Ratios are averaged per frequency class (1 = lowest frequency; 10 = highest frequency) and are based on a pool of 500 words, with 50 words per frequency class.

It is precisely this characteristic of word frequency distributions that is overlooked in the analysis of the effect of word frequency. If the evolution of relative word frequency with more exposure follows a trajectory that is analogous to the evolution of relative frequency with increase in corpus size, this alone can account for differences in the size of the FE. On these grounds, an interaction of proficiency and corpus frequency is *expected,* but it should not be attributed to qualitative differences between poor and good readers, or between a categorical difference between monolinguals and bilinguals. As we already mentioned, when assuming lower exposure to all items in the lexicon and using raw corpus word frequencies in the analyses, a larger FE *slope* is expected. When we log transform these word frequencies we do not necessarily expect a larger FE slope as long as the ratios between the relative frequencies stay the same. The importance of changes for low frequency words but not for high frequency words is exactly what a logarithmic transformation accounts for; differences in the frequency effect due to a lower exposure to all words in the lexicon should not be found if a logarithmic transformation is used and if there are no changes in *relative* word frequency. However, if relative subjective frequencies do not stay constant, this difference should lead to a difference in the size or slope of the frequency effect when a logarithmic transformation is applied to the frequencies. It should be noted that the reasoning that differences in the size of the frequency effect are only due to the logarithmic relationship between word frequencies and word processing times, is therefore incomplete (e.g., Duyck, Vanderelst, Desmet & Hartsuiker, 2008; Schmidtke, 2014).

LANGUAGE EXPOSURE

The weaker links theory is consistent with the individual differences account of Kuperman and Van Dyke (2013) in the sense that differences in the FE are attributed to the degree of exposure rather than to qualitative differences originating from the acquisition of multiple languages. However, the weaker links theory makes the general claim that a) there is an overall lower (absolute) exposure to language for bilinguals than for monolinguals and b) that this results in a larger FE for bilinguals.

A pure exposure-based account leaves open the possibility that bilinguals may have the same degree of exposure to one (or, in principle, more) of their two languages as monolinguals have and this account can specify the exact locus of the modulation of the size of the FE, namely that it arises from differences in ratios of high and low relative frequencies for individuals with different levels of exposure.

As already discussed, language exposure should be an important determinant of the shape and size of the FE. It is therefore of vital importance to have a good measurement for this variable. Most experiments use subjective measures like questionnaires to assess exposure, some try to quantify exposure by measuring language proficiency. Because there is a direct relation between the obtained measure of vocabulary size and the degree of exposure (e.g., Baayen, 2001), we prefer the use of a vocabulary test to assess language proficiency. By using vocabulary growth curves (see Figure 1), we can see a tight relationship between language exposure (word tokens on the x-axis) and vocabulary size (word types on the y-axis). Word tokens are counted as every word in a language corpus, including repetitions and word types are unique words. As the number of word tokens grows, so does the number of word types.

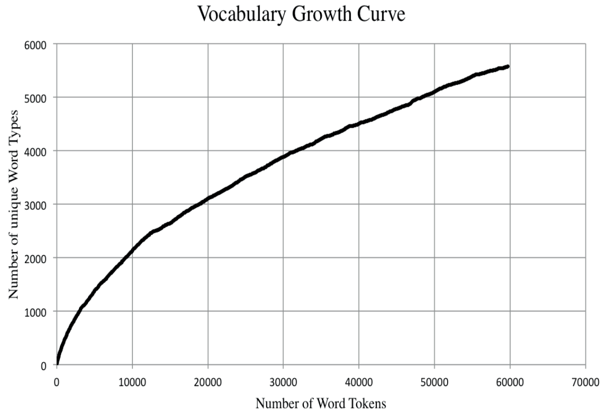


Figure 1. An example of a vocabulary growth curve. This plot shows the number of word tokens encountered (on the x-axis) and the amount of encountered word types (on the y-axis) when reading the Dutch version of the novel 'A mysterious affair at Styles'.

When vocabulary size is small, the probability that the next encountered word will be a hitherto unseen type is large, but as exposure grows the probability that the next word will be a new type decreases. As a result, to double vocabulary size requires much more than twice the amount of exposure. Concurrently, the more exposure one has, the smaller the increase in vocabulary size that is associated with additional exposure. Assuming no large differences in the complexity of material that one is exposed too, a similar vocabulary score indicates similar exposure and an increase in vocabulary scores indicates a higher degree of exposure. For subjects with an equal but very high vocabulary score, it becomes more uncertain that they have the exact same amount of language exposure. Nevertheless, on the whole, when participants have equal proficiency scores, we do not expect differential FE's, because language exposure should be quite similar.

Kuperman and Van Dyke (2013) note that robust interactions between language proficiency and word frequency have been found in a wide range of studies concerning individual reading differences: More proficient readers showed a smaller frequency effect on reaction times. (For examples see Chateau & Jared, 2000 and Diependaele et al. 2013)

Although this is indeed a robust finding, it must be noted that some authors have claimed that this finding might be an artifact of the base-rate effect (Butler & Hains, 1979; Faust et al. 1999; Yap et al., 2012). The base-rate effect is the observation that the magnitude of lexical effects correlates positively with reaction latencies. This would mean that the larger frequency effects for participants with a lower language proficiency score would be mainly due to the fact that their reaction times are longer than higher skilled participants. However, Kuperman and Van Dyke (2013) showed that the interaction between word frequency and language skill is still present after z-transforming reaction times per subject, thus eliminating any kind of base rate effect.

BILINGUAL RESEARCH

As shown by analyses that find larger frequency effects for L2 word recognition when word frequencies are log-transformed (Diependaele et al., 2013; Duyck et al., 2008; Lemhöfer et al., 2008; Whitford & Titone, 2011), exposure does have a systematic relation with the size of the word frequency effect that cannot be accounted for by the logarithmic relation between word processing times and word frequencies alone.

In the tradition of the weaker links account and as evidence for reduced lexical entrenchment for bilinguals compared to monolinguals, the bilingual FE has been compared with the monolingual FE. Indeed, when we look at the experimental findings, we find that when proficiency is equal across groups, no differences in the size of the FE are found: Gollan et al. (2011) found a similar FE in an English lexical decision and a sentence reading task for balanced Spanish-English bilinguals as for English monolinguals; Duyck et al.'s (2008) study did not find a difference between the L1 FE of unbalanced Dutch-English bilinguals and the FE of English monolinguals in lexical decision times either. The studies that did find a larger bilingual FE used bilingual participants with lower proficiency, and thus lower exposure, for the tested language than the monolinguals; also the tested language was acquired later than their other language. This means that the corpus frequencies were probably overestimated for the lower frequent words for the bilingual group, inflating reaction times for the low range. For example, Lehtonen et al. (2012) found a larger FE in a Finnish lexical decision task for balanced Finnish-Swedish bilinguals than for Finnish monolinguals. When we look at the Finnish proficiency scores we see that the bilinguals scored significantly lower than the monolinguals. Also, Lemhöfer et al. (2008) found a larger FE for different groups of bilinguals in English, their L2, than for English monolinguals in a word identification task. Gollan et al. (2011) showed that the L2 FE for Dutch-English bilinguals in a lexical decision task was larger than for English monolinguals. Naturally the bilinguals had less exposure to their L2 than the monolinguals had for their L1. These two last studies used raw frequencies.

In short, the results of all of these studies are congruent with our expectations, namely that language exposure could account for all differences found between bilingual and monolingual FE's.

Indeed, Diependaele et al. (2013) reinvestigated Lemhöfer et al.'s (2008) English word identification times, using log-transformed word frequencies. They hypothesized that target language proficiency is the determining factor for identification times both in the L1 of the monolinguals and in L2 of the bilinguals, without a qualitative difference between L1 and L2 processing. They found a larger FE for bilinguals' word identification times in L2, than for the monolinguals' word identification times in L1. When they added target language proficiency in their model, the FE modulation by group was no longer significant. Higher target language proficiency reduced the size of the FE and this effect was the same for both groups.

As already discussed, within the unbalanced bilingual's lexicon, we assume lower exposure to L2 words than to L1 items. For this reason, a larger FE for bilinguals reading in L2 is expected compared to reading in L1, even when word frequencies are log-transformed. Duyck et al.'s (2008) data confirm this hypothesis. They used an English and Dutch lexical decision task to test Dutch-English unbalanced bilinguals. Using a dichotomous (low vs. high) corpus frequency manipulation, they found that the L2 FE is about twice as large as the L1 FE. Whitford and Titone (2011) used eye movement measures of L1 and L2 paragraph reading of unbalanced English-French and French-English bilinguals. Bilinguals reading in L2 showed larger FE's in gaze durations and total reading time than they did in L1. On top of that, they found a modulation of the L1 and L2 FE by L2 exposure. Bilinguals with a higher L2 exposure showed a smaller FE when reading in L2 than the bilinguals with a lower L2 exposure.

In sum, the findings of FE modulation in the bilingual field are compatible with the account that Kuperman and Van Dyke (2013) propose for individual differences in FE's for monolingual participants. Quantitative differences between language exposure, resulting in a different ratio of relative frequencies for low compared to high exposure items, can account for the differences between bilingual and monolingual language processing, but also for the differences found within groups for L1 and L2 processing.

THIS STUDY

Our study is the first to investigate the difference between the first acquired and dominant L1 FE of unbalanced bilinguals, and the monolingual FE in natural reading. Duyck et al.'s (2008) study compared the same groups (Dutch-English bilinguals and English monolinguals) but merely used an isolated word recognition task. This lexical decision task contained a limited number of 50 target words (25 low frequency and 25 high frequency words) per participant and provided only a small amount of data per participant. On top of that, the *isolated-word method* used in their experiment, represent an oversimplification of the natural way in which words are encountered, limiting ecological validity. When reading in a natural context, word processing takes place while other language processing is going on, e.g. integrating words in context, parsing of syntax, etc. Also a lexical decision task involves a behavioral response, which might require mental processes or strategic factors that are normally not associated with reading.

Until now, only 2 studies compared the frequency effects for L1 and L2 visual word recognition (Duyck et al., 2008; Whitford & Titone, 2011). In Whitford and Titone's (2011) study, comparing L1 with L2 FE's, participants read 2 paragraphs each containing only about 50 content words. In our study, the largest bilingual eye tracking data corpus (Cop, Drieghe & Duyck, 2014), bilingual and monolingual participants read a whole novel containing around 29 000 content words. Not only is this a much larger and thus more generalizable, assessment of bilingual reading, it is also an even more naturalistic setting than paragraph reading, since people often read text in the context of a coherent story.

This study also attempts to resolve a concern we have with most cited studies, namely a poor measurement of L2 proficiency and a lack of assessment of L1 proficiency. We follow Luk and Bialystok (2014) in their assertions that there are multiple dimensions of bilingualism and follow their recommendation to use both methods of subjective and objective proficiency assessments. By triangulating these different measurements, we calculated a composite proficiency score for both L1 and L2 language proficiency. Both the individual measurements as this composite score can then be used to assess differences in proficiency between the tested groups. The way this composite score was calculated is described in the method section.

Most studies on the bilingual FE use self-reported L2 language exposure as a measure of proficiency (cf. Whitford & Titone, 2011) or do not measure the language proficiency of their participants at all (cf. Duyck et al. 2008). For our analyses we use the LexTALE scores because this test has been validated as an indication of vocabulary size, a central concept in this study. Kuperman and Van Dyke (2013) explain the different individual FE's precisely by vocabulary size. On top of that, the LexTALE score has been used in multiple bilingual studies, ensuring an easy comparison between the results and replication of the effects of this score.

Interestingly, no study has ever investigated the differential effects of L1 vs. L2 proficiency for bilinguals on frequency effects. This is the first study to even add L1 proficiency to the analysis of the FE of bilingual reading data. Neither Whitford and Titone (2011), Duyck et al. (2008) nor Diependaele et al. (2013) used this variable in analyzing the bilingual data, while it is expected that the proficiency of L1, which is an indication of lexicon size and exposure, is of importance to the actual frequencies of the word forms in the bilingual lexicon.

Concerning proficiency, the weaker links account (Gollan & Acenas, 2004) always assumed a trade off between the two scores: A high L2 exposure will imply a lower L1 exposure. The proliferation of lexical items in bilinguals should necessarily lead to a lower exposure to other items and eventually to weaker links between lexical representations and their word forms. For unbalanced bilinguals we assume that the mentioned trade-off between L1 and L2 exposure will be much more unclear. We might even expect that the L1 and L2 proficiency scores should correlate positively with each other, when we assume that innate language aptitude plays a role in language acquisition. Many studies in the monolingual domain have found that participants with increased vocabulary size show a reduced response time and a higher accuracy rate in lexical decision tasks (Yap, Balota, Sibley, & Ratcliff, 2012) for both familiar and unfamiliar words (e.g. Chateau & Jared, 2000). On top of that Perfetti, Wlotko and Hart (2005) observe that individuals who are better at comprehending text or have a higher reading skill, require fewer exposures to learn new words. This means that a person with a large L1 proficiency score, will be faster at establishing a connection between a new word form and its meaning (Perfetti et al., 2005) and might thus be more likely to also have a larger L2 proficiency score.

For monolingual L1 and bilingual L1 reading, we expect that L1 proficiency should have a large influence on the size of the frequency effect, with smaller L1 FE's for higher L1 proficiency. The relationship between L1 proficiency and the FE should be the same for both groups. For the comparison between the bilingual L2 reading, L1 proficiency might have a similar effect on the size of the FE, within the vocabulary size rationale discussed above. Given the robust effects of L2 proficiency on the size of the FE in previous studies, we might expect this effect to persist even in the presence of L1 proficiency. If it does, a higher L2 proficiency is expected to reduce the FE in L2 reading but not in L1 reading.

METHOD

This method section is partly taken from Cop, Drieghe, and Duyck (2014) because the data in this analysis is a subset from a large eye movement corpus described in Cop et al. (2014).

PARTICIPANTS

Nineteen unbalanced bilingual Ghent University and fourteen monolingual Southampton University undergraduates participated either for course credit or monetary compensation. The bilingual participants' dominant language was Dutch, their second language English. They had a relatively late L2 age of acquisition (mean=11 [2.46]). The monolingual participants had knowledge of only one language: English. Bilingual participants completed a battery of language proficiency tests including a Dutch and English spelling test (GLETSHER and WRAT4), the LexTALE (Lemhöfer & Broersma, 2011) in Dutch and English, a Dutch and English lexical decision task (for results see Table B.1 in Appendix B) and a self-report language questionnaire (based on the LEAP-Q, Marian, Blumenfeld, & Kaushanskaya, 2007). Monolinguals completed an English spelling test, the English LexTALE and an English lexical decision task. We calculated a composite L1 and L2 proficiency score by averaging the score on the spelling test, the score on the LexTALE and the adjusted score of the lexical decision task. This composite score and the LexTALE scores show that bilinguals score significantly higher on L1 proficiency than they do on L2 proficiency, and that the bilinguals and monolinguals are matched on L1 proficiency. The LexTALE score is used in the analysis. Participants had normal or corrected-to-normal vision. None of the participants reported to have any language and/or reading impairments. For detailed scores see Table B.1 in Appendix B.

MATERIALS

The participants were asked to read the novel "The mysterious affair at Styles" by Agatha Christie (Title in Dutch: "De zaak Styles"). This novel was selected out of a pool of books that was available via the Gutenberg collection. The books were judged on length and difficulty, indicated by the frequency distribution of the words that the book contained. We selected the novel whose word frequency distribution was the most similar to the one in natural language use (Subtlex database). The Kullback–Leibler divergence was used to measure the difference between the two probability distributions (Cover and Thomas, 1991).

In English, the book contains 56 466 words and 5 212 sentences (10.83 words per sentence); in Dutch it contains 60 861 words and 5 214 sentences (11.67 words per sentence). The average word length in Dutch was 4.51 characters and 4.27 characters in English. The average word log frequency of the content words in the book was 3.82 for both books. Only the non-cognate content words of the novel were analyzed. The Dutch novel contained 30 817 content words and the English novel 28 108. From those words, 5 207 Dutch and 4 676 English words were individually distinct types. This means that each participant read ± 5000 different content words. See Table 2 for the description of the content words in Dutch and English. Although both word frequency and word length show minor differences across languages, these variables will be included in the higher order interactions in our linear mixed model.

Table 2. *Summary of the characteristics of the content non-cognate words of the novel: Number of Words, Average Content Word Frequency and Average Word Length. Standard deviations are in brackets.*

	Dutch	English
Number of Words	22 919	20 695
Average Content Word Frequency	3.74 [1.23]	3.79 [1.20]
Average Word Length	5.95 [2.56]	5.47 [2.23]

APPARATUS

The bilingual and monolingual eye movement data were recorded with the EyeLink 1000 system (SR-Research, Canada) with a sampling rate of 1 kHz. Reading was binocular, but eye movements were recorded only from the right eye. Text was presented in black 14 point Courier New font on a light grey background. The lines were triple spaced and 3 characters subtended 1 degree of visual angle or 30 pixels. Text appeared in paragraphs on the screen. A maximum of 145 words was presented on one screen. During the presentation of the novel, the room was dimly illuminated.

PROCEDURE

Each participant read the entire novel in four sessions of an hour and a half, except for one bilingual participant who only read the first half of the novel in English. The other bilinguals read half of the novel in Dutch, the other half in English. The order was counterbalanced.

The participants were instructed to read the novel silently while the eye tracker recorded their eye movements. It was stressed that they should move their head and body as little as possible while they were reading. The participants were informed that they would be presented with multiple-choice questions about the contents of the book after each chapter. This was done to ensure that participants understood what they were reading and paid attention throughout the session.

The text of the novel appeared on the screen in paragraphs. When the participant finished reading the sentences on one screen, they were able to press the appropriate button on a control pad to move to the next part of the novel.

Before starting the practice trials, a nine-point calibration was executed. After this, the calibration was done every 10 minutes, or more frequently when the experiment leader deemed necessary.

RESULTS

Words that had an orthographically overlapping translation equivalent in the other language were categorized as identical cognates and were excluded for the frequency analysis (Dutch: 8.1%, English: 13.7%). The first and last word on a line were excluded from the analysis (Dutch: 18.8%, English: 16.9%), because their processing times also reflect sentence wrap-up effects (e.g. Rayner et al., 1989).

In Table 3 we report the average single fixation duration, gaze duration, skipping rates and the frequency effects for monolinguals and bilinguals reading in L1 and L2. A single fixation duration is the duration of the fixation on target words that were fixated only once. The gaze duration is the time spent on the word prior to moving the eye towards the right of that word. This means that first pass refixations are included in this measure. The skipping rate of a word is the likelihood that that word will be skipped the first time it is encountered. For the sake of visualization, words were median-split by frequency to create a low and high frequency set.

In this article we report the analysis of the single fixation durations. We prefer this measure because eye movements are complex and can reflect different processes. For example, first fixation durations are used most commonly as an early measure of lexical access. However, these can consist of either the single fixation duration but also of the first fixation of multiple fixations on a word. This measure sometimes shows reversed word length effects because the first of a fixation on a longer word will be shorter because of the need to fixate a longer word multiple times (e.g. Rayner, Sereno & Raney, 1996). If there is only a single fixation on a target word, we assume that the target word is processed sufficiently with this one fixation because there is no refixation prior to moving to the next word or after doing so. Thus we prefer the measurement of single fixation duration because this would most accurately reflect lexical access time for the target word. The size of the corpus allows us to exclude words that are refixated whilst maintaining ample amount of statistical power. For the analyses of the other 3 dependent variables, we refer to Tables S1-S6 in the online supplementary materials.

Table 3. *Average Single Fixation Durations, Gaze durations and Skipping Rates for low [0.01-3.98] and high [3.99-5.90] frequent words (LF, HF) and the L1 and L2 bilingual and monolingual frequency effects (FE).*

	Bilingual L1			Bilingual L2			Monolingual		
	LF-Words	HF-Words	FE	LF-Words	HF-Words	FE	LF-Words	HF-Words	FE
Single Fixation duration	217.9	210.7	7.2	239.3	224.9	14.4	223.9	215.1	8.8

(ms)									
Skipping Rate (%)	27.6	48.9	21.3	23.8	44.0	20.2	29.9	51.0	21.1
First Fixation Duration (ms)	216.6	210.2	6.4	233.4	223.0	10.4	221.5	214.9	6.5
Gaze duration (ms)	241.8	223.9	17.9	277.9	244.6	33.3	245.3	227.4	17.9

BILINGUAL L1 READING VS. MONOLINGUAL READING

For the comparison between monolinguals and bilinguals reading in L1, all words that were either not fixated or were fixated more than once were excluded (46,63%). Single fixations that differed more than 2.5 standard deviations from the subject means were excluded from the dataset (2.23%). This left us with 265 756 data points. The dependent variable was log transformed to normalize the distribution as suggested by the Box-Cox method. This transformation did not change the functional relationship between the single fixation durations and the log-transformed word frequencies (see Figure A.1 in Appendix A). This data was fitted in a linear mixed model using the lme4 package (version 1.1-7) of R (version 3.0.2). The model contained the fixed factors of Bilingualism (L1 or mono), log 10 word frequency (continuous), L1 proficiency (continuous) and the control variable of word length (continuous). As proficiency variable we used the score on the L1 LexTALE (Lemhöfer & Broersma, 2011). For the word frequency, the subtitle word frequency measures (English: Brysbaert & New 2009; Dutch: Keuleers, Brysbaert & New, 2010) were log transformed with base 10 to normalize their distribution. All continuous predictors were centered. The maximum correlation between fixed effects in the final model was -0.063.

In the model we included a random intercept per subject. This ensured that differences between subjects concerning genetic, developmental or social factors were modeled. We also included a random intercept per word because our stimuli sample is not an exhaustive list of all words in a language. The model was fitted using restricted maximum likelihood estimation (REML). First a full model, including all of the interactions between the fixed effects and the two random clusters, was fitted. The optimal model was discovered by backward fitting of the fixed effects, then forward fitting of the random effects and finally again backward fitting the fixed effects. We strived to include a maximal random structure (Barr, Levy, Scheepers & Tily, 2013). For the final model see Table 4.

Our two groups did not differ in single fixation durations: L1 reading was equally fast for mono- and bilinguals (β=-0.019, SE=0.015, t-value=-1.25). We did find an overall frequency effect (β=-0.0082, SE=0.00095, t-value=-8.59), which was not larger for bilinguals than for monolinguals (β=0.00051, SE=0.0013, t-value=0.39).

No main effect of L1 proficiency was found. Proficiency did however interact with word frequency (β=0.00017, SE=0.000077, t-value=2.19). The score on the L1 LexTALE has a larger impact on the single fixation durations on low frequency words than on high frequency words (See Figure 2). This results in a smaller FE for participants with higher L1 proficiency scores. What is striking is that the relationship between frequency and single fixation duration is the same for monolinguals and bilinguals reading in L1. Because word length is not matched across languages (0.48 letter difference), we added word length to this higher order interaction. The 3-way interaction was not significant and did not render the significant 2-way interaction between L1 proficiency and frequency insignificant.

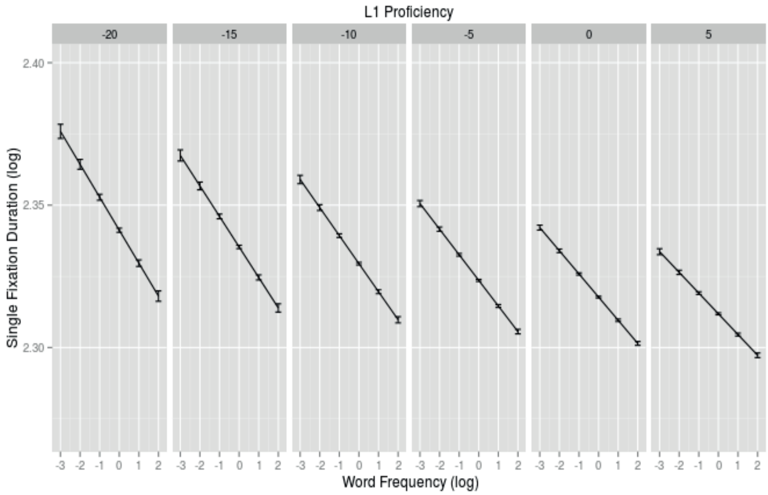


Figure 2. The effect of L1 Language Proficiency (centered on panels) and Word Frequency (centered and log-transformed on the x-axis) on Single Fixation Durations (log-transformed on the y-axis) for monolinguals and bilinguals reading in their L1. This graph is plotted using the model estimates of the relevant effects of the final model for Single Fixation Durations.

Table 4. *Estimates, standard errors and t-values for the fixed and random effects of the final linear mixed effect model for Single Fixation Durations of the comparison between L1 bilingual and monolingual reading.*

Bilingual L1 vs. Monolingual			
	Estimate	SE	t-value
Fixed Effects			
(Intercept)	2.33	0.012	194.06
Word Frequency	-0.0082	0.00095	-8.59
Bilingualism	-0.019	0.015	-1.25
L1 Proficiency	-0.0012	0.0012	-0.99
Word Frequency*L1 Proficiency	0.00017	0.000077	2.19
Word Frequency * Bilingualism	0.00051	0.0013	0.39
Control variables			
Word Length	0.0020	0.00044	4.52
Word Frequency * Word Length	-0.0013	0.00021	-6.16
L1 Proficiency * Word Length	-0.00013	0.000049	-2.55

	Variance	SD
Random Effects		
Word		
(Intercept)	0.00026	0.016

Subject		
(Intercept)	0.0024	0.048
Word Frequency	0.0000087	0.0030
Word Length	0.0000045	0.0021
Word Frequency * Word Length	0.00000078	0.00088

BILINGUAL L1 READING VS. BILINGUAL L2 READING

Again, all words that were either not fixated or were fixated more than once were excluded from the dataset (50.8%). Single fixations that differed more than 2.5 standard deviations from the subject means were also excluded (2.27%). This left us with 221 953 data points. The dependent variable was log transformed with base 10 to normalize the distribution. As we have already demonstrated, this transformation did not change the functional relationship between the dependent variable and the log-transformed word frequencies (see Figure A.1 in Appendix A). This data was fitted in a linear mixed model using the lme4 package (version 1.1-7) of R (version 3.0.2). The model contained the fixed factors of language (L1 or L2), log 10 word frequency (continuous), L1 and L2 proficiency (continuous) and the control variables of word length (continuous) and age of L2 acquisition (continuous). As proficiency variables we used the score on the L1 and L2 LexTALE (Lemhöfer et al.). We computed the frequency variable the same way as in the previous comparison. Again, all continuous predictors were centered. The maximum correlation in the final model between fixed effects was -0. 643. Again, we included a random intercept per subject and a second random intercept per word. The model was fitted using restricted maximum likelihood estimation (REML). First a full model, including all of the interactions between the fixed effects, was fitted. The optimal model was discovered by backward fitting of the fixed effects, then forward fitting of the random effects and finally again backward fitting of the fixed effects. We strived to include a maximal random structure (Barr, Levy, Scheepers & Tily, 2013). For the final model see Table 5. Our bilinguals fixated on average longer when reading in L2 than in L1 (β=-0.034, SE=0.0011, t-value=-11.37). We find an overall frequency effect (β=-0.011, SE=0.0011, t-value=-9.89) and a modulation of the FE by language (β=0.0031, SE=0.00099, t-value=3.10). The FE is larger in L2 than in L1, which is caused by a larger disadvantage for low frequency L2 words (See Figure 3).

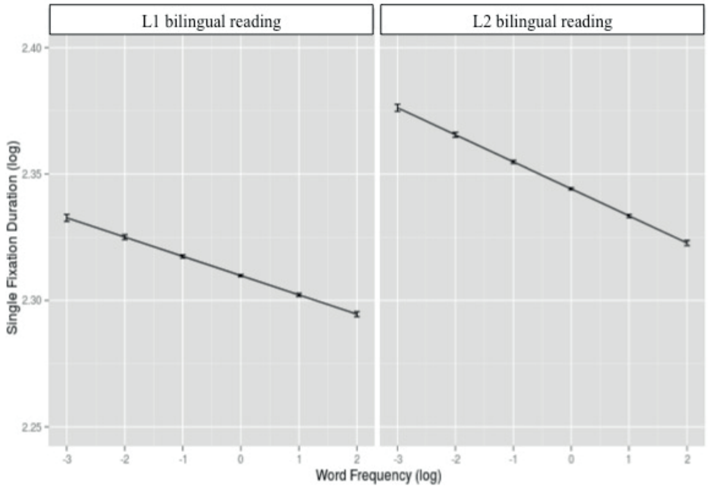


Figure 3. Single fixation durations (log-transformed) dependent on word frequency (log transformed and centered on the x-axis) and for bilinguals reading in L1 and L2 (panels). Standard Errors are indicated by whiskers. This graph is plotted using the model estimates of the relevant effects of the final model for Single Fixation Durations.

No main effects of L1 or L2 proficiency were found, but L1 proficiency modulates the frequency effect (β=0.00026, SE=0.00010, t-value=2.48). This modulation is the same when reading in L1 or L2. The FE is smaller when L1 proficiency is higher, both when the bilinguals read in L1 and in L2. We thus replicate the modulation by L1 proficiency of the FE. Figure 4 shows that the modulation of the FE by L1 proficiency is driven by speeded lexical access for low-frequent words both in L1 and L2 reading.

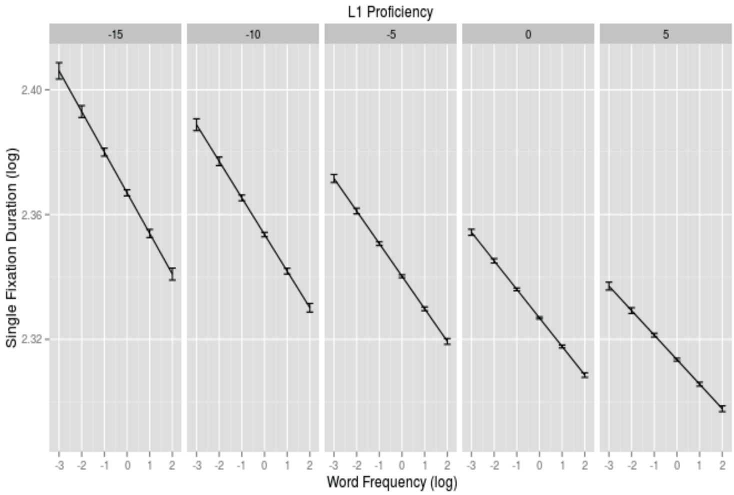


Figure 4. The effect of L1 Language Proficiency (centered on the panels) and Word Frequency (log-transformed and centered on the x-axis) on Single Fixation Durations (log-transformed on the y-axis) for bilinguals reading in L1 and L2. This graph is plotted using the model estimates of the relevant effects of the final model for Single Fixation Durations.

L2 proficiency interacted with language (β=0.00082, SE=0.00020, t-value=4.16). This means that when the bilinguals were reading in L2, there was an advantage for participants scoring high on L2 proficiency: they make shorter single fixations. For L1 reading an opposite effect was found: a higher score on L2 proficiency made the single fixation durations longer (See Figure 5).

Because word length is not matched across languages, we again added word length to the higher order interactions. These 3-way interactions were not significant and did not render the other 2-way interactions insignificant. This means that the effects described generalize for both short and long words.

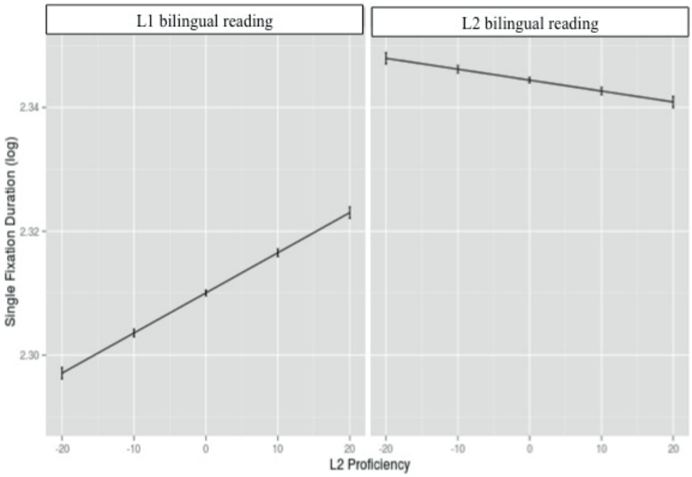


Figure 5. Single fixation durations (log-transformed) dependent on L2 Proficiency score (on the x-axis) for bilinguals reading in L1 and L2 (panels). Standard errors are indicated by whiskers. This graph is plotted using the model estimates of the relevant effects of the final model for Single Fixation Durations.

Table 5. *Estimates, standard errors and t-values for the fixed and random effects of the final linear mixed effect model for Single Fixation Durations of the comparison between bilingual L1 and L2 reading.*

Bilingual L1 vs. Bilingual L2			
	Estimate	SE	t-value
Fixed Effects			
(Intercept)	2.34	0.011	213.81
Word Frequency	-0.011	0.0011	-9.89
Language	-0.034	0.0030	-11.37
L1 Proficiency	-0.0027	0.0019	-1.41
L2 Proficiency	-0.00019	0.00096	-0.19
Word Frequency * Language	0.0031	0.00099	3.10
Word Frequency * L1 Proficiency	0.00026	0.00010	2.48
Language * L2 Proficiency	0.00082	0.00020	4.16
Control variables			
Word Length	0.0046	0.00076	6.09
Age of Acquisition L2	-0.0020	0.0035	-0.58
Word Frequency*Word Length	-0.0012	0.00014	-8.25
L1 Proficiency* Word Length	-0.00025	0.00010	-2.42
L2 Proficiency * Word Length	0.00012	0.000050	2.35
Language * Word Length	-0.0024	0.00049	-4.88

	Variance	SD
Random Effects		
Word		
(Intercept)	0.00025	0.016
Subject		
(Intercept)	0.0023	0.048
Language	0.00015	0.012
Word Frequency	0.000015	0.0038
Word Length	0.0000086	0.0029
Language * Word Frequency	0.0000048	0.0022
Language * Word Length	0.0000015	0.0012

GENERAL DISCUSSION

This paper compared the monolingual and bilingual (L1 and L2) FE in text reading. Participants read an entire novel containing ± 29 000 content words, of which ± 8 000 were nouns. Bilinguals read the novel half in Dutch (L1), half in English (L2). In the analyses of single fixation durations on non-cognate content words, we found similarly sized FE's for bilinguals and monolinguals reading in their mother tongue. A rise in L1 proficiency reduced the slope of the L1 FE. The bilinguals showed a larger FE when reading in their L2 compared to reading in their L1. We also found a modulation of the bilingual L2 FE by L1 proficiency. A rise in L1 proficiency reduced the slope of the L1 and L2 FE. L2 proficiency did not modulate the FE, but it did have a differential effect across languages. In L2 reading, a rise in L2 proficiency speeds up single fixations, for L1 reading a rise in L2 proficiency does the opposite. This trade-off of reading speed is in line with the idea of 'weaker links'. To account for both these and previous results, we propose an account that fits within the framework of the weaker links hypothesis, suggesting not only a lower exposure to all lexical items but a disproportionate overestimation of corpus word frequencies for low frequency words for smaller vocabularies. Our proposal is consistent with a purely *exposure* based explanation of language processing speed.

BILINGUAL VS. MONOLINGUAL L1 FE

We find a similarly sized FE for bilinguals reading in L1 and monolinguals reading in their mother tongue. Our findings seem at odds with the weaker links account, which predicts that due to a lower exposure to all items in the bilingual lexicon, bilinguals would show an overall larger FE in both their languages compared to a monolingual. Gollan and Acenas (2008), who mostly tested balanced Spanish-English populations, make the implicit assumption with their weaker links account that the total language exposure is equal for all people. While this maybe the case for their participants, it is definitely not true for all groups of bilinguals. Our population of unbalanced bilinguals usually acquires a 2nd language in a classroom context, thus increasing their total language exposure, not per se substantially decreasing their L1 exposure. The acquisition of a second language for adults might be more defined by actively seeking more language exposure in a second language, resulting in indeed a larger lexicon, but also a higher total exposure. The hypothesis that bilingual exposure to L1 is not substantially lowered by bilingualism is supported by the fact that the L1 proficiency of our monolinguals was equal to the L1 proficiency of the bilinguals7. The similar proficiency scores indicate a similar sized vocabulary and thus a similar exposure to L1 for both groups. This contrast with most studies reporting differential FE's for bilinguals compared to monolinguals which use balanced bilingual populations and/or report lower target language proficiency for bilinguals than for monolinguals (Gollan et al., 2011; Lemhöfer et al., 2008; Lehtonen et al., 2012). To conclude, the weaker links account connects lower language exposure, leading to lower proficiency, to a larger FE. We nuance this rationale by pointing out that not all bilingual groups necessarily have lower L1 exposure than monolinguals do. This means that as long as there are no differences in language exposure as measured by language proficiency, we do not expect differently sized FE's. We would only predict a perceivable disadvantage for bilinguals in L1 compared to monolinguals when vocabulary size, and thus exposure, would be considerably smaller for the bilinguals.

[7] All 4 methods measuring L1 proficiency (LexTALE, lexical decision task, spelling test and the proficiency questionnaire) do not yield different scores for the two groups (see Table B.1 in Appendix B for a summary of the objective measures). This makes it highly unlikely that we fail to pick up on existing language proficiency differences between our two groups.

The second important observation in our data is the reduction of the monolingual and bilingual L1 FE as L1 proficiency rises. This is consistent with multiple findings in the literature. For example Ashby, Rayner and Clifton 's (2005) eye tracking experiment found that underperforming adults show a larger frequency effect especially for low frequency words. Also, Kuperman and Van Dyke (2011) showed that individual language skill scores in rapid automatized naming and word identification modulated frequency effects for fixation times. Participants scoring high on language skill, showed a smaller frequency effect. Diependaele et al. (2013) showed that both for monolinguals and bilinguals, the rise of target language proficiency makes the size of the FE of word identification times smaller. Kuperman and Van Dyke (2013) observed that the relative amount of exposure to high corpus based frequency words will be virtually identical for individuals with different language experiences, whereas the low corpus frequency words will yield a larger difference in exposure, i.e. lexical entrenchment, for different groups.

In short, a higher L1 proficiency score reflects the size of the lexicon and the exposure to the items in that lexicon. Our results show, consistent with ideas formulated in Diependaele et al. (2013), that target language proficiency explains the size of the FE in both monolingual and bilingual groups and that the relationship between proficiency and FE is exactly the same for these two groups. This implies that we do not need qualitatively different lexical processing mechanisms to explain the size of L1 FE's for monolinguals and unbalanced bilinguals.

When we look at the mechanisms behind this modulation of the FE, we can draw conclusions about the location on the word frequency range this effect takes place. As we see a modulation of the FE by L1 proficiency even when word frequency is log transformed, this means that L1 proficiency does not measure absolute L1 exposure but is more sensitive to the L1 exposure for low frequency L1 items.

BILINGUAL L1 VS. BILINGUAL L2 FE

Bilinguals show a larger effect of frequency in the processing of L2 text than in the processing of L1 text. This finding is compatible with findings of Duyck et al. (2008) and Whitford and Titone (2011), who also found larger L2 FE's for unbalanced bilinguals, respectively for sentence reading and paragraph reading.

This finding is compatible with accounts of word recognition that implement implicit learning. In unbalanced bilingual populations, L2 words are learned later than L1 words and they have received on average less exposure than L1 words, thus making the threshold for activation for L2 items lower or the representations of these L2 words less accurate. Because we used corpus word frequencies in our analyses, the actual word exposure is overestimated for L2 reading compared to L1 reading. Kuperman and Van Dyke showed that this is especially the case for words with a low corpus frequency. This results in a larger FE in L2 mainly driven by a disproportional slower processing of low frequency words (See Figure 3).

Both in L1 and L2, a larger L1 proficiency reduces the slope of the FE. The effect of L1 proficiency on L1 reading is explained extensively in the above section: the processing time becomes disproportionally faster for low frequency than for high frequency words as exposure rises, causing a smaller FE.

The effect of L1 proficiency on L2 reading is much more surprising. Apparently, increased vocabulary size in the mother tongue facilitates access to low frequency words in a second language. To accommodate this finding, we have to assume that the L1 vocabulary size is measuring something more than exposure to the mother tongue. It is reasonable to assume that the amount of L1 exposure should be approximately the same for subjects with similar SES, education and age. Given that we do find different L1 proficiency scores, we are probably picking up on a more abstract reading skill or general language aptitude by measuring L1 vocabulary size. This assumption makes it more understandable that L1 proficiency modulates the FE in L2 reading in much the same way as it does in L1 reading. This line of reasoning is compatible with the idea proposed by Perfetti et al. (2005) that there is some individual variable that determines the speed of learning connections between word forms and meaning. We seem to capture this variable with our measure of L1 proficiency.

Diependaele et al. (2013) showed that proficiency explained the difference in FE across groups. In our data proficiency modulated the FE, but it did not eliminate the interaction between frequency and group. This means that the size of the FE was not totally explained by proficiency score. This is not that surprising, given that eye movement measures are more complex than identification times. Also, Whitford and Titone (2011) 's results are in line with ours, seeing that they still found differences across groups after proficiency was added to their model.

In our data L2 proficiency did not have an effect on the size of the FE, neither in L1 reading nor in L2 reading. Higher L2 proficiency scores did however reduce L2 reading speed, which validates the measure. For reading in L1, a rise in L2 proficiency made the single fixation durations longer. High L2 proficiency does seem to reduce reading speed in L1, congruent with the idea of weaker links. These are the only effects of L2 proficiency we find in our reading data. It seems that while L1 proficiency has a disproportioned impact on low frequency words in both languages, L2 proficiency has an equally large impact on low and high frequency words, but an opposite effect in both languages. Our results thus show that, despite the high correlation between the two, L1 and L2 proficiency are distinct concepts. L1 vocabulary size seems to be a measure for a general language aptitude, while L2 vocabulary size might be more linked to actual L2 exposure.

Although we tested similar populations (unbalanced bilinguals[8]) in a similar task (natural reading), Whitford and Titone (2011) found that more L2 exposure was linked to a larger L1 FE, but to a smaller L2 FE. So in their data L1 and L2 FE's are a function of L2 exposure, while our data shows

8 Note that the languages of the tested populations were different. In our study Dutch-English bilinguals were tested, in Whitford and Titone's (2011) English-French bilinguals were tested.

that L1 and L2 FE's are a function of L1 proficiency. A large factor to take into account when trying to reconcile our data with those of Whitford and Titone is that their analysis did not actually include L1 proficiency of the bilinguals. Given that L1 and L2 proficiency are highly correlated, it is plausible that removing one of the factors from the analysis will have an impact on the significance of the other. Another factor is that they use a subjective estimate of L2 exposure in their analysis, while we use an objective vocabulary score to approximate language exposure. When we enter the subjective L2 exposure ratings in our analysis without L1 proficiency, we see that Subjective L2 exposure does have an effect on the slope of the L1 and L2 FE, just as in Whitford and Titone. A higher subjective exposure to L2, reduces the slope of the FE in L1 and L2. Again, a lower exposure, inflates the FE. So, the fact that L2 exposure influences the size of the FE is compatible with Whitford and Titone's results. What is not compatible is that we do not find a differential effect of this subjective L2 exposure on L1 and L2 reading. In our data, the effect of L2 exposure is the same in L1 and L2 reading, with smaller FEs for both languages.

Another possible reason for these different findings is that Whitford and Titone (2011) use gaze durations and total reading time as dependent variables. As already explained we prefer single fixation durations due to the complexity of eye movement variables. In their appendix they do report analyses of first fixation duration and skipping rates, but not single fixation durations. Their results for first fixation durations patterned with their results for gaze durations.

Our results are compatible with the assumption that the interaction between language proficiency and word frequency reported across a number of studies is caused by the use of corpus based word frequencies. Kuperman and Van Dyke (2013) show that in eye movement data the interaction between proficiency and frequency disappears when the objective corpus frequencies are replaced in the analysis by subjective frequencies, acquired by familiarity ratings. These subjective frequencies are supposed to be a closer approximation of the exact number of times a person has been exposed to a word form. For future studies, we recommend the use of more accurate estimates of actual word frequencies of bilingual populations to study the bilingual and monolingual FE.

A possible criticism to our comparison of English and Dutch text is that the larger FE's for L2 compared to L1 reading could be explained by inherent language differences between English and Dutch, not controlled for in the experimental design. Given that the monolingual (English) - L1 Bilingual (Dutch) comparison did not yield any significant differences across groups, the differences we did find across languages in L1 (Dutch) and L2 (English) are very unlikely to be due to inherent language characteristics. Also the two most important lexical variables, word length and word frequency, were included in all of the higher order interactions in each model. This ensures that the reported effects are not due to any differences between the English and Dutch texts regarding word frequency or word length.

Even so, it could be pointed out that, although the Dutch language is very closely related to English, English has a deeper orthography than Dutch (Aro & Wimmer, 2003). This means that the mapping from orthography to phonology is less transparent for English than for Dutch. This deeper orthography could, according to the orthographic depth hypothesis (Katz & Feldman, 1983) lead to more reliance on the orthographic route of visual word recognition leading to more coarse-grained language processing. In this view, one could assume that this larger reliance on lexical representations for deep orthographies could cause larger word frequency effects on lexical access in those languages. This orthographic depth hypothesis is not without challenge (e.g. Besner & Hildebrandt, 1987; Lukatela & Turvey, 1999; Seidenberg, 1985, 1992;Tabossi & Laghi, 1992). For example Besner and Hildebrandt (1985) compared naming in two Japanese syllabic orthographies and show that Japanese readers always use the orthographic route, regardless of the orthographic depth of the script they are reading. Second, looking at data supporting the orthographic depth hypotheses, no cross-lingual comparison has found a modulation of the size of the frequency-effect by the orthographic depth of a language (Frost, Katz & Benin, 1987; Seidenberg & Vidanovic, 1985) and, to our knowledge, no study finds effects of orthographic depth on eye movements. As far as we know, the only evidence for a modulation of the frequency effect by depth of orthography comes from a study by Frost (1994). He compared naming of words in two scripts of Hebrew; an unpointed (deep) and a pointed (shallow) variant. He found a frequency effect for unpointed Hebrew words and no frequency effect for pointed Hebrew words. The absence of any frequency effect in the pointed script is probably caused by a) the very transparent nature of the script and the task used, which makes it sufficient to use strict grapheme to phoneme conversion rules without activating the correct lexical representation and/or b) the low frequent use of this particular script. Both of these factors are not applicable to reading Dutch. According to the same orthographic depth hypothesis, language learners rely more on phonology than adult skilled readers, regardless of language (e.g. Katz & Feldman, 1983). This means that L2 reading of English should rely less on the orthographic route, than L1 reading. So this hypothesis would actually predict a smaller frequency effect for L2 readers of English compared to L1 readers of English or Dutch, the opposite of what we observed.

CONCLUSION

A systematic exploration of the bilingual and monolingual FE in text reading showed that the FE is modulated by L1 proficiency, both for monolinguals and for bilinguals in L1 and L2.

The size of the FE was comparable for bilinguals and monolinguals when both groups read in their mother tongue. . Bilinguals displayed no disadvantages in any of the L1 proficiency (see Appendix B) or any of the L1 reading measures under investigation (see results and supplementary materials) compared to monolinguals. A higher score on L1 proficiency reduced the size of the FE equally for both groups. The size of the FE was larger for bilinguals reading in L2 compared to bilinguals reading in L1. Bilinguals showed clear proficiency (see Appendix B) and reading disadvantages (see results and supplementary materials) in L2 compared to L1. The size of the FE was reduced for participants with higher scores on L1 proficiency, both for L1 and L2 reading. Whereas objective L2 proficiency had no effect on the slope of the FE, neither in L1 reading nor in L2 reading, a subjective rating of L2 exposure did modulate the size of the FE. A higher subjective exposure to L2 reduces the slope of the FE in L1 and L2. Because of the log transformation of the word frequency measure, we can attribute the modulation of the frequency effect to a disproportionate lower exposure to words with a low corpus frequency in L2 compared to L1.

These results are easily reconcilable with the weaker links account and a) provide evidence for the assumption that the same qualitative relationship between exposure frequency and word recognition exists for all language users and b) clarify that it is not a lowering of exposure to all items in the lexicon, but a disproportional lowering of the exposure to words with a low corpus word frequency that inflates the FE.

REFERENCES

Ashby,J., Rayner, K., & Clifton, C.J. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *The Quarterly Journal of Experimental Psychology, 58A* (6), 1065-1086.

Baayen, R.H. (2001). *Word frequency Distributions.* Dordrecht, The Netherlands: Kluwer Academic Publishers.

Balota, D., Pilotti, M., Cortese, M. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory and Cognition, 29*, 639-647

Besner, D., & Hilderbrandt, N. (1987). Orthographic and phonological codes in the oral reading of Japanese kana. *Journal of Experimental Psychology: Learning, Memory and Cognition, 13*, 335-343.

Brysbaert, M. & Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? *Bilingualism: Language and Cognition, 13* (3), 359-371.

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41* (4), 977–990.

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bülte, J., & Bühl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimate. *Experimental Psychology, 58* (5), 412-424.

Butler, B., & Hains, S. (1979). Individual differences in word recognition latency. *Memory & Cognition, 7,* 68–76.

Chateau, D., & Jared, D. (2000). Exposure to print and word recognition processes. *Memory & Cognition, 28*, 143-153.

Christie, A. *The Mysterious Affair at Styles*. N.p.: John Lane, 1920. *Project Gutenberg*. 1 Mar. 1997. Web. 07 Nov. 2012. <http://www.gutenberg.org/>.

Connine, C. M., Mullennix, J., Shernoff, E., & Yelen, J. (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16(*6), 1084-1096.

Cop, U., Drieghe, D., & Duyck, W. (2014). Eye Movements in Natural Reading: a Comparison of Monolingual and Bilingual Novel Reading. Manuscript submitted for publication.

Diependaele, K., Lemhöfer, K. & Brysbaert, M. (2013). The Word Frequency Effect in First- and Second-language word recognition: A lexical entrenchment account. *The Quarterly journal of experimental Psychology, 66*(5), 843-863.

Dijkstra, T., & Van Heuven, W.J.B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5* (3), 175-197.

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Learning, Memory, and Cognition*, *33*(4), 663-679.

Duyck, W., Vanderelst, D., Desmet, T. & Hartsuiker, R.J. (2008). The frequency effect in second-language visual word recognition. *Psychonomic Bulletin & Review, 15*(4), 850-855.

Faust, M., Balota, D., Spieler, D., & Ferraro, F. (1999). Individual differences in information-processing rate and amount: Implications for group differences in response latency. *Psychological Bulletin, 125,* 777–799.

Forster, K. I. F. & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, 12, 627-35

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for Visual Word Recognition and Orthographical Depth: A Multilingual Comparison. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(1), 104-115

Frost, R. (1994). Prelexical and Postlexical Strategies in Reading: Evidence From a Deep and a Shallow Orthography. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(1), 116-129.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General, 113*(2), 256-281.

Gollan, T. H., & Acenas, L. A. (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*, 246–269.

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition (Pre-2011), 33*(7), 1220-1234.

Gollan, T.H., Montaya, R.I., Sera, C., & Sandoval, T.C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism and the weaker links hypothesis. *Journal of Memory and Language, 58*(3), 787-814.

Gollan, T. H., & Silverberg, N. B. (2001). Tip-of-the-tongue states in Hebrew–English bilinguals. *Bilingualism: Language and Cognition, 4,* 63– 84.

Gollan, T.H., Slattery, T.J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental psychology, 140*(2), 186-209.

Howes, D.H., & Solomon, R.L. (1951). Visual duration threshold as a function of word-probability. *Journal of Experimental Psychology*, **41**(6), 401-410

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica, 127,* 277-288.

Katz, L. & Feldman, L. (1983). Relation between Pronunciation and Recognition of Printed Words in Deep and Shallow Orthographies. *Journal of Experimental Psychology: Learning, Memory and Cognition, 9* (1), 157-166

Keuleers, E., Brysbaert, M. & New, B. (2010). SUBTLEX-NL: A new frequency measure for Dutch words based on film subtitles. *Behavior Research Methods, 42*(3), 643-650.

Kuperman, V. & Van Dyke, J.A. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language, 65,* 42-73.

Kuperman, V. & Van Dyke, J.A. (2013). Reassessing word frequency as a Determinant of word recognition for skilled and unskilled readers. *Journal of Experimental Psychology, 39* (3), 802-823.

Lehtonen, M., Hulten, A., Rodriguez-Fornells, A., Cunillera, T., Tuomainen, J., & Laine, M. (2012). Differences in word recognition between early bilinguals and monolinguals: Behavioral and ERP evidence. *Neuropsychologia, 50,* 1362-1371.

Lemhöfer, K., & Broersma, M. (2011, advance online publication). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*. doi:10.3758/s13428-011-0146-0.

Luk, G. & Bialystok, E. , 2014. Bilingualism is not a categorical variable: Interaction between Language proficiency and usage. *Journal of Cognitive Psychology, 25:5,* 605-621.

Lukatela, G., & Turvey, M.T. (1999). Reading in two alphabets. *American Psychologist*,
53, 1057–1072.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. Journal of Speech, Language, and Hearing Research, 50, 940–967.

McClelland, J., & Rumelhart, D. (1981). An interactive activation model of context effects in letter perception. An account of basic findings. *Psychological Review*, *88*, 375–407.

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General, 118,* 43-71.

Monsell, S. (1991). The nature and locus of word frequency effects in reading. In D. Besner & G.W. Humphreys (Eds.), *Basic Process in reading: Visual word recognition* (pp. 148-197). Hillsdale, NJ: Erlbaum.

Morton, J. (1970). A functional model of human memory: In D. A. Norman (Ed.), *Models of human memory* (pp. 203-254). New York: Academic Press

Murray, W.S., & Forster, K.I. (2004). Serial mechanisms in lexical access: the rank hypothesis. *Psychological Review, 111,* 721-756

Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), Reading acquisition (pp. 145–174). Hillsdale, NJ: Lawrence Erlbaum

Associates.
Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. Scientific Studies of Reading, 11, 357–383.
Perfetti, C., Wlotko, E., & Hart, L. (2005). Word learning and individual differences in word learning reflected in event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1281-1292.
Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*, 4(special issue), 21–49.
Rayner, K., Sereno, S.C. & Raney, G.E. (1996). Eye movements Control in Reading: A Comparison of two types of models. *Journal of Experimental Psychology, 22* (5), pp. 1188-1200.
Rubenstein, H., Garfield, L. & Millikan, J.A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior, 9,* 487-492.

Scarborough, D.L., Cortese, C., & Scarborough, H.S. (1977). Frequency and repetition effects in lexical memory, *Human Perception and Performance*, *3,* 3-12
Schmidtke, J. (2014). Second language experience modulates word retrieval effort in bilinguals: evidence from pupillometry. *Frontiers in Psychology,5, doi:10.3389/fpsyg.2014.00137*

Seidenberg, M.S. (1985). The time-course of phonological code activation in two writing systems. *Cognition, 19*, 1-30.
Seidenberg, M.S. (1992). Beyond orthographic depth in reading: Equitable division of labor. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 85–118). Amsterdam: Elsevier.
Seidenberg, M.S., & Vidanovic, S. (1985). Word Recognition in Serbo-Croatian and English: Do they differ? Paper presented at the XXVI Annual Meeting of the Psychonomic Society, Boston.
Solomon R. L., & Howes, D. H. (1951). Word frequency, word value, and visual duration thresholds. *Psychological Review*, 58, 256- 270.
Tabossi, P., & Laghi, L. (1992). Semantic Priming in the Pronunciation of Words in Two Writing Systems: Italian and English. *Memory and*

*Cognition, 20,*315-328.

Van Assche, E., Duyck, W., & Hartsuiker, R.J. (2012). Bilingual word recognition in a sentence context. *Frontiers in psychology,*, *3*:174, doi:10.3389/fpsyg.2012.00174 .

Van Assche, E., Duyck, W., Hartsuiker, R., & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. *Psychological science*, *20*(8), 923-927.

Van Heuven, W.J.B., Mandera, P. , Keuleers. E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology, 67* (6), 1176-1190.

Whitford, V., & Titone, D. (2011). Second-language experience modulates first-and second-language word frequency effects: Evidence from eye movement measures of natural paragraph reading. *Psychonomic Bulletin & Review,* doi: 10.3758/s13423-011-0179-5.

Yap, M.J., Balota, D.A., Sibley, D.E. & Ratcliff, R. (2012) Individual differences in visual word recognition: Insights from the English lexicon project. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 53-79

APPENDIX A: THE FUNCTIONAL RELATIONSHIP BETWEEN WORD FREQUENCY AND SINGLE FIXATION DURATION

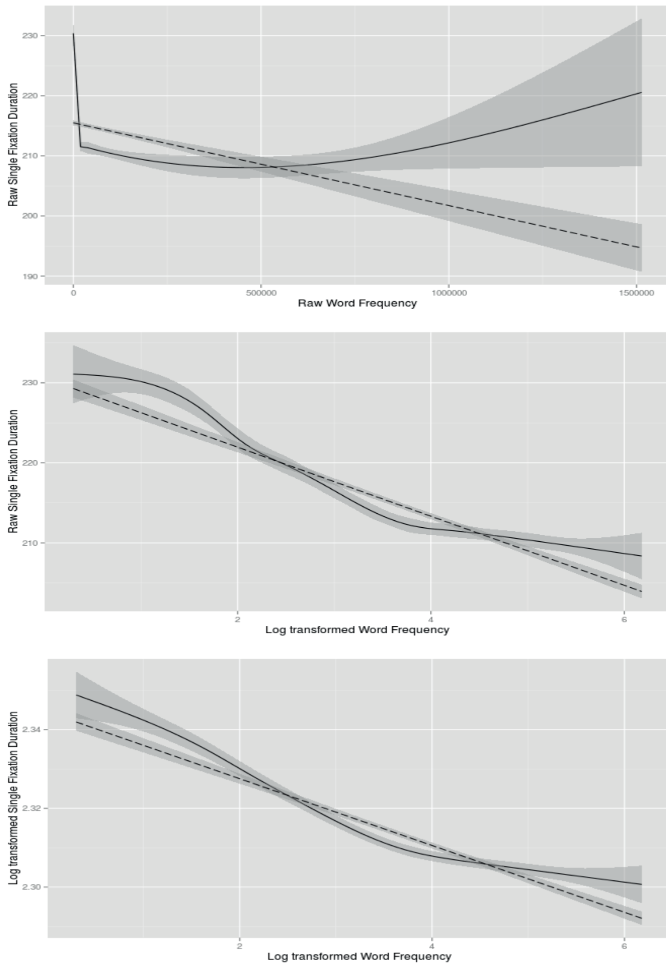


Figure A.1. The functional relationship between corpus word frequency and single fixation durations for non-cognate nouns for Dutch-English bilinguals reading the novel in Dutch. Dashed lines show the best linear fit, full lines show the best non-parametric additive fit. The first panel shows the relationship when both variables are untransformed. The second panel shows the relationship between untransformed fixation durations and log10 transformed word frequencies. The third panel shows the relationship when both variables are log-transformed with base 10. The second and third panels look similar, because the transformation of the dependent variable only caused a small change.

APPENDIX B: PROFICIENCY SCORES

Due to the lack of a standardized cross lingual spelling test, we tested the English spelling with the spelling list card of the WRAT 4 (Wilkinson & Robertson, 2006) and the Dutch spelling with the GLETSCHR (Depessemier & Andries, 2009). The LexTALE (Lexical Test for Advanced Learners of English) is an unspeeded lexical decision task, which is an indicator of language proficiency for intermediate to highly proficient language users, validated for English, Dutch and German (Lemhöfer & Broersma, 2011). A classical speeded lexical decision task was also administered in Dutch and English. The mean accuracy scores for the LexTALE and the percentage of correct word trials corrected for false alarms for the lexical decision task are reported in Table B.1.

Table B.1. *Average percentage scores (standard deviations between brackets and range between square brackets) on the LexTALE, Spelling test and Lexical Decision task for the bilingual and monolingual group.*

	Monolinguals	Bilinguals L1	Bilinguals L2	t-value L1-L2	t-value L1-mono
LexTALE- score (%)	91.07(8.92) [71.25-100]	92.43 (6.34) [73.75-100]	75.63(12.87) [51.25-98.75]	7.59 ***	0.49
Spelling score (%)	80.78 (7.26) [73.81-90.48]	83.16(7.80) [67.00-93.00]	69.92 (8.74) [52.00-83.00]	8.15 ***	0.99
Lexical Decision score (%)	77.89 (12.01) [54.61-95.23]	80.47 (5.45) [68.87-88.76]	56.75 (11.01) [38.46-75.86]	9.87 ***	0.67
Composite Proficiency Score (%)	83.25 (8.30) [67.35-94.40]	85.54 (4.68) [77.87-95.25]	67.81 (9.72) [52.49-86.76]	11.78***	0.93

* p<0.05, **p<0.01, ***p<0.001

The Dutch (L1) proficiency of the bilinguals was matched with the English proficiency of the monolinguals (See Table B.1 in Appendix B), indicating that both groups were equally proficient in their first language. Neither the LexTALE (t=0.488, df=22.254, p=0.630), the spelling test (t=0.989, df=29.282, p=0.331), nor the lexical decision tasks (t=0.667, df=17.092, p=0.514) yielded significant differences for these two groups performing in L1. Consequently, there are no differences between the composite proficiency scores ($t=-0.932$, $df=19.051$, $p=0.363$). The bilingual L2 LexTALE scores were significantly lower than their L1 scores ($t=7.587$, $df=18$, $p<0.001$). The bilingual L2 Spelling scores were lower than the L1 scores ($t=8.154$, $df=18$, $p<0.001$). The performance of the bilinguals on the classic lexical decision task was significantly better in L1 ($t=9.873$, $df=18$, $p<0.001$) than in L2. Bilinguals have lower composite proficiency scores in L2 than in L1 ($t=11.777$, $df=18$, $p<0.001$).

CHAPTER 5
BILINGUALS READING A NOVEL: AN EYE MOVEMENT STUDY OF COGNATE FACILITATION IN L1 AND L2 NATURAL READING [9][10]

This study examined how noun processing in bilingual reading is influenced by either complete or incomplete orthographic overlap with its translation equivalents in another language. L1 and L2 eye movements of Dutch-English bilinguals reading an entire novel were analyzed.

In L2 we found a facilitating effect of orthographic overlap. Additional identical cognate facilitation was found for later eye movement measures. This shows that the complex, semantic context of a novel does not eliminate cross-lingual activation in natural reading.

In L1 we detected non-identical cognate facilitation for first fixation durations for long nouns. Identical cognate facilitation was found for high frequent nouns for total reading times. This study is the first to show cognate facilitation in L1 reading of natural text: When reading a novel in the mother tongue, lexical access is not restricted to the target language.

These results support a two-morpheme, one-orthography view on identical cognate representation.

---

[9] This chapter is based on a manuscript co-authored by Eva Van Assche, Denis Drieghe and Wouter Duyck: Cop, U., Van Assche, E., Drieghe, D., & Duyck, W. (2014). Bilinguals reading a novel: An eye movement study of cognate facilitation in L1 and L2 natural reading, submitted for publication.

[10] Manuscript submitted for publication in Bilingualism: Language and Cognition. Revision in progress.

INTRODUCTION

Reading entails the identification of word forms, the retrieval of their meaning, and subsequently the integration of that meaning in the context of the sentence, paragraph or story. When a person has knowledge of two or more languages, an important question arises: Are words from these different languages co-activated during bilingual reading? A popular method to attempt to answer this question is to study responses to words that share orthography and/or meaning across the different languages of a bilingual. If the responses to these words are different than the responses to control words, this can be considered as evidence that words belonging to the non-target language were activated. These activated words can either inhibit, or facilitate the activation of orthographic forms and the subsequent mapping on semantic representations in the target language. Examples of words that share characteristics across languages are cross-lingual homographs. These words share the same orthography, but have different meanings. For example, the word *room* exists in Dutch and English, but means *cream* in Dutch. Dijkstra, Timmermans, and Schriefers (2000) tested Dutch-English bilinguals in a go/no-go task in which they had to press a button only if the presented word was an English word. Reaction times for inter-lingual homographs were slower than for control words. This suggests that the Dutch representation of the homograph was activated and interfered with the lexical access of the English word.

This theoretical question about co-activation is related to the question of how lexical items are stored in the bilingual lexicon. Van Heuven and Dijkstra (1998) provide evidence for non-selective lexical access and a shared bilingual lexicon, in which words from both languages are stored in an integrated manner, using an orthographic neighborhood manipulation. An orthographic neighbor is any word that differs by one letter from the target word, respecting the other letters' position. For example the Dutch word *tolk*, meaning *translator* in English, has the English word *toll* as a neighbor. In monolingual studies word identification and naming are sensitive to the number of within-language neighbors of that word (Grainger, O'Regan, Jacobs, & Segui, 1989; Snodgrass & Mintzer, 1993). Van Heuven et al.

(1998) reported orthographic neighborhood effects across languages. The recognition of exclusively English target words by Dutch-English bilinguals was slower when this target word had a larger number of orthographic neighbors in Dutch. This shows that words from a non-target language are activated during word recognition, which then compete for lexical selection with target-language representations. For an overview of the large number of studies providing evidence for language independent activation of words and a shared bilingual lexicon see chapter 4 in de Groot (2011).

COGNATE FACILITATION

Most studies investigating language non-selective activation of words have used cognates. Cognates are translation equivalent words that not only overlap in meaning but also in orthography. An example of an identical cognate, for which the orthographic overlap is complete, is the word “piano” in English and in Dutch. An example of a non-identical cognate is the Dutch word “tomaat”, of which the English translation equivalent is “tomato”. Identical and non-identical cognates are recognized faster and more accurately than control words in behavioral studies that present words in isolation, such as lexical decision tasks (e.g. Bultena, Dijkstra, & van Hell, 2013; Dijkstra, Grainger, & van Heuven, 1999; Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010; Peeters, Dijkstra, & Grainger, 2013), translation priming tasks (Davis et al., 2010; Sanchez-Casas, Davis, & García-Albea, 1992), or progressive demasking tasks (e.g., Dijkstra et al., 2010; Lemhöfer et al., 2008). In second language (L2) processing, cognate facilitation is larger than in native language (L1) processing (e.g.,Kroll, Dijkstra, Janssens, & Schriefers, 1999), although cognate facilitation has also been found in strict L1 contexts (e.g., Van Hell & Dijkstra, 2002).

These cognate facilitation effects for isolated visual word recognition are not necessarily informative about whether or not both languages of a bilingual are activated during reading in actual natural contexts (e.g., reading a newspaper). The fact that the sentence the word is embedded in is composed in one language might restrict lexical access to one language. Another reason to investigate cognate processing in a sentence context instead of in isolation is that most isolated-word methods entail a decision component. This component recruits processes that do not necessarily involve language processing, thus possibly disguising the actual effects reflecting lexical access in bilinguals.

A series of recent experiments have therefore explored and replicated cognate facilitation effects for target words in an L2 sentence context (Bultena, Dijkstra, & van Hell, 2014; Duyck, Van Assche, Drieghe, & Hartsuiker, 2007; Libben & Titone, 2009; Schwartz & Kroll, 2006; Van Assche, Drieghe, Duyck, Welvaert, & Hartsuiker, 2011; Van Assche, Duyck, & Brysbaert, 2013; van Hell & de Groot, 2008). This suggests that the mere representation of words in a sentence and the language cue that a sentence provides, do not restrict dual-language activation in the bilingual language system (e.g., Bultena et al., 2014; Duyck et al. 2007). These cognate effects are modulated by the predictability of the target word in the particular sentence context. When the sentence is of low constraint, comparable facilitation effects are found as in isolation studies (e.g. Schwartz & Kroll, 2006; Van Hell & de Groot, 2008). Mixed results have been found when the sentence context provides semantic constraints for lexical activation, but recent eye tracking studies have shown that a high semantic constraint does not necessarily eliminate cross-lingual activation in the bilingual language system (e.g., Van Assche et al., 2011), at least for early interaction effects reflected in early eye movement measures (e.g., Libben & Titone, 2009).

Very few studies have tested whether there can be cognate facilitation in an L1 sentence context (Titone, Libben, Mercier, Whitford, & Pivneva, 2011; Van Assche, Duyck, Hartsuiker, & Diependaele, 2009). Both of these eye tracking studies embedded target words in low constrain L1 sentences and found cognate facilitation. Titone et al. (2011) presented low and high-constraint L1 sentences to English-French bilinguals. Cognate facilitation was present on early reading time measures, but this effect was only present for bilinguals who acquired the L2 early in life. The L2 age of acquisition did not affect cognate facilitation effects on late reading time measures, but here semantic constraint did. Cognate facilitation was only found in low-constraint sentences: when the words were embedded in high constraint sentences, no cognate facilitation was found. Van Assche et al. (2009) also found non-identical cognate facilitation for Dutch-English bilinguals reading low constraint sentences in Dutch (L1). Orthographic overlap had a continuous effect on first fixation durations, gaze durations and go past times: words that shared more overlap with the translation equivalent were read faster. However, these few studies that find cognate facilitation in L1 sentence reading have a rather limited stimulus set: Van Assche et al. used 40 cognates with varying degree of orthographic overlap, whereas Titone et al. used 32 form-identical cognates.

Earlier studies investigating cognate effects in sentence context have often made discrete distinctions between identical cognates, non-identical cognates and control words (e.g., Duyck et al., 2007) or identical cognates and control words (Libben & Titone, 2009; Schwartz & Kroll, 2006). However, identical overlap in spelling is not required to facilitate the processing of cognates in L2 sentence contexts (e.g. Davis et al., 2010; van Hell & de Groot, 2008) or even L1 sentence contexts (Van Assche et al., 2009). Therefore, to fully understand cognate processing, it is necessary to investigate the influence of the gradual similarity between translation equivalent words. Two studies have explicitly tested the effect of degree of orthographic overlap of target words in L2 sentence contexts (Bultena et al., 2014; Van Assche et al., 2011). Both studies have shown continuous effects of orthographic overlap. If a target word has a larger overlap with its translation equivalent, it is read faster. Bultena et al. (2014) found facilitation for nouns only in go past times, while Van Assche et al. (2011) found cognate facilitation effects for early and late eye movement measures.

COGNATE REPRESENTATION

Despite the abundance of behavioral studies that report cognate facilitation effects, there is no consensus about the mechanisms leading to the easier processing of words with orthographically overlapping translation equivalents. An important issue here is to understand in which way cognate words are represented in the bilingual lexicon. Over the years, several theoretical accounts have been proposed.

According to one account, cognate facilitation can be framed within the BIA+ (Bilingual Interactive Activation Plus) model of visual word recognition (Dijkstra & van Heuven, 2002). The BIA+ assumes that L1 and L2 lexical items are stored in an integrated manner: orthographic, phonological and semantic representations of words are accessed in a language non-selective way. This model is the successor of the original BIA model (Dijkstra & Van Heuven, 1998), which is a bilingual adaptation of the Interactive Activation model (McClelland & Rumelhart, 1981). When a word is encountered, matching orthographic candidates are activated through bottom-up activation, dependent on their similarity to the printed word and their resting-level activation, determined by the subjective frequency. As L2 items tend to be lower in subjective frequency, their representations are activated somewhat slower than L1 items. The activation spreads from the orthographic candidates to the connected phonological and semantic representations. Every word in the lexicon is connected to one of the available language nodes, representing the language membership of that word. In BIA+, these nodes also represent the global lexical activity of a language. These nodes do not feed activation back to the orthographic or phonological level so they cannot function as a language selection mechanism. Instead, in order to account for differences in word recognition depending on tasks and other non-linguistic variables (instructions, expectations of the participants, ...) a task/decision system is proposed. See Figure 1 for a schematic representation of the BIA+ model.

Within the BIA+ framework, the combination of meaning and form overlap gives rise to the cognate facilitation effect. The degree of this cross-linguistic overlap will determine the amount of facilitation from these overlapping representations. For non-identical cognates, the input word will activate all lexical candidates, including the target representation and the form-overlapping cognate in the non-target language. For example: TOMATO in English will activate both the Dutch orthographic representation TOMAAT and the English correct orthographic representation TOMATO. The overlapping semantic representation of TOMAAT will facilitate the recognition of the target word TOMATO. As described above, cognate facilitation may indeed emerge when orthographic overlap is incomplete. On top of that the size of the facilitation effects depends on the cross-linguistic overlap: more overlap results in stronger cognate facilitation effects (Bultena et al., 2014; Dijkstra et al., 2010; Van Assche et al., 2011, 2009). For unbalanced bilinguals, another hypothesis arises. Because L1 lexical representations are used more often than L2 lexical representations, the higher resting activation for L1 items will result in larger cognate facilitation in L2 processing than in L1 processing.

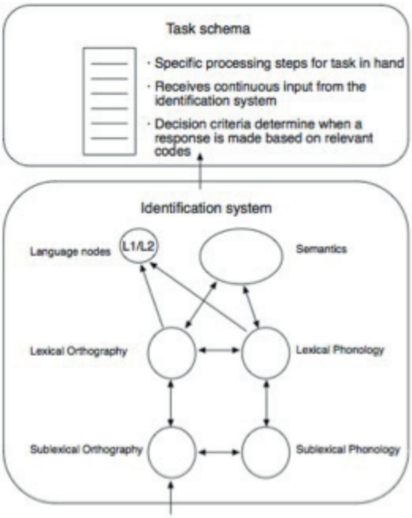


Figure 1. The architecture of the BIA+ model (Taken from Dijkstra and Van Heuven, 2002, page 183).

An important question is whether identical cognates are represented in the same way as non-identical cognates and whether they share one orthographic representation or instead have two distinct orthographic representations. The shared orthographic representation option would distinguish identical cognates from other translation equivalents and non-identical cognates, which by definition need to be represented twice in the bilingual lexicon, because of their orthographic difference. As Dijkstra et al. (2010) note, the BIA+ model leaves open the option for both possibilities. In case of two

distinct orthographic codes, it can be expected that at some point during the word recognition process, two lexical candidates receive the same amount of activation. In the BIA+ framework, this will cause lateral inhibition. When it is assumed that as identical cognates share one orthographic representation, this lateral inhibition will be absent. So in the latter case, an additional facilitation effect for identical cognates on top of the effect of orthographic overlap should be present. Following this reasoning, the facilitation for identical cognates should be detected relatively fast. The facilitation for non-identical cognates would take place later as an effect of shared semantic representation feedback (Dijkstra et al., 2010).
To be able to distinguish between the two possibilities Dijkstra et al. (2010) explicitly investigated the two distinct effects of identical cognate status and orthographic overlap. They tested Dutch-English bilinguals' performance on an L2 lexical decision task (Experiment 1, Dijkstra et al., 2010). They showed that both identical cognate status and orthographic overlap had a facilitating effect on reaction times. Larger orthographic overlap between a word and its translation equivalent yielded faster recognition. In addition to the expected facilitation by orthographic overlap, identical cognates showed much faster reaction times. This large discontinuous facilitation for identical cognates implies that identical cognates may be represented differently than non-identical cognates and may share one orthographic representation.

Another viewpoint on cognate representation assumes that cognates have a supra-lexical connection with their cross-language translation equivalents. This supra-lexical representation transcends each language specific lexicon (Cristoffanini, Kirsner, & Milech, 1986; Davis et al., 2010; Kirsner, Lalor, & Hird, 1993; Lalor & Kirsner, 2000; Sanchez-Casas et al., 1992; Sanchez-Casas & García-Albea, 2005). This idea is similar to Giraudo and Grainger's (2003) idea of a shared morphemic representation. More specifically, words that share an etymological root share a representation at the morphological level, located between the form and the lemma level. This cross-language connection for cognates could facilitate recognition. As a consequence, cognate facilitation should be sensitive to the cumulative frequency of the shared morphemic representation: Reaction times for cognates should be more affected by this cumulative frequency measure than by individual frequencies. However, it remains unclear what the necessary degree of form overlap should be to create such a cross-language connection or shared morpheme. This shared-morpheme view suggests that once this threshold of orthographic overlap necessary to create a shared morpheme is crossed, equal facilitation for all cognates should be found.

In contrast to the shared-morpheme explanation, Peeters et al. (2013) proposed a two-morpheme view in which identical cognates are represented by one orthographic representation and have two distinct language-specific morphological representations. The BIA+ architecture is used to explain the cognate facilitation effect, namely activation spreading from orthographic codes to other representations. This two-morpheme view allows for cognates to have different gender and plural information and a separate subjective frequency in the two languages. This account also explains the larger facilitation for identical cognates compared to non-identical cognates (Dijkstra et al., 2010). Peeters et al. provide evidence for two-separate rather than one morphological representation (e.g., Sánchez-Casas & García-Albea, 2005) by testing late French-English bilinguals on an L2 lexical decision task. Because cognates with a low L2 and high L1 frequency have a higher subjective cumulative frequency than those with high L2 and low L1 frequency, the shared-morpheme account predicts that the former words would be responded to faster in L2 than the latter. Peeters et al.'s results provided evidence against the shared-morpheme account: Cognates with a high L2 frequency and a low L1 frequency were processed more quickly than cognates with a low L2 frequency and a high L1 frequency. Peeters et al. claim that for late bilinguals, two separate morphological representations for identical cognates are plausible and might develop, because of the different learning contexts (class room vs. at home).

PRESENT STUDY

The first aim of the present study is to investigate whether cognate facilitation is restricted to reading of experimental materials, or whether it is strong enough to influence reading of continuous, meaningful natural text (i.e. a novel). Although the sentence reading studies described above (e.g., Duyck et al., 2007; Libben & Titone, 2009; Schwartz & Kroll, 2006; Titone et al., 2011; Van Assche et al., 2011; Van Hell & De Groot, 2008) tried to mimic the natural reading process by embedding target words in L1 or L2 sentences, this is still a very contrived situation. When people read in the real world, they usually read sentences that have their place in a larger meaningful whole, such as a novel or a newspaper. It may be possible that the more constrained semantic context of a sentence embedded in a larger text, reduces the cross-lingual activation causing cognate effects. Therefore it is possible that cognate effects observed in single sentences may be lab artifacts with little relevance for everyday reading. Also, the difference in goals people have for reading isolated sentences in an experimental setting compared to meaningful text, could elicit different reading strategies for these different contexts. There is indeed evidence that reading a continuous text or story is not the same as reading isolated sentences. Radach, Huestegge, and Reilly (2008) showed that the total reading time of words is longer for reading passages, but also the earlier eye movements are faster than when reading isolated sentences. Radach et al. explained this by suggesting that readers of passages of text perform a fast first pass across the text followed by a rereading of the passage. This may make it more difficult to detect cognate effects in natural reading.

There is only one study that did investigate cognate facilitation in a larger textual context (Balling, 2013), but only in L2. Balling (2013) instructed Danish-English bilinguals to read paragraphs of texts in their L2. Her results showed no clear cognate facilitation in first fixation durations, but in gaze duration and total reading time, morphologically simple words were read faster when they were cognates. This is indeed evidence for the relevance of cross-lingual interactions in reading, but in every day life we do not encounter solely monomorphemic words. On the contrary, most content words are morphologically complex.

We aim to replicate Balling's findings for L2 reading in a more extended and authentic semantic context, namely an entire novel. Also, we will try to extend these results to L1 reading. Cognate facilitation in L1 visual word recognition is usually smaller than in L2 reading and has not been reported very often (for a few exceptions, see van Hell & Dijkstra, 2002; Van Assche et al., 2009, Titone et al., 2011). The BIA+ model (Dijkstra & Van Heuven, 2002) indeed predicts smaller cognate facilitation effects in L1 versus L2.

On top of that, the question whether L1 cognate facilitation can be found in a naturalistic reading context has never been tested.

The second aim is to investigate the difference between the facilitation effects for identical and non-identical cognates, because this difference reveals how cognates might be represented in the bilingual brain. These two distinct effects were already investigated together for L2 word recognition in isolation by Dijkstra et al. (2010). As described above, L2 lexical decision times were faster as the orthographic overlap was larger, but an added drop in reaction times was found for identical cognates. We will investigate these effects simultaneously in a natural reading setting. If we find similar results, this offers evidence for the viewpoint that assumes that identical cognates are represented by one orthographic representation, while non-identical cognates have two separate orthographic codes (e.g., Dijkstra & Van Heuven, 2002; Peeters et al., 2013). If we do not replicate the additional drop in eye movement durations for identical cognates, this may be a task artifact, for instance due to the decision component that lexical decision entails.

Given the architecture of the BIA+ model (Dijkstra & Van Heuven, 2002), we expect to find early facilitation effects for identical cognates and later effects of non-identical cognate facilitation. The absence of lateral inhibition on the orthographic level for identical cognates sharing an orthographic representation would cause a very early difference to arise between these and other words. For non-identical cognates, two lexical orthographic representations are activated that initially inhibit each other via lateral inhibition, but leading to later cognate facilitation effects via semantic resonance.

METHOD

This method section is partly taken from Cop, Drieghe, and Duyck (2015) because the data in this analysis is a subset from a large eye movement corpus described in Cop et al.

PARTICIPANTS

Nineteen unbalanced bilingual Ghent University undergraduates participated either for course credit or monetary compensation. The participants' dominant language was Dutch and their second language was English. The participants were all advanced L2 learners with a relatively late L2 age of acquisition (mean=11 [2.46]). All have had formal education of English in the Belgian school system from age 12 or 13. Participants completed a battery of language proficiency tests including a Dutch and English spelling test (GLETSHER and WRAT4), the LexTALE (Lemhöfer & Broersma, 2011) in Dutch and English, a Dutch and English lexical decision task and a self-report language questionnaire (based on the LEAP-Q, Marian, Blumenfeld, & Kaushanskaya 2007). We calculated a composite L1 and L2 proficiency score by averaging the score on the spelling test, the score on the LexTALE and the adjusted score of the lexical decision task. This composite score and the LexTALE scores shows that bilinguals score significantly higher on L1 proficiency than they do on L2 proficiency. The LexTALE score is used in the analysis. Participants had normal or corrected-to-normal vision. None of the participants reported to have any language and/or reading impairments. For detailed scores on all proficiency measures see Table A.1 in Appendix A.

APPARATUS

The bilingual eye movement data were recorded with the EyeLink 1000 system (SR-Research, Canada) with a sampling rate of 1 kHz. Reading was binocular, but eye movements were recorded only from the right eye. Text was presented in black 14 point Courier New font on a light grey background. The lines were triple spaced and 3 characters subtended 1 degree of visual angle or 30 pixels. Text appeared in paragraphs on the screen. A maximum of 145 words was presented on one screen. During the presentation of the novel, the room was dimly illuminated.

MATERIALS

The participants were asked to read the novel “The mysterious affair at Styles” by Agatha Christie (Title in Dutch: “De zaak Styles”). This novel was selected out of a pool of books that was available via the Gutenberg collection. The books were judged on length and difficulty, indicated by the frequency distribution of the words that the book contained. We selected the novel whose word frequency distribution was the most similar to the one in natural language use (Subtlex database). The Kullback–Leibler divergence was used to measure the difference between the two probability distributions (Cover & Thomas, 1991). See Table 1 for characteristics of the target nouns in the novel. Both word frequency and word length show minor differences across languages, these variables will be included in the higher order interactions in our linear mixed model to ensure statistical control of these predictors that were not experimentally controlled.

Table 1. *Summary of the characteristics of all of the nouns and the identical cognates in the novel. Standard deviations (SD) are in brackets.*

	Nouns		Identical Cognates	
	Dutch	English	Dutch	English
Number of Words	7 988	7 640	548	675
Number of Unique Words	1 776	1 742	94	142
Average Word frequency	3.16 [1.02]	3.36 [0.99]	3.10 [1.21]	3.08 [1.20]
Average Word Length	6.68 [2.73]	5.94 [2.30]	5.63 [2.48]	5.81 [2.51]
Average Orthographic Overlap	0. 36[0.30]	0.38[0.32]	1 [0]	1 [0]

Only the nouns were selected for the current analyses. Because we used the authentic text of a novel, we had to assess the cognate status of the nouns because we did not select our stimuli on the basis of their cognate status. Balling (2013) already showed that the cognate status of words in sentence context must be evaluated relative to the context-appropriate translation equivalent. For example, when the Dutch word *arm* is placed in a context as a noun, it is a cognate with the English word *arm,* whereas when it is placed in a Dutch context as an adjective, it means *poor* in English and can no longer be considered a cognate.

We therefore manually assessed all the possible appropriate translations in context for each noun in the novel. We then selected the translation that was orthographically closest to the target word. When this translation was orthographically identical to the target word it was classified as an identical cognate. For all translation pairs we calculated the corrected Levenshtein distance (Schepens, Dijkstra, & Grootjen, 2012). For the formula see Appendix B. This variable was used as a measure for continuous orthographic overlap in our analyses. For the frequency distribution of orthographic overlap and some examples of translation pairs see Figure 2.

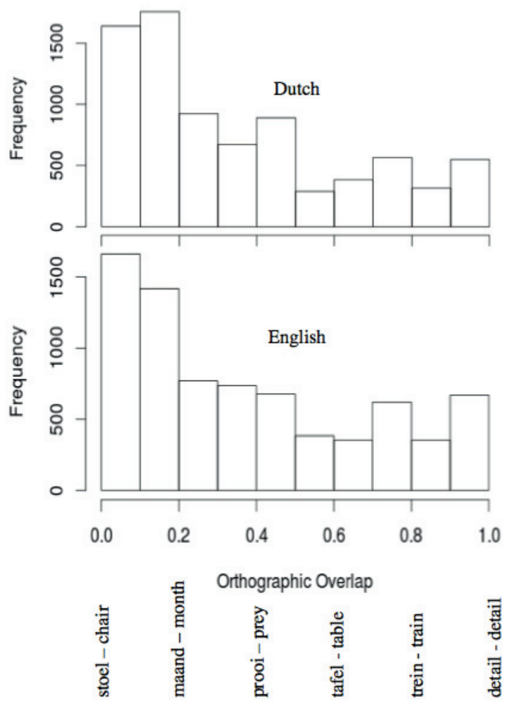


Figure 2. The frequency distribution of orthographic overlap for all nouns in the Dutch and English version of the novel. Examples of translation equivalent pairs dependent on orthographic overlap are given below the graphs.

PROCEDURE

Each participant read the entire novel in four sessions of an hour and a half. They read half of the novel in Dutch, the other half in English. The order was counterbalanced.

The participants were instructed to read the novel silently while the eye tracker recorded their eye movements. It was stressed that they should move their head and body as little as possible while they were reading. The participants were informed that they would be presented with multiple-choice questions about the contents of the book after each chapter. This was done to ensure that participants understood what they were reading and paid attention throughout the session.

The text of the novel appeared on the screen in paragraphs. A maximum of 145 words were presented on the screen in one trial. When the participant finished reading the sentences on one screen, they pressed the appropriate button on a control pad to move to the next part of the novel. After each chapter, multiple-choice questions were given to the participant. Participants were given the choice to pause for a maximum of 10 minutes after each chapter.

Before starting the practice trials, a nine-point calibration was executed. After this, the calibration was done every 10 minutes, or more frequently when the experiment leader deemed necessary.

ANALYSES

We analyzed four eye movement measures that reflect early language processes such as initial lexical access: a) First fixation duration, the duration of the first fixation on the target noun the first time they land on it; b) Single fixation duration, first pass fixation duration on a word that is fixated exactly once; c) Gaze duration, the sum of all fixation durations during first passage before the eyes move out of the word and, d) probability of first pass skipping of a word. We analyzed two eye movement measures of reading times of the nouns that reflect later, higher-order, language processes such as semantic integration: a) Go past time, the sum of all fixation durations on the target word including all of the regressions to previous words until the eyes move rightward from the target word; b) Total reading time, the sum of all fixation durations on the target word, including refixations. Fixations shorter than 100ms were excluded from the dataset (Rayner, 1998).

Reading time measures and skipping probabilities were fitted in (generalized) linear mixed models using the lme4 package (version 1.1-7) of R (version 3.0.2). All of the initial models contained the fixed factors of Language (L1 or L2) and Cognate status (Identical Cognate or not) and the covariates Orthographic Overlap (continuous), L1 proficiency (continuous) and L2 Proficiency (continuous) and the control variables of word frequency (continuous) and word length (continuous). As proficiency variables we used the score on the L1 and L2 LexTALE (Lemhofer et al.). For the word frequency, the subtitle word frequency measures (English: Brysbaert & New 2009; Dutch: Keuleers, Brysbaert & New, 2010) were log transformed to normalize their distribution. To reduce collinearity, all continuous predictors were centered.

We included a random intercept per subject in all initial models. This ensured that differences between subjects concerning genetic, developmental or social factors were modeled. We also included a random intercept per word, to be able to generalize to other nouns, because our stimuli sample is not an exhaustive list of all nouns in a language. The models were fitted using restricted maximum likelihood estimation (REML). First a full model, including the two random factors and all of the 2-and 3-way interactions between the fixed effects, was fitted. The optimal model was discovered by backward fitting of the fixed effects, then forward fitting of the random effects and finally again backward fitting the fixed effects (Barr, Levy, Scheepers & Tily, 2013). We strived to include a maximal random structure in the final models (Barr et al., 2013).

RESULTS

For an overview of the fitted values for the effect of identical cognate status and orthographic overlap of the final models see Table 2.

FIRST FIXATION DURATION

First Fixation durations that differed more than 2.5 standard deviations from the subject means per language (2.15% for Dutch, 2.21% for English) were excluded. This left us with 87 980 data points. The dependent variable was log transformed to normalize the distribution as suggested by the Box-Cox method (Box & Cox, 1964) The outcome of the final model for first fixation duration is presented in Table 3. The maximum correlation between fixed effects in the final model was -.692 for L1 and L2 proficiency.

A significant main effect of language was found. First fixations on nouns were longer in L2 (226.3ms) than in L1 (212.0ms). We also found an effect of orthographic overlap: Target words with larger orthographic overlap with their translation equivalents yielded shorter first fixation durations (see Figure 3). This variable did not interact with language, indicating a comparable cognate facilitation effect in L1 and L2.

Because of the importance of this finding, we fitted additional models to test for the significance of the effect of orthographic overlap for each language separately. For L2 reading, the effect of orthographic overlap was significant (β=-0.018, sd=0.0086, t=-2.13). For L1 reading this effect was only marginally significant (β=-0.011, sd=0.0074, t=-1.50). The interaction of orthographic overlap with word length was also marginally significant (β=-0.0052, sd=0.0034, t=-1.54). Planned contrasts showed that for words of 9 characters or longer, there is a facilitating effect of orthographic overlap (χ^2=3.96, df=1, p-value<0.05).

The effect of identical cognate status was not significant. This means that there is no additional facilitation for identical cognates compared to non-identical cognates that cannot be explained by a linear decrease in fixation duration due to the increase in orthographic overlap.

Table 2. *Mean fitted values and difference score for First Fixation duration, single fixation duration, gaze duration, total reading time, go past time and skipping rate for the effect of identical cognate status and orthographic overlap.*

		First Fixation Duration		Single Fixation Duration		Skipping Rate		Gaze Duration		Total Reading Time		Go past Time	
		L1	L2	L1	L2	L1	L2	L1	L2	L1	L2	L1	L2
Effect of identical cognate status	Cognate	203.6	217.0	207.1	218.1	27.4	26.4	213.4	228.4	230.1	256.8	252	277.1
	Other	203.0	214.9	205.1	219.8	29.1	26.0	212.1	230.9	234.5	265.1	251	291.2
	Difference	0.1	2.1	1.9	- 1.7	1.7	0.4	1.3	-2.5 *	-4.4 *	-8.2 *	1.0	-14.0*
Effect of orthographic overlap	Max= 1	202.0	213.1	203.8	218.5	28.6	27.3	210.7	230.0	233.3	263.0	253.4	298.4
	Min=0	203.7	216.1	206.0	220.2	29.2	25.4	212.8	230.9	234.6	265.0	250.0	286.0
	Difference	-1.7 *	-2.9*	-2.2	-1.7	-0.7	1.9 *	-2.2	-0.9	-1.3	-2	3.4	12.4

* t-value >1.96 in final model

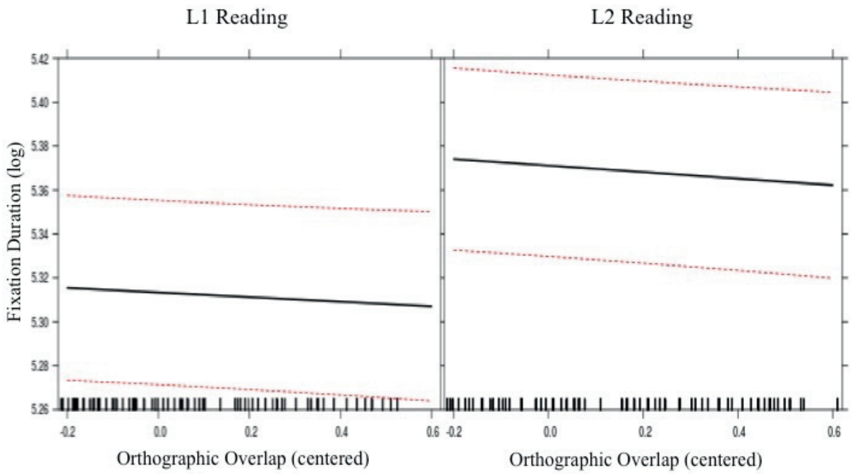


Figure 3. Effect of orthographic distance on first fixation duration (log transformed) for L1 and L2 reading. The 95% Confidence Intervals (CI's) are depicted as dotted lines.

Table 3. *Estimates, standard errors (SE) and t-values for the fixed effect and variance and SD's for the random effects for the final model for first fixation durations.*

First Fixation duration			
	Estimate	SE	t-value
Fixed Effects			
(Intercept)	5.31	0.020	60.80
Cognate Status	0.0068	0.0074	.92
Orthographic Overlap	0.013	0.0057	2.19
Language	0.069	0.0076	.17
L1 proficiency	0.0058	0.0043	1.35
L2 proficiency	0.0022	0.0021	.03
Control variables			
Word Length	0.0050	0.0010	.96
Word Frequency	0.015	0.0027	5.72
Language * Word Frequency	0.0089	0.0028	3.15

	Variance	SD
Random Effects		
Word		
(Intercept)	0.0020	0.045
Subject		
(Intercept)	0.0077	0.088
Language	0.00088	0.030
Word Frequency	0.000037	0.0061
Word Length	0.000010	0.0032

SINGLE FIXATION DURATIONS

Single Fixation durations that differed more than 2.5 standard deviations from the subject means were excluded per language (2.17% for Dutch, 2.22% for English). This left us with 61 860 data points. The dependent variable was log transformed to normalize the distribution as suggested by the Box-Cox method (Box & Cox, 1964). The outcome of the final model for single fixation duration is presented in Table 4. The maximum correlation between fixed effects in the final model was -.692 for L1 and L2 proficiency.

Table 4. *Estimates, SE's and t-values for the fixed effect and variance and SD's for the random effects for the final model for Single fixation durations.*

Single Fixation duration			
	Estimate	SE	t-value
Fixed Effects			
(Intercept)	5.36	0.024	226.47
Cognate Status	-0.00094	0.0041	-0.23
Orthographic Overlap	-0.0093	0.0064	-1.44
Language	0.042	0.0043	9.96
L1 proficiency	-0.0032	0.0042	-0.76
L2 proficiency	0.0019	0.0021	0.91
Control variables			
Word Length	0.010	0.0016	6.31
Word Frequency	-0.030	0.0036	-8.21
		Variance	SD
Random Effects			
Word			
(Intercept)		0.0021	0.045
Subject			
(Intercept)		0.010	0.10
Language		0.00028	0.017
Word Frequency		0.00017	0.013
Word Length		0.000036	0.0060

A significant main effect of language was found. Single fixations on nouns were longer in L2 (235.6ms) than in L1 (218.4ms).

Neither the effect of orthographic overlap, nor the effect of identical cognate status reached significance.

GAZE DURATION

Gaze durations that differed more than 2.5 standard deviations from the subject means were excluded per language (2.44% for Dutch, 2.45% for English). This left us with 87 643 data points. The dependent variable was transformed with the Box-Cox transformation (1) to normalize the distribution (Box & Cox, 1964). The value for lambda was set at -0.5.

$$y_{transformed} = \frac{y^{-0.5} - 1}{-0.5} \tag{1}$$

The outcome of the final model for gaze durations is presented in Table 5. The maximum correlation between fixed effects in the final model was -.691 for L1 and L2 proficiency.

Table 5. *Estimates, SE's and t-values for the fixed effect and variance and SD's for the random effects for the final model for gaze durations.*

Gaze duration			
	Estimate	SE	t-value
Fixed Effects			
(Intercept)	1.87	0.0019	1004.8
Cognate Status	-0.000027	0.00032	-0.1
Orthographic Overlap	0.00053	0.00048	-1.1
Language	0.0034	0.00039	8.9
L1 proficiency	-0.000092	0.00024	-0.4
L2 proficiency	0.000011	0.00012	0.1
Cognate Status * Language	-0.00023	0.00017	-1.4
Control variables			
Word Length	0.0014	0.00019	7.1
Word Frequency	-0.0022	0.00024	-8.9
Cognate Status * Word Length	0.00011	0.00012	0.9
Language * Word Length	0.00032	0.000064	5.0
Cognate Status * Language * Word Length	0.00013	0.000063	2.1

	Variance	SD
Random Effects		
Word		
(Intercept)	0.000018	0.0042
Subject		
(Intercept)	0.000063	0.0080
Language	0.0000022	0.0015
Word Frequency	0.00000068	0.00082
Word Length	0.00000039	0.00062

A significant main effect of language was found: first pass reading was faster in L1 (234.7ms) than in L2 (260.8ms). The effect of orthographic overlap was not significant. The main effect of identical cognate status was not significant but a 3-way interaction between language, identical cognate status and word length was found. Separate analyses per language revealed that there was no effect of identical cognate status in L1 (β=0.00048, sd=0.00092, t=0.5) and no interaction with word length (β=-0.00016, sd=0.00033, t=-0.5). In L2, a marginal significant interaction effect of identical cognate status and word length was found (β=0.00051, sd=0.00029, t=1.7).

Although the main effect of identical cognate status did not reach significance for L2 reading (β=-0.00083, sd=0.00080, t=-1.0), this marginal interaction indicated that identical cognates were read faster than other nouns (see Figure 4). Planned contrasts showed that gaze durations of nouns of 4 characters or less were facilitated when the target noun was an identical cognate (χ^2=3.19, df=1, p-value=0.074), but this effect only reached full significance when the target noun was 2 characters or less (χ^2=3.85, df=1, p-value$<$0.05).

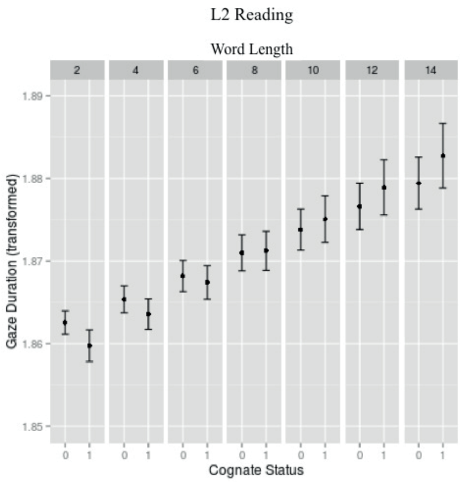


Figure 4. Effect of identical cognate status (1=Identical Cognate, 0=Other) on gaze durations (transformed) dependent on word length for L2 reading. The 95% CI's are depicted as whiskers.[11]

Probability of skipping

For skipping probability a logistic linear mixed model was fitted with a binary dependent variable. We analyzed 116 695 observations. The outcome of the final model for skipping probabilities is presented in Table 6. The maximum correlation between fixed effects in the final model was -.435 for L1 and L2 proficiency.

We found a main effect of language. Bilinguals skipped nouns more often when reading in L1 (32.5%) than when reading in L2 (30%). We did not find a main effect of orthographic overlap. The interaction of orthographic overlap and word length was significant and the interaction between orthographic overlap and language was significant. Separate analyses showed that for L1 reading the effect of orthographic overlap was not

[11] Note that the CI's on the graph are not informative for the significance of the effect, since the data are not independent

significant (β=-0.036, sd=0.052, t=-0.70), neither was the interaction of this variable with word length (β=-0.020, sd=0.020, t=-1.02). In L2, the effect of Orthographic Overlap was significant (β=0.11, sd=0.049, t=2.15) and interacted with Word Length (β=-0.040, sd=0.020, t=-2.02) (see Figure 5). Planned contrasts showed that nouns shorter than 6 characters were skipped more often when they had a larger orthographic overlap (χ^2=4.27, df=1, p-value <.05).

Table 6. *Estimates, SE's and t-values for the fixed effect and variance and SD's for the random effects for the final model for skipping rates.*

Skipping Rate			
	Estimate	SE	z-value
Fixed Effects			
(Intercept)	-0.77	0.10	-7.62
Cognate Status	-0.12	0.059	-1.98
Orthographic Overlap	-0.037	0.051	-0.74
Language	-0.37	0.075	-4.91
L1 proficiency	4.26	1.24	3.45
L2 proficiency	-2.19	0.75	-2.93
Cognate Status * Language	0.13	0.071	1.82
Orthographic Overlap * Language	0.13	0.070	1.80
Control variables			
Word Length	-0.25	0.01	-24.32
Word Frequency	0.15	0.017	8.68
Orthographic Overlap * Word Length	-0.03	0.014	-2.12
Language * Word Length	0.018	0.0084	2.20
L2 Proficiency * Word Length	-0.23	0.074	-3.06

	Variance	SD
Random Effects		
Word		
(Intercept)	0.038	0.19
Subject		
(Intercept)	0.19	0.43
Language	0.096	0.31
Word Frequency	0.0031	0.055
Word Length	0.0013	0.036

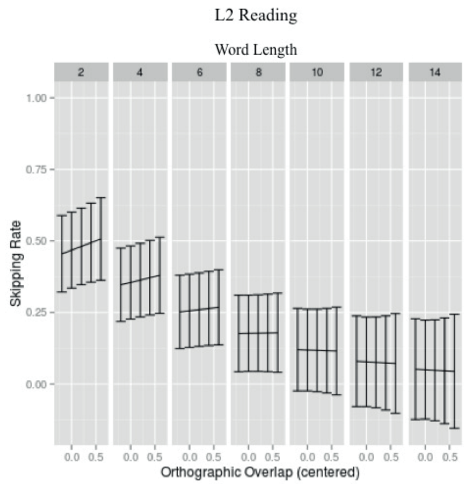


Figure 5. Effect of orthographic distance (centered) on skipping rates for L2. The 95% CI's are depicted as dotted lines.[12]

We found a main effect of identical cognate status and the interaction of identical cognate status with language was marginally significant. Separate analyses showed no significant effect of identical cognate status in either language (L1: β=-0.086, sd=-0.063, t=-1.36; L2: β=0.019, sd=0.056, t=0.34).

We also found a significant effect of L1 proficiency on skipping rates. When L1 proficiency scores were higher, participants are more likely to skip words. The effect of L2 proficiency was also significant. When L2 proficiency scores were higher, participants were less likely to skip a noun.

TOTAL READING TIMES

Total reading times that differed more than 2.5 standard deviations from the subject means were excluded per language (2.83% for Dutch, 2.82% for English). This left us with 87 348 data points. The dependent variable was transformed using the Box-Cox transformation (1) to normalize the distribution (Box & Cox, 1964). The value for lambda was set at -0.5.

[12] Note that the CI's on the graph are not informative for the significance of the effect, since the data are not independent.

The outcome of the final model for total reading times is presented in Table 7. The maximum correlation between fixed effects in the final model was -.693 for L1 and L2 proficiency.

Table 7. *Estimates, SE's and t-values for the fixed effect and variance and SD's for the random effects for the final model for total reading times.*

Total Reading Time			
	Estimate	SE	t-value
Fixed Effects			
(Intercept)	1.87	0.0018	1022.2
Cognate Status	-0.00068	0.00096	-0.7
Orthographic Overlap	-0.00047	0.00054	-0.9
Language	0.0097	0.00090	10.8
L1 proficiency	0.000061	0.00023	0.3
L2 proficiency	0.000088	0.00011	0.8
Cognate Status * Language	-0.0012	0.00081	-1.5
Control variables			
Word Length	0.0015	0.00017	8.9
Word Frequency	-0.0026	0.00033	-7.7
Cognate Status * Word Length	-0.00075	0.00042	-1.8
Cognate Status * Word Frequency	-0.0019	0.00086	-2.2
Language * Word Length	0.00017	0.00014	1.2
Language * Word Frequency	-0.0014	0.00041	-3.4
Cognate Status * Language * Word Length	0.0010	0.00043	2.3
Cognate Status * Language * Word Frequency	0.0021	0.00094	2.2
	Variance		SD
Random Effects			
Word			
(Intercept)	0.000023		0.0048
Subject			
(Intercept)	0.000062		0.0079
Language	0.000013		0.0036
Word Frequency	0.00000095		0.00098
Word Length	0.00000043		0.00065
Language * Word Frequency	0.00000076		0.00087

A main effect of language was found. Total reading times on nouns were on average shorter in L1 (270.9ms) than in L2 (312.7ms). The main effect for identical cognate status was not significant, but the 3-way interaction with language and word frequency and the 3-way interaction with language and word length were. Separate analyses per language showed that the effect of identical cognate status was not significant in L1 (β=-0.0012, sd=0.0012, t=-1.0). The interaction of identical cognate status and word frequency was significant (β=-0.0022, sd=0.00095, t=-2.3) (see Figure 6). Planned contrasts showed that for high frequent words (> 4.6 log word frequency) there was identical cognate facilitation for total reading times (χ^2=3.92, df=1, p-value<0.05). For very low frequent nouns (<0.8 log word frequency) we found identical cognate inhibition (χ^2=3.94, df=1, p-value<0.05).

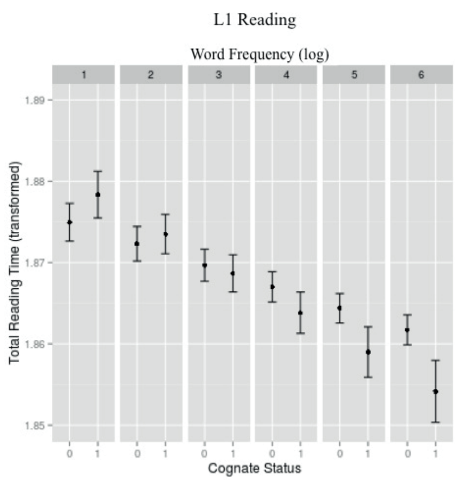


Figure 6. Effect of cognate status (1=Identical Cognate, 0=Other) on total reading time (transformed) for L1 reading dependent on word frequency (log-transformed). The 95% CI's are depicted as whiskers.[13]

The effect of identical cognate status was significant in L2 (β=-0.0020, sd=0.00087, t=-2.3). For L2 reading, identical cognates were read faster in total reading times (307ms) than other words were (313ms) (See Figure 7). The interactions of identical cognate status with word length and word frequency were not significant for L2 reading.

[13] Note that the CI's on the graph are not informative for the significance of the effect, since the data are not independent.

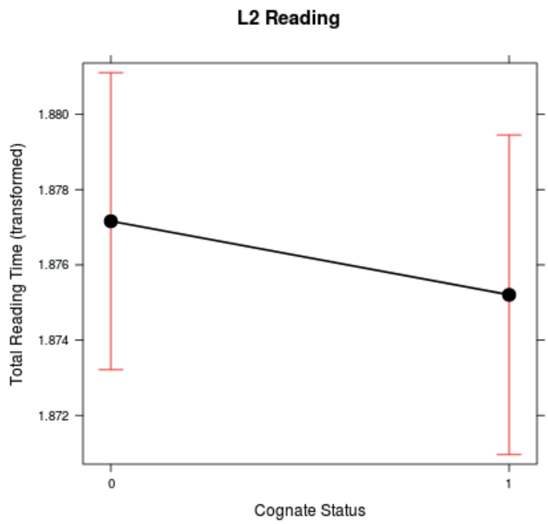


Figure 7. Effect of cognate status (1=Identical Cognate, 0=Other) on total reading time for L2 reading. The 95% CI's are depicted as whiskers.[14]

GO PAST TIMES

Go past times that differed more than 2.5 standard deviations from the subject means were excluded per language (2.38% for Dutch, 2.36% for English). This left us with 87 799 data points. The dependent variable was transformed using the Box-Cox transformation (1) to normalize the distribution (Box & Cox, 1964). The value for lambda was set at -0.5.

In Table 8 the outcome of the final model for go past times can be found. The maximum correlation between fixed effects in the final model was -.622 for L1 and L2 proficiency,

14 Note that the CI's on the graph are not informative for the significance of the effect, since the data are not independent. The point estimation of the difference of reading Identical Cognates vs. other words in L2 is 0.00196. The 95% CI for this difference is [0.000324; 0.00359].

Table 8. *Estimates, SE's and t-values for the fixed effect and variance and SD's for the random effects for the final model for go past times.*

Go Past Time			
	Estimate	SE	t-value
Fixed Effects			
(Intercept)	1.87	0.0021	910.5
Cognate Status	0.0020	0.00090	2.2
Orthographic Overlap	-0.00055	0.00059	-0.9
Language	0.011	0.00085	12.5
L1 proficiency	-0.012	0.038	-0.3
L2 proficiency	0.0032	0.017	0.2
Cognate Status * Language	-0.0042	0.00079	-5.3
Language * L1 Proficiency	-0.033	0.010	-3.1
Control variables			
Word Length	0.0012	0.00020	6.4
Word Frequency	-0.0032	0.00032	-10.0
Language * Word Length	0.00060	0.00014	4.3
L1 Proficiency * Word Frequency	0.0099	0.0044	2.2
		Variance	SD
Random Effects			
Word			
(Intercept)		0.000029	0.0054
Subject			
(Intercept)		0.000079	0.0089
Language		0.000011	0.00085
Word Frequency		0.0000012	0.0011
Word Length		0.00000059	0.00077
Language * Word Length		0.00000012	0.00034

A main effect of language was found. Go past times were on average shorter in L1 (305.6ms) than they were in L2 (365.8ms).

We found a main effect of identical cognate status. Identical cognate status interacted with language. Separate analyses revealed no effect of identical cognate status in L1 (β=0.0015, sd=0.0011, t=1.4). In L2 there was indeed a main effect of identical cognate status (β=-0.0020, sd=0.00096, t=-2.1). Identical cognates had a shorter go past time (352ms) than other nouns (367ms) (See Figure 8). This indicates that during regressions identical cognates were looked upon for a shorter amount of time than non-identical cognates.

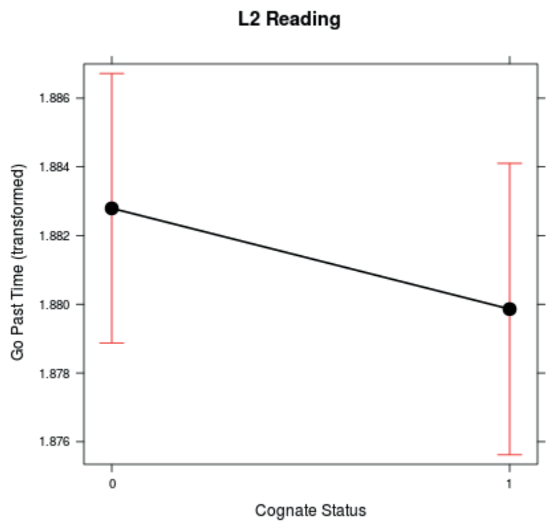


Figure 8. Effect of cognate status (1=Identical Cognate, 0=Other) on go past times for L2 reading. The 95% CI's are depicted as whiskers.[15]

The effect of orthographic overlap was not significant. The interaction between L1 proficiency and language was also significant. Separate analyses per language showed that L1 proficiency didn't reach significance in either language (L1: β=-0.00011, sd=0.00042, t=-0.3; L2: β=-0.00040, sd=0.00035, t=-1.1).

DISCUSSION

We studied the effect of identical cognate status and orthographic overlap for translation equivalent nouns in a naturalistic reading context. The eye movements of late Dutch-English bilinguals who read an entire novel in L1 and L2 were analyzed.

We found cognate facilitation in early and late eye movement measures in both L1 and L2 reading.

[15] Note that the CI's on the graph are not informative for the significance of the effect, since the data are not independent. The point estimation of the difference of go past times for Identical Cognates vs. other words in L2 is 0.00196. The 95% CI for this difference is [0.000414; 0.00350].

L2 COGNATE FACILITATION

The analyses of the early reading measures showed clear cognate facilitation effects for reading in L2. First fixation durations were facilitated by cross-lingual orthographic overlap. Because we only found an effect for first fixation durations, not for single fixation durations or gaze durations, this effect was driven exclusively by the first fixation landing on the target noun: when reading in second language, a word with more cross-lingual orthographic overlap will elicit a shorter first fixation. Additionally, skipping probabilities for short words were higher when the orthographic overlap was higher. These results show that nouns were more likely to not receive a fixation on first pass reading when the orthographic overlap with their translation equivalent became larger, and when they did lexical access was faster for words with a larger orthographic overlap. The failure to find this effect for words with seven letters or more, might be the result of a low skipping rate for longer words (8%), causing floor effects and also reduced parafoveal processing for the final letters of these words preventing a more thorough lexical processing.

For first fixation durations and skipping rates, we did not find an additional effect of identical cognates: words that have complete orthographical overlap across languages were not processed faster than would be expected if the effect was only due to orthographic similarity.

In the later reading measures under investigation, identical cognate facilitation was found for gaze durations, total reading times and go past times. This means that for later word processing, only identical cognates are read faster than control words. Non-identical cognates are not. Given that regressions indicate semantic integration difficulties, the results for go past times imply that identical cognates read in L2 are easier to integrate in the larger semantic context.

The findings for gaze durations and total reading times replicate and extend those of Balling's (2013) L2 paragraph reading experiment. While they only found cognate facilitation for morphologically simple words, we found it for a set of complex and simple words. On top of that we found cognate facilitation for earlier measures, namely first fixation durations and skipping rates.

L1 COGNATE FACILITATION

For the early eye movements under investigation, we detected cognate facilitation in L1 reading. Orthographic overlap of the L1 target noun with its L2 translation equivalent shortened the first fixation duration on the target nouns for words of 9 characters or longer. We found no additional facilitation when the target noun was an identical cognate.

These results indicate that cognate facilitation is detectable in L1 natural reading during the earliest stages of word recognition. This finding is compatible with results found in single L1 sentence contexts (Titone et al., 2011; Van Assche et al., 2009). In these experiments the target nouns were presented without the larger and much more complex semantic context that is present when reading a novel. In this way this study goes beyond the previous findings on cognate processing in L1 and provides compelling evidence for cross-lingual interaction in early L1 language processing.

As we predicted, the effect size of the cognate facilitation was rather small. The difference between the fitted value for first fixation duration for words with the smallest and the highest orthographic overlap for L1 reading was only 1.7ms. Van Assche et al. (2009) also report a rather small effect size for cognate facilitation (5ms) for first fixation durations in L1 reading. Titone et al. (2011) report a small cognate facilitation effect of about 1-3ms for first fixation durations. The size of the effects is probably partly due to floor-effects because the first fixation durations are rather short (212ms on average in L1 reading). The interaction we found with word length is compatible with this idea: for longer words, for which the first fixation durations were longer (219ms) than for short words (209ms), the effect of orthographic overlap became fully significant.

For single fixation durations, gaze durations and skipping rates, no cognate facilitation was found. This is compatible with Titone et al.'s (2011) findings. Van Assche et al. (2009) do find cognate facilitation for gaze durations, but the robustness of that effect appeared to be unreliable: when they replicated this study with a different stimulus set, the effect for gaze durations became marginally significant.

In later eye movement measures we detected identical cognate facilitation for total reading times, but not for go past times. Total reading times were shorter for high frequent nouns, when this noun was an identical cognate. This is in part compatible with Titone et al.'s (2011) and Van Assche et al's (2009) results. Titone et al. found identical cognate facilitation for total reading times and go past times in an L1 low constraint sentence context, but not in a high constraint sentence context. Van Assche et al. found identical and non-identical cognate facilitation for go past times in low constraint sentences. They did not analyze total reading times.

In our experiment, participants read a book. This means that a lot of the presented nouns were highly constrained by the sentence context, or by the context of the novel as a whole. As we described in the introduction, cognate facilitation is harder to detect in this kind of sentences but it represents a more realistic equivalent of normal reading. These results show that late L1 cognate facilitation can be found even in contextually constrained contexts, contrary to Titone et al's (2011) results.

Interestingly, we detected identical cognate inhibition on total reading times when the cognates were of very low frequency. Although we did not predict this, it could be the case that for those low frequent nouns the translation equivalent in L2 is unknown to the participant and so even though the noun has an orthographically identical translation, this noun does not function as a cognate. For example the word 'legaat' in Dutch, 'legate' in English, has a Dutch log word frequency of 0.30 and a Levenshtein distance of 0.66 or the identical cognate 'begonias' has a Dutch log word frequency of 0.30. It is very plausible that these words are not known in the second language of the participants. It is still a question why these nouns would elicit inhibition compared to the control nouns. Future research efforts should be made to determine the right conditions for cognate inhibition detection.

To sum up, for the first time, early and late cognate facilitation has been found in L1 natural reading, without the use of a restricted, contrived set of low constraining sentences but with bilinguals reading a real novel containing a large diversity of semantic contexts. This shows that a bilingual reading in his or her most dominant and first-acquired language is influenced by knowledge of translations in another language.

COGNATE REPRESENTATION

The present study showed continuous effects of orthographic similarity, with more cross-lingual orthographic overlap leading to faster first fixation durations in L1 and L2 and higher skipping rates in L2. Also, effects in L2 are larger and were present across more eye movement measures than effects in L1. These results are in line with models, which assume that cognate facilitation arises from converging cross-lingual lexical activation from activated lexical candidates (e.g., Dijkstra & Van Heuven, 2002; Midgley et al., 2011). Considering the BIA+ architecture (Dijkstra & Van Heuven), non-identical cognates and other translation equivalents have two distinct language-specific orthographic representations connected to a shared semantic representation. For translation equivalents with some orthographic overlap, the cross-lingual activation of similar orthographic representations results in more activation spreading to the same semantic representation. The more orthographic overlap between the written target word and the translation equivalent, the more activation spreads towards the shared semantics. This mechanism might explain the linear non-identical cognate facilitation we found in the current study. Viewpoints that assume qualitative differences such as differences at a morphological level between cognates and non-cognates (e.g., Sánchez-Casas & García-Albea, 2005) could not account for these continuous effects of orthographic overlap.

For later eye movement variables; gaze durations, total reading times and go past times; we found facilitation only for identical cognates. This means that later in the reading process there is a processing difference for identical cognates compared to all other nouns, regardless of the orthographic overlap between them and their translation equivalent. To explain these kinds of effects, it has been proposed that identical cognates may be represented differently from other words at the lexical level (Dijkstra et al., 2010). Peeters et al. (2013) suggested that identical cognates share one orthographic representation (e.g. Dijkstra & Van Heuven, 2002; Dijkstra et al., 2010; Midgley et al., 2011), and thus lack the lateral inhibition on that level.

This architecture seems to predict that both of the effects should be present at the same time: there should be additional facilitation for identical cognates compared to non-identical cognates, on top of orthographic overlap effects. Dijkstra et al. (2010) found exactly this: a continuous effect of orthographic overlap of translation equivalents on reaction times and an additional drop in reaction times for identical cognates during a lexical decision task. Also, identical cognate effects should be detectable earlier in the language process, because lateral inhibition takes place early in the visual word recognition process, than non-identical cognate effects, because semantic activation and feedback occur later in this process. Also, the orthographic form of an identical cognate is encountered more often than the orthographic forms of non-identical cognates and non-cognates, so the subjective frequency should be higher for these nouns resulting in higher resting activation levels and may be activated more quickly (Gollan, Forster, & Frost, 1997).

Our results diverge on two points from this proposed mechanism. First of all, no additional identical cognate facilitation is found above and beyond the effect of orthographic overlap. We do find both effects, but they are not present at the same time in the reading process. Our results suggest that this additional facilitating effect of identical cognate status is limited to certain tasks, like the lexical decision task of Dijkstra et al. (2010) and is not necessarily generalizable to a natural reading context.

Second, a delineation of early identical cognate effects and late non-identical cognate effects was not found. Actually, what we find looks more like the opposite. The earliest indication of language processing, skipping probability, was only affected by continuous orthographic overlap, not by identical cognate status. Another early reading measure, namely first fixation durations, also showed non-identical cognate facilitation. Later eye movement measures, total reading time and go past time, did not show convincing non-identical cognate facilitation, but did show identical cognate facilitation.

The BIA+ model (Dijkstra & Van Heuven, 2002) only hypothesizes inhibitory links between lexical representations. Motivated by our findings of the early linear effects of orthographic overlap, we propose that translation equivalents could be connected through excitatory connections of which the weight varies with the orthographic overlap between the two. This kind of connection could emerge when learning a second language and noticing orthographic similarities for certain translations. These direct links could be an efficient way to speed up L2 learning and lexical retrieval. By also assuming two orthographic representations for identical cognates, this assumption predicts early continuous effects of orthographic overlap without an additional boost for identical cognates. The late effects of identical cognate status might indicate stronger semantic feedback for identical cognates compared to non-identical cognates. A paper by Tokowicz, Kroll, de Groot, and Van Hell (2002) found that for nouns with only a single or very dominant translation, ratings of meaning overlap across languages are higher than for words with multiple translations. This is supported by studies finding slower translation latencies for words with multiple translation possibilities (e.g.Tokowicz & Kroll, 2007). Also, cognates are less likely than non-cognates to have multiple translations (Tokowicz et al., 2002), meaning that cognates could have a larger cross-lingual overlap in semantic representations, leading to larger semantic facilitation for the target word. This could result in a larger late cognate facilitation. Of course this mechanism necessitates the existence of separate but overlapping semantic representations for cognates.

Although we have some preliminary ideas, the processes underlying this particular unfolding of cross-lingual interactions are not yet clear. Further tailored research might clarify the precise way in which cognates are processed in the bilingual lexicon.

CONCLUSION

This paper examined the cognate effect in L1 and L2 text reading of a complete novel. The effect of identical cognate status and the continuous effect of orthographic overlap were investigated, both in early and in late reading measures.

We found early and late L1 and L2 cognate facilitation effects. These results provide an important insight into the processing of cognates by bilinguals. By using a large, naturalistic body of text, we have shown that a highly constrained and unilingual context does not eliminate cross-lingual activation effects in L1 or L2, and therefore that these effects are real and sufficiently meaningful to influence everyday reading.

This is the first time L1 cognate effects have been studied in a semantic linguistic context that is larger than one sentence. We found L1 identical cognate facilitation for high frequent words for total reading times. Long words that share part of their orthography with their translation equivalent also elicited facilitation for first fixation durations. These effects demonstrate that even when reading in the mother tongue, lexical access is not restricted to the target language.

Our findings of linear facilitating effects of orthographic overlap can be framed within the BIA+ model and are consistent with the idea of cross-lingual orthographic-semantic resonance leading to cognate facilitation. Frameworks that assume a shared-morpheme (e.g. Sanchez-Casas et al., 1992) for cognates cannot accommodate the linear effect of orthographic overlap.

We also found identical cognate facilitation, although not in conjunction with non-identical cognate facilitation. This could point towards a 'special' status for identical cognates compared to non-identical cognates or control words. A possibility is the existence of one shared orthographic representation for identical cognates, thus removing the lateral inhibition of two activated orthographic representations.

An alternative explanation of the results entails excitatory connections between translation equivalents, weighted by orthographic overlap, combined with separate representations for identical cognates sharing more semantic overlap than non-identical cognates.

In all, this study is the first to indicate just how ubiquitous cognate effects are in both L1 and L2 daily reading. Future research will have a large role in determining what the conditions and lexical variables are that determine the exact size and maybe even the direction of these effects.

REFERENCES

Balling, L. W. (2013). Reading authentic texts: What counts as cognate? *Bilingualism: Language and Cognition*, *16*(03), 637–653. http://doi.org/10.1017/S1366728911000733

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society*, *26*(2), 211–252.

Bultena, S., Dijkstra, T., & van Hell, J. G. (2013). Cognate and word class ambiguity effects in noun and verb processing. *Language and Cognitive Processes*, *28*(9), 1350–1377. http://doi.org/10.1080/01690965.2012.718353

Bultena, S., Dijkstra, T., & van Hell, J. G. (2014). Cognate effects in sentence context depend on word class, L2 proficiency, and task. *The Quarterly Journal of Experimental Psychology*, *67*(6), 1214–1241. http://doi.org/10.1080/17470218.2013.853090

Cop, U., Drieghe, D., & Duyck, W. (2015). Presenting GECO: An Eye-Tracking Corpus of Monolingual and Bilingual Sentence Reading. *Submitted for Publication*.

Cristoffanini, P., Kirsner, K., & Milech, D. (1986). Bilingual lexical representation: The status of Spanish-English cognates. *The Quarterly Journal of Experimental Psychology*, *38*, 367–393.

Davis, C., SáNchez-Casas, R., GarcíA-Albea, J. E., Guasch, M., Molero, M., & Ferré, P. (2010). Masked translation priming: Varying language experience and word type with Spanish–English bilinguals. *Bilingualism: Language and Cognition*, *13*(02), 137. http://doi.org/10.1017/S1366728909990393

De Groot, A. M. B. (2011). *Language and cognition in bilinguals and multilinguals: An introduction.* New York: Psychology Press.

Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of Cognates and Interlingual Homographs: The Neglected Role of Phonology. *Journal of Memory and Language*, *41*(4), 496–518.

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language*, *62*(3), 284–301. http://doi.org/10.1016/j.jml.2009.12.003

Dijkstra, T., Timmermans, M., & Schriefers, H. (2000). On being blinded by your other language: effects of task demands on inter-lingual homograph recognition. *Journal of Memory and Language*, *42*, 445–464.

Dijkstra, T., & Van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. Jacobs (Eds.), *Localist Connectionist Approaches to Human Cognition* (pp. 189–225). New York: Mahwah.

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(03). http://doi.org/10.1017/S1366728902003012

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 663–679. http://doi.org/10.1037/0278-7393.33.4.663

Grainger, J., O'Regan, J. K., Jacobs, A., & Segui, J. (1989). On the role of competing word units in visual word recognition:The neigborhood frequency effect. *Perception and Psychophysics*, *45*, 189–195.

Kirsner, K., Lalor, E., & Hird, K. (1993). The Bilingual Lexicon: Exercise, meaning and morphology. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 215–248). Amsterdam: John Benjamins.

Kroll, J. F., Dijkstra, T., Janssens, N., & Schriefers, H. (1999). Cross-language lexical activity during production: Evidence from cued picture naming. In A. Vandierendonck, M. Brysbaert, & K. Van der Goten (Eds.), *Proceedings of the 11th congress of the European society for cognitive psychology* (p. 92). Ghent: ESCOP/Academic Press.

Lalor, E., & Kirsner, K. (2000). Cross-lingual transfer effects between English and Italian cognates and non-cognates. *International Journal of Bilingualism*, *4*, 385–398.

Lemhöfer, K., Dijkstra, T., Schriefers, H., Baayen, R. H., Grainger, J., & Zwitserlood, P. (2008). Native language influences on word recognition in a second language: A megastudy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 12–31. http://doi.org/10.1037/0278-7393.34.1.12

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, *35*(2), 381–390. http://doi.org/10.1037/a0014875

McClelland, J., & Rumelhart, D. (1981). An interactive activation model of context effects in letter perception. An account of basic findings. *Psychological Review*, *88*, 375–407.

Peeters, D., Dijkstra, T., & Grainger, J. (2013). The representation and processing of identical cognates by late bilinguals: RT and ERP effects. *Journal of Memory and Language*, *68*(4), 315–332. http://doi.org/10.1016/j.jml.2012.12.003

Radach, R., Huestegge, L., & Reilly, R. (2008). The role of global top-down factors in local eye-movement control in reading. *Psychological Research*, *72*(6), 675–688. http://doi.org/10.1007/s00426-008-0173-3

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422.

Sanchez-Casas, R., Davis, C., & García-Albea, J. E. (1992). Bilingual Lexical Processing -Exploring the cognate non-cognate distinction. *European Journal of Cognitive Psychology*, *4*(4), 293–310.

Sanchez-Casas, R., & García-Albea, J. E. (2005). The representation of cognate and noncognate words in bilingual memory. Can cognate status be characterized as a special kind of morphological relation? In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches.* (pp. 226–250). Oxford: Oxford University Press.

Schepens, J., Dijkstra, T., & Grootjen, F. (2012). Distributions of cognates in Europe as based on Levenshtein distance. *Bilingualism: Language and Cognition*, *15*(01), 157–166. http://doi.org/10.1017/S1366728910000623

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language*, *55*(2), 197–212. http://doi.org/10.1016/j.jml.2006.03.004

Snodgrass, J. G., & Mintzer, M. (1993). Neigborhood effects in isual word recognition: Facilitatory or inhibitory? *Memory & Cognition*, *21*, 247–266.

Titone, D., Libben, M., Mercier, J., Whitford, V., & Pivneva, I. (2011). Bilingual lexical access during L1 sentence reading: The effects of L2 knowledge, semantic constraint, and L1–L2 intermixing. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, *37*(6), 1412–1431. http://doi.org/10.1037/a0024492

Tokowicz, N., & Kroll, J. F. (2007). Number of meanings and concreteness: Consequences of ambiguity within and across languages. *Language and Cognitive Processes*, *22*(5), 727–779. http://doi.org/10.1080/01690960601057068

Tokowicz, N., Kroll, J. F., de Groot, A. M. B., & Van Hell, J. G. (2002). Number-of-translation norms for Dutch-English translation pairs: A new tool for examining language production. *Behavior Research Methods Instruments & Computers*, *34*(3), 435–451.

Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *Journal of Memory and Language*, *64*(1), 88–107. http://doi.org/10.1016/j.jml.2010.08.006

Van Assche, E., Duyck, W., & Brysbaert, M. (2013). VERB PROCESSING BY BILINGUALS IN SENTENCE CONTEXTS. *Studies in Second Language Acquisition*, *35*(02), 237–259. http://doi.org/10.1017/S0272263112000873

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. *Psychological Science*, *20*(8), 923–927.

Van Hell, J. G., & de Groot, A. M. B. (2008). Sentence context modulates visual word recognition and translation in bilinguals. *Acta Psychologica*, *128*(3), 431–451. http://doi.org/10.1016/j.actpsy.2008.03.010

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review*, *9*(4), 780–789.

Van Heuven, W. J. B., & Dijkstra, T. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language*, *39*, 458–483.

APPENDIX A: PROFICIENCY SCORES

Due to the lack of a standardized cross lingual spelling test, we tested the English spelling with the spelling list card of the WRAT 4 (Wilkinson & Robertson, 2006) and the Dutch spelling with the GLETSCHR (Depessemier & Andries, 2009). The LexTALE (Lexical Test for Advanced Learners of English) is an unspeeded lexical decision task that contains a high proportion of words with a low corpus frequency. First developed as a vocabulary test, Lemhofer and Broersma (2011) have validated this as a measure of general English Proficiency. This test has been later extended to Dutch and German. The mean accuracy scores for the LexTALE are reported in Table A. 1. A classical speeded lexical decision task was also administered in Dutch and English. In Table A.1, the percentage of correct word trials corrected for false alarms is shown.

Table A. 1. *Average percentage scores (standard deviations between brackets and range between square brackets) on the LexTALE, Spelling test and Lexical Decision task for the bilingual and monolingual group*.

	Bilinguals L1	Bilinguals L2	t-value L1-L2
LexTALE- score (%)	92.43 (6.34) [73.75-100]	75.63(12.87) [51.25-98.75]	7.59 ***
Spelling score (%)	83.16(7.80) [67.00-93.00]	69.92 (8.74) [52.00-83.00]	8.15 ***
Lexical Decision score (%)	80.47 (5.45) [68.87-88.76]	56.75 (11.01) [38.46-75.86]	9.87 ***
Composite Proficiency Score (%)	85.54 (4.68) [77.87-95.25]	67.81 (9.72) [52.49-86.76]	11.78***

$p < 0.001$ ***

APPENDIX B: FORMULA FOR ORTHOGRAPHIC OVERLAP

The formula for the Corrected Levenshtein Distance was taken from Schepens et al. (2012).

$$Orthographic\ Overlap = 1 - \frac{Distance}{Length}$$

Distance = min (number of insertions, deletions and substitutions needed to edit target word into translation word)

Length = max (length of target word, length of translation word)

CHAPTER 6
CROSS-LINGUAL NEIGHBORHOOD EFFECTS IN GENERALIZED LEXICAL DECISION AND NATURAL READING[16][17]

The present study assessed intra- and cross-lingual neighborhood effects, using both a generalized lexical decision task and an analysis of a large-scale bilingual eye-tracking corpus (Cop, Keuleers, Drieghe & Duyck, in press).

Using novel neighborhood density and frequency measures, the general lexical decision task yielded an inhibitory cross-lingual neighborhood density effect on reaction times to second-language words, replicating van Heuven, Dijkstra and Grainger (1998). Reaction times for native language words were not influenced by neighborhood density or frequency but error rates showed cross-lingual neighborhood effects modulated by target word frequency.

The analysis of the large-scale eye movement corpus confirmed effects of cross-lingual neighborhood on natural reading, even though participants were reading a novel in a unilingual context. Both native language and second language reading were influenced by the presence of lexical candidates from the non-target language, although these effects in natural reading were largely facilitatory.

Our results provide direct evidence for activation of lexical candidates of the non-target language during bilingual visual word recognition.

---

[16] This chapter is based on a manuscript co-authored with Nicolas Dirix, Denis Drieghe, and Wouter Duyck: Cop, U., Dirix, N., Drieghe, D. & Ducyk, W. (2015). Cross-lingual Neighborhood Effects in Generalized Lexical Decision and Natural Reading, submitted for publication.

[17] Manuscript submitted for publication in Journal of Memory and Language.

INTRODUCTION

During written word recognition, we are faced with the complex task of activating and identifying the correct lexical representation among a large group of orthographically similar, but not identical, representations. The term orthographic *neighbor,* coined by Coltheart, Davelaar, Jonasson, and Besner (1977), is used to denote such a similar word. Coltheart et al.'s (1977) definition of such a neighbor is any word that can be created by changing one letter of the target word while preserving letter positions (example: *house* is a neighbor of the word *horse*). We will refer to this kind of neighbor as a s*ubstitution* neighbor from now on. Most studies examining neighbor effects used this definition. The number of neighbors of a particular target word is called the *neighborhood density* (N density).

All of the relevant monolingual models of word recognition hypothesize that a written word activates a set of possible lexical candidates. This means that at some point the correct target word has to be selected out of a number of neighbors. The search model (Forster, 1976) and the activation verification model (Paap, Newsome, McDonald, & Schvaneveldt, 1982) both predict that the neighborhood density will affect language performance because the actual decision is established by a frequency-ordered lexical search within those candidates. The longer the list of neighbors, the longer it would take to select the correct representation. The most popular model of word recognition, the interactive activation model (IA model, McClelland & Rumelhart, 1981), also makes the prediction that the number of candidates should affect lexical access but proposes that the reason for this is lateral inhibition (see also the multiple read-out model, Grainger & Jacobs, 1996). In the IA model, word identification starts with letter identification. These letters feed forward activation to lexical candidates. Each of these representations has a resting level of activation, which is determined by the frequency of the word. The activated representations feed activation backwards to the letter level. Word recognition is the end result of a competitive process between the activated lexical candidates, each inhibiting the others activation. The representation whose activation level rises significantly above that of the other candidates first, is selected.

The most intuitive hypothesis formed by the IA model is that words with more orthographically similar lexical items would receive more lateral inhibition from these neighbors and this would slow lexical access to the target word (e.g. Grainger & Jacobs, 1993). On the other hand, a facilitative effect of a larger neighborhood is also not impossible within the IA model. More neighbors could cause greater overall excitation in the lexicon, which could help in specific tasks like the lexical decision task (e.g. Andrews, 1997). Also, the feedback activation of multiple lexical candidates to particular letters, again activating the target representation, could facilitate activation of the correct lexical representation. The IA model assumes that these competitive processes between lexical candidates and activation of the lexical candidates occur in one single stage. In the IA model, word frequency determines the resting level activation of representations. For this reason, it could be expected that recognition of low frequent words would show larger effects of neighborhood density and frequency. This is because a low frequent representation will need more time to accumulate enough activation to significantly rise above the activation levels of the higher frequent neighbors, thus delaying lexical access to the target word.

Within the IA architecture, precise predictions about the time course of neighborhood effects and whether the combination of these counteracting effects would result in facilitation or inhibition of recognition are difficult to make. With its complex interactions between parallel activation of letters and words and lateral inhibition among words, the IA model can account for a lot of different effects. Indeed, model simulations of the IA model have shown that effects can be both inhibitory (Grainger & Jacobs, 1996; Jacobs & Grainger, 1992) or facilitating (Coltheart & Rastle, 1994; Pollatsek, Perea, & Binder, 1999) depending on stimulus materials and small adjustment to the parameters of the model. As we will see below, empirical investigations of neighborhood effects have also yielded a complicated mix of findings, with multiple moderating variables.

MONOLINGUAL NEIGHBORHOOD EFFECTS

ISOLATED WORD STUDIES.

In the empirical search for neighborhood effects, mainly two variables have been manipulated. The first one is the neighborhood density. Coltheart et al. (1977) were the first to show neighborhood density effects for isolated word recognition. In a lexical decision task, they found inhibitory effects for non-words with increasing neighborhood density, but no effects for words (see also Holcomb, Grainger, & O'rourke, 2002). After this, multiple authors investigated the effects of neighborhood density on lexical decision word performance. As Andrews (1997) argued in a review paper, large neighborhoods are almost always associated with better performance in standard lexical decision tasks. Indeed, most of these experiments pointed towards a facilitating effect of increasing neighborhood size for the speed and accuracy of lexical decision (Andrews, 1989, 1992; Carreiras, Perea, & Grainger, 1997; Forster & Shen, 1996; Huntsman & Lima, 2002; Johnson & Pugh, 1994; Laxon, Coltheart, & Keating, 1988; Pollatsek et al., 1999; Sears, Campbell, & Lupker, 2006; Sears, Hino, & Lupker, 1995). Similar results were found for naming (Peereman & Content, 1995; Sears et al., 1995) and semantic categorization tasks (Carreiras et al., 1997; Forster & Shen, 1996). Perceptual identification tasks have shown mixed results. Carreiras et al. (1997) reported slower reaction times for words with a large neighborhood density, whereas Snodgrass and Mintzer (1993) found a null effect in their experiment 1, facilitation in experiment 2 and inhibition in experiment 3, 4 and 5. Andrews (1997) concluded that inhibitory effects of large neighborhoods observed for perceptual identification tasks are the result of unusual stimulus environments or elaborate guessing strategies.

Another neighborhood measure that is used regularly is whether the target word has a higher frequent neighbor or not. We will refer to this factor as *neighborhood frequency* (N frequency). In lexical decision tasks it is usually found that reaction times are longer and accuracy is lower when a higher frequent neighbor is present (Carreiras et al., 1997; Davis & Taft, 2005; Grainger, 1990; Grainger & Jacobs, 1996; Grainger, Oregan, Jacobs, &

Segui, 1992; Grainger, O'regan, Jacobs, & Segui, 1989; Grainger & Segui, 1990; Huntsman & Lima, 1996; Perea & Pollatsek, 1998). This effect is also present for perceptual identification tasks (Carreiras et al., 1997; Grainger & Jacobs, 1996; Grainger & Segui, 1990).

Although research on neighborhood effects has predominantly used isolated word tasks, such as lexical decision tasks and naming tasks, there is some debate as to whether these tasks capture the cognitive processes underlying lexical access (e.g Balota & Chumbley, 1984; Rayner & Pollatsek, 1989). The most important argument is that isolated word tasks entail a decision component or behavioral response, decreasing the validity of the measure (e.g., Paap & Johansen, 1994; Rayner & Liversedge, 2011, Snodgrass & Mintzer, 1993). Because of this decision component, the lexical decision task in the specific case of neighborhood effects is insensitive to the cases where the participant activates the lexical representation of a higher frequent neighbor of the target word instead of the target representation and still responds with a correct "Yes" answer.

Kuperman, Drieghe, Keuleers and Brysbaert (2013) indeed showed that the lexical decision task and a more ecologically valid method, reading in context are distinguishable and measure, to a large extent at least, different language processes. They found that reaction times only explained 5-17% of the variance in gaze durations on target words embedded in sentences after partialling out the effects of word frequency and word length. This dropped to 0.2% of the variance in fixation durations in natural reading when not only the target words, but all words in the sentences are analyzed.

EYE TRACKING STUDIES.

It becomes clear that a more ecologically valid method, like a natural reading task, could produce measures that are a closer approximation of natural language processes and thus provide more direct evidence for the influence of neighborhood measures on lexical access. In the case of neighborhood effects, eye tracking can be especially useful because it has a very high temporal resolution This enables the study of the time course of language processing through multiple dependent variables, whereas with the lexical decision task only reaction times and accuracy scores can be investigated. Eye tracking during natural reading would therefore contribute to the study of neighborhood effects above and beyond lexical decision results.

So far, only a handful of studies investigated neighborhood effects in sentence reading using eye tracking. Only one of those investigated the effect of neighborhood density (Pollatsek, Perea, & Binder, 1999). In experiment 2 of Pollatsek et al. (1999), English participants read target embedded sentences for comprehension. Half of the target words had a lot of neighbors (average= 8.5), the other half few (average=2.2). All of these targets had at least one higher frequent neighbor. Their first analysis showed an inhibitory effect of neighborhood density for gaze duration and total reading time. Because in this analysis the number of neighbors was confounded with the number of higher frequent neighbors, Pollatsek et al. conducted another analysis, in which they held the number of higher frequent neighbors constant. Now they found that words with more low frequent neighbors were skipped more often, but these words were also regressed to more often. The authors noted that the facilitatory effect on skipping rates might be due to initial misidentification of the target word. However they did find a facilitatory effect in gaze durations that could not be due to such misidentification because it was stronger in the sentences where the highest frequent neighbor was implausible in the sentence context.

Perea and Pollatsek (1998) conducted another reading study, this time investigating the effect of neighborhood frequency. In their experiment 2 they instructed English participants to read sentences for comprehension. The embedded target words in these sentences were matched on number of neighbors. Half of the target words had an orthographic neighbor with a higher word frequency and the other half did not. The results showed more regressions towards the target word when it had a higher frequency neighbor than when it did not. Also, spillover effects were larger when the target word had a higher frequent neighbor. These effects were larger for low frequent target words. Davis, Perea, and Acha (2009) and Slattery (2009) conducted similar reading studies and confirmed that inhibitory effects of neighborhood frequency might occur late in the reading process. Davis et al. (2009) found an inhibitory effect of neighbor frequency for gaze durations and total reading time. Although there were also more regressions toward words with a higher frequent neighbor, this effect was not significant. Slattery (2009) found an inhibitory effect of the presence of a higher frequent neighbor in a sentence-reading task. More regressions were made and the total reading time was longer when the target word had a higher frequent neighbor. He pinpointed this effect on the initial misidentification of the target word, by showing that these effects are no longer present when the higher frequent word is not compatible with the sentence context. However, Sears et al., (2006) failed to report similar neighborhood frequency effects in an extensive set of reading experiments. They concluded that, at least in English, neighborhood frequency has no direct effect on reading times and has little to no effect on post-identification processes.

It becomes clear that all previous experiments examining neighborhood effects, either in isolated word studies or eye tracking studies, have focused on one of the two neighborhood variables, density or frequency, while holding the other one constant. It is not clear what the net result would be of either variable in natural reading when both vary simultaneously.

BILINGUAL NEIGHBORHOOD EFFECTS

In the field of bilingualism, one of the most important questions has been whether word recognition involves activation of lexical candidates from the non-target language. This question is tied in with the architecture of the bilingual lexical models, which may have one integrated, or two separate lexicons. Some have argued that lexical access for bilinguals is language-selective, meaning that when reading one language, only representations of that language are activated (Caramazza & Brones, 1979; Scarborough, Gerard, & Cortese, 1984). More recently however, a consensus has evolved in the literature that word recognition involves cross-lingual activation for bilinguals (for an overview see Dijkstra, 2007). The evidence for this mechanism comes mostly from studies using words that share features across two languages, such as inter-lingual homographs (words sharing orthography but not meaning across languages) and cognates. The latter are translation equivalent words that not only overlap in meaning but also in orthography (example of an identical cognate is the word “piano” in English and in Dutch). Cognates are recognized faster and more accurately than control words in behavioral studies that present words in isolation, such as lexical decision tasks (Bultena, Dijkstra, & van Hell, 2013; Dijkstra, Grainger, & van Heuven, 1999; Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010; Duyck, Van Assche, Drieghe, & Hartsuiker, 2007; Peeters, Dijkstra, & Grainger, 2013; Van Assche, Drieghe, Duyck, Welvaert, & Hartsuiker, 2011). Similarly, cognate effects have been observed when bilinguals read text (Duyck et al, 2007; Van Assche et al. 2011; Van Assche, Duyck, & Brysbaert, 2013), even in the native language (Van Assche, Duyck, Hartsuiker, & Diependaele, 2009). This is remarkable because the language of a text might serve as a useful cue in restricting access to the target language and therefore could speed up word recognition in this way (Altarriba, Kroll, Sholl, & Rayner, 1996; Van Assche, Duyck, & Hartsuiker, 2012). Generally, these cognate effects are attributed to spreading activation between representations of both languages. Alternatively, because cognates share the exact same orthography and often almost the exact same phonology and semantics, it has been argued that identical cognates could have a single representation across languages (see Dijkstra et al., 2010). This is not without importance, because there is only very indirect evidence that cognates would actually be represented separately for each language, which is necessary for an explanation in terms of cross-lingual lexical activation. So, a more conservative test of cross-lingual lexical activation would be one in which representations that are certainly language-specific, such as neighbors, interact with other language-specific representations of the other language. This is why the most compelling evidence for cross-lingual lexical access comes from cross-lingual neighborhood effects in bilingual reading. The only study so far providing such evidence is one by van Heuven, Dijkstra, and Grainger (1998). In this study, Dutch-English bilinguals performed a blocked and mixed-progressive demasking task, a generalized lexical decision task and an English lexical decision task. They constructed four item conditions by orthogonally manipulating the number of English and Dutch substitution neighbors in the CELEX database of the target words. For example the word *farm* was included in the large English and Dutch N condition, whereas *coin* was included in the condition with large English N and small Dutch N.

In the progressive demasking task, participants had to identify four-letter words that gradually appeared on a screen as fast as possible. In the blocked version of the task, the experiment consisted of two blocks, one containing only L1 words, the other containing only L2 words. Both in the English and Dutch block of the progressive demasking task, they found an inhibitory effect of non-target N density, but this effect only reached full significance in the L2 block. In the mixed progressive demasking task, L1 and L2 words were presented in a random order. Here the authors expected to find larger effects, because in a mixed language setting, both languages have to be active to perform the task. In this experiment, inhibition from the non-target neighbors was found for English and Dutch items. In the generalized lexical decision task, participants had to decide as fast and accurately as possible whether the target stimulus was a word (Dutch or English) or not. For the generalized lexical decision task, they again found inhibition of Dutch N and facilitation for English N for reaction times to the English items. No neighborhood effects were found for the Dutch items. In the English lexical decision task, monolingual and bilingual participants had to decide whether the presented stimulus was an English word or not. Here, again an inhibitory effect of Dutch N was found, showing that cross-lingual activation is not limited to mixed language contexts. All of these results were taken as evidence that words automatically activate substitution neighbors both pertaining to the target and non-target language. Although van Heuven et al.'s (1998) results were never replicated, two ERP studies supported the existence of cross-lingual N density effects, by showing a more negative N400 ERP component for words with more cross-lingual neighbors (Grossi, Savill, Thomas, & Thierry, 2012; Midgley, Holcomb, van Heuven, & Grainger, 2008). It is interesting to note that van Heuven et al. (1998) did not find any effects in a blocked or selective L1 setting of cross-lingual N density. Because this is the only study reporting cross-lingual N effects, so far there has been no direct evidence of cross-lingual activation of neighbors in L1 reading in a purely unilingual context. Eye tracking of bilingual natural reading could provide this evidence.

Also, the effect of cross-lingual N frequency has never been investigated. It is clear that the presence of a higher frequent neighbor influences reaction times and error rates in lexical decision tasks (e.g. Carreiras et al., 1997; Davis & Taft, 2005; Grainger & Jacobs, 1996; Grainger, 1990; Grainger, O'regan, Jacobs, & Segui, 1989; Perea & Pollatsek, 1998). Also, several studies provided evidence for an important role of this factor in N density effects (Carreiras et al., 1997; Davis & Taft, 2005; Grainger & Jacobs, 1996; Perea & Pollatsek, 1998). We will address this gap in the literature concerning cross-lingual neighborhood effects in the current study.

BIA+ MODEL

The findings on cross-lingual activation in bilingual reading described above have led to the development of the BIA+ model (Dijkstra & van Heuven, 2002). This model is the successor of the original BIA model (Dijkstra & van Heuven, 1998), which is a bilingual adaptation of the Interactive Activation model (McClelland & Rumelhart, 1981). The BIA+ model is a language non-selective model of lexical access, which entails an integrated bilingual lexicon (see Figure 1). Like in the IA model, a set of orthographic candidates is activated through bottom-up activation when a written word is encountered. Depending on their similarity to the printed word and their resting-level activation, determined by the word frequency, these representations are partly activated. As L2 items tend to be lower in subjective frequency for unbalanced bilinguals, their representations are activated somewhat slower than L1 items. The activation spreads from the orthographic candidates to the connected phonological and semantic representations. Every word in the lexicon is connected to one of the available language nodes, representing the language membership of that word. In BIA+, these nodes also represent the global lexical activity of a language. These nodes do not feed activation back to the orthographic or phonological level so they cannot function as a language selection mechanism. Instead, in order to account for differences in word recognition depending on tasks and other non-linguistic variables (for example instructions and expectations of the participants) the BIA+ model consists of a word identification system and a task/decision system. This architecture for the bilingual lexicon predicts that within and cross-lingual orthographic neighbors should take a prominent place in the lexical access process during visual word recognition.

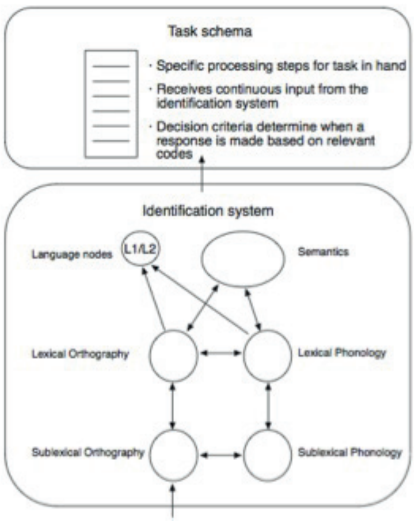


Figure 1. Schematic of the BIA+ model (taken from (Dijkstra & van Heuven, 2002, page 183)

Dijkstra, van Heuven, and Grainger (1998) explained van Heuven et al.'s (1998) cross-lingual neighborhood effects with the help of simulations of their results in the BIA model, because the BIA+ was then not formulated yet. Inhibition from non-target neighbors is explained by the mechanism of lateral inhibition on the lexical level. The facilitation of a larger within-language N density for bilinguals in English is explained by the relative activation of the two languages depending on word frequency in combination with asymmetric top-down inhibition from the language nodes implemented in the BIA-model (which disappeared in the BIA+ model). They also reported this facilitatory effect in monolinguals and refer to

Grainger and Jacobs (1996), who showed that facilitatory effects of large N could be simulated with the IA model with the help of read-out criteria. This means that Dijkstra et al. (1998) explained the monolingual and bilingual facilitatory N density effects in English in two different ways, which is not very parsimonious. Another challenge for Dijkstra et al.'s interpretation is that the top down activation from language nodes is not implemented in the BIA+ model. It is unaddressed how the authors would explain the facilitatory effects of target and non-target N density within the BIA+ model.

Another complicating factor is that it has become clear that defining N densities by only including substitution neighbors is insufficient. For example, Davis et al. (2009) found an additional effect of *addition neighbors* (by adding a letter to a word, e.g., *frog* is an addition neighbor of *fog*) and *deletion neighbors* (by deleting a letter from a word, e.g., *rash* is a deletion neighbor of *trash*) above and beyond the effect of substitution neighbors. Word recognition models with fixed letter positions such as the IA and BIA+ have problems explaining these effects, because lexical competition only occurs between representations of identical word length (see Davis & Bowers (2006) for an overview). Alternatively, there are monolingual models of word recognition with a relative positional nature that can account for effects of addition and deletion neighbors (e.g., the SOLAR model, Davis & Bowers, 2004; the SERIOL model, Whitney, 2001; the Overlap model, Gomez, Ratcliff, & Perea, 2008). van Heuven et al. (1998) counted the number of Dutch and English substitution neighbors of the target word using the CELEX database (Baayen et al., 1993). We might get a more accurate picture of cross-lingual N effects when we include addition and deletion neighbors in the N density measure. This new measure might be more sensitive in detecting cross-lingual effects in L1.

To conclude, despite the development of the BIA+ model and the abundance of papers addressing other effects of cross-lingual activation in visual word recognition, such as cognate effects (e.g. Bultena et al., 2013; Dijkstra et al., 1999; Duyck et al., 2007, Peeters et al., 2013; Van Assche, et al. 2011), there has been only one study that has provided direct evidence for parallel activation of lexical representations in bilingual word recognition by showing neighborhood effects. This study by van Heuven et al. (1998) used lexical decision instead of natural reading, and found no indications of L2 activation during pure L1 reading. We will therefore begin by replicating van Heuven et al.'s generalized lexical decision task, both using their categorization of stimuli and a broader N density measure, including addition and deletion neighbors. Next, we will investigate whether these cross-lingual N effects are present in a large database of bilingual eye movements of natural reading (Cop et al., in press), as a conservative test of parallel access to target language and non-target language representations of the bilingual lexicon.

EXPERIMENT 1

In Experiment 1 we replicate the generalized lexical decision task of van Heuven et al. (1998), using the exact same stimuli as them to investigate cross-lingual N density effects in a new group of bilingual Dutch-English participants. Based on their findings, we expect within-language facilitation and cross-language inhibition for L2 reading and only a small within-language inhibitory effect and no cross-lingual effect for L1 reading.

We will also present linear mixed effects analyses in addition to Van Heuven et al.'s ANOVAs, including English and Dutch N frequency variables. Also, by using a more inclusive measure of N density, we expect to find stronger effects of N density for L2 words and we might detect cross-lingual effects for L1 words.

Because of the architecture of the BIA+ model we expect larger effects of N density and frequency for low frequent target words (Perea & Pollatsek, 1998; Williams, Perea, Pollatsek, & Rayner, 2006).

METHOD

PARTICIPANTS

Thirty undergraduates received course credit for their participation in this experiment (19 females, M_{age} = 19.07 [2.08]). All students were unbalanced

Dutch-English bilinguals. Participants were tested for language proficiency with the Dutch and English version of the LexTALE (Lexical Test for Advanced learners of English, Lemhöfer & Broersma, 2012) and a self-report questionnaire (see Table B.1 in Appendix B for detailed proficiency scores). For the questionnaire, participants rated how good they were at listening, speaking, reading and writing in both languages on a 5-point Likert scale.

MATERIALS

The 160 words (80 Dutch and 80 English) and 160 nonwords were identical to those of Experiment 3 of van Heuven et al. (1998) (see Table 1 for word characteristics; see Appendix C for all stimuli). We calculated some additional word characteristics because they were not provided in the original study (e.g., bigram frequency) or because more up-to-date, and improved, measures exist nowadays (e.g., Subtlex frequencies, (Keuleers, Brysbaert, & New, 2010; van Heuven, Mandera, Keuleers, & Brysbaert, 2013) instead of CELEX (Baayen, Piepenbrock, & Van Rijn, 1993) word frequencies). Each participant saw each stimulus once, which resulted in 320 trials. All stimuli were presented in black against a white background. The font was Courier New, size 18 bold. Instruction language (Dutch or English) and response mapping (pressing the left button for a word, right for a nonword or vice-versa) were counterbalanced across participants.

The experiment was programmed in E-Prime 2.0 software (Schneider, Eschman, & Zuccolotto, 2012). Stimuli were presented on a Benq XL2411Z 24 inch LED monitor. The computer used for the experiment was a Dell Optiplex 3020 mini-tower with a 3.2GHz Intel Core i5-4570 processor. Participants had to respond by pressing left and right buttons on a RB-730 Cedrus responsebox.

Table 1. *Descriptive Statistics for the Stimuli Used in Experiment 1 by language and neighborhood density (standard deviations between parentheses).*

	Neighbors[a]		Number of neighbors[b]		Higher frequent neighbor[c]		Word Frequency[d]	Average Bigram Frequency[e]	CLD[f]
	Dutch	English	Dutch	English	Dutch	English			
Dutch	Large	Large	7 (2.49)	7.1 (4.17)	.85 (.36)	.90 (.30)	2.257 (0.61)	1828.47 (854.07)	0.30 (0.24)
	Large	Small	7.95 (2.94)	3.95 (2.8)	.90 (.30)	.75 (.43)	2.457 (0.86)	2533.67 (1665.93)	0.25 (0.24)
	Small	Large	4.05 (2.5)	6.6 (2.78)	.85 (.36)	.90 (.30)	2.364 (0.92)	1947.32 (961.72)	0.38 (0.27)
	Small	Small	3.45 (2.27)	4.6 (4.46)	.65 (.48)	.75 (.43)	2.368 (0.45)	2194.75 (1227.3)	0.26 (0.25)
English	Large	Large	5.35 (2.8)	8.15 (3.44)	.70 (.46)	.90 (.30)	3.576 (0.57)	1370.93 (541.67)	0.36 (0.23)
	Large	Small	2.15 (1.71)	8.3 (3.69)	.50 (.50)	.80 (.40)	3.758 (0.35)	1300 (608.32)	0.29 (0.27)
	Small	Large	5.9 (6.2)	5.4 (2.58)	.70 (.46)	.55 (.50)	3.434 (0.65)	1324.74 (668.89)	0.30 (0.33)
	Small	Small	1.9 (1.7)	4.15 (2.85)	.30 (.46)	.50 (.50)	3.505 (0.62)	1282.15 (653.78)	0.26 (0.27)
Nonwords	Large	Large	5.675 (2.59)	6.2 (2.94)					
	Large	Small	4.975 (2.19)	3.675 (2.41)					
	Small	Large	3.125 (2.27)	6.475 (3.14)					
	Small	Small	2.35 (1.92)	3.375 (1.84)					

[a]N densities as defined by van Heuven et al. (1998); [b]Total CLEARPOND N densities (Marian, Bartolotti, Chabal, & Shook, 2012); [c] The proportion of words with a higher frequent Neighbor, [d] Log10 Subtlex frequencies: Subtlex-NL for Dutch words (Keuleers et al., 2010), Subtlex-UK for English words (van Heuven et al., 2013); [e] Summated bigram frequencies (calculated using WordGen, (Duyck, Desmet, Verbeke, & Brysbaert, 2004) were normalized for corpus size and then divided by word length to obtain average bigram frequencies. Bigram frequencies could not be calculated for the nonwords: since van Heuven et al. (1998) didn't specify which of the nonwords were matched with which language, we couldn't determine which language corpus to use to calculate bigram frequencies; [f]Corrected Levenshtein distance was calculated with the formula in Appendix A by comparing the word with its closest translation in NIM (Guasch, Boada, Ferré, & Sánchez-Casas, 2013).

PROCEDURE.

Participants were seated in a comfortable chair at approximately 45-60cm from the screen. All instructions were presented on the screen. They were told they had to judge whether a presented letter string was either a word (in Dutch or English) or a non-word by pressing the according button. They were instructed to decide as quickly and accurately as possible. After the instructions, participants had to perform a practice block with 10 trials (five words and non-words each), which was repeated if their accuracy was below 80%. None of the stimuli used in the practice block were used in the experimental block. Afterwards the experimental block followed, with a presentation of the stimuli in a pseudo-random order. No more than four consecutive words or nonwords were presented. Halfway the experiment participants could take a short break.

A trial started with the presentation of a fixation cross (800ms), followed by a blank screen of 300ms. The stimulus was presented until the participant responded or for a maximum duration of 2500ms. The inter-trial interval was kept constant at 700ms.

After finishing the experiment, participants were presented with the English and Dutch version of the LexTALE and the self-reported questionnaire. The entire session lasted about 45 minutes.

ANOVA ANALYSES

First, our data was analyzed using the exact same analytic procedure that van Heuven et al. (1998) used in their Experiment 3. We applied a 2 (Dutch N density: small vs large) x 2 (English N density: small vs large) within-subjects ANOVA separately for each language and the non-words, with the same N density categorization that van Heuven et al. (1998) applied in their Experiment 3. This resulted in 20 words per condition.

F1 AND F2 ANALYSIS

All analyses were performed in R version 3.1.2 (R Core Team, 2014). We present the results for the words below. Results for nonwords can be found in Table S.1 of the online supplementary materials. We analyzed the data by means of a participant (*F*1) and item (*F*2) analysis, to allow a direct comparison with analyses of van Heuven et al. (1998). Data-trimming was also conducted as in van Heuven et al.: responses that were more than 2 standard deviations above or below participants' or item's mean RT (9.82% of the data) were excluded. Additionally, for the RT analyses, incorrect responses were excluded (6.71% of the data). See Table 2 for mean RTs and error rate.

Table 2. *Average RTs (in milliseconds) and Error Rate (percentage) for each language and Dutch x English N density in Experiment 1 (standard deviations between parentheses).*

Dutch N	English N	English		Dutch		Nonwords	
		RT	Error rate	RT	Error rate	RT	Error rate
Large	Large	688 (104)	5.76 (6.1)	662 (89)	5.72 (7.07)	756 (106)	6.39 (7.28)
Large	Small	677 (90)	5.86 (5.7)	644 (89)	6.65 (5.97)	746 (102)	6.55 (8.14)
Small	Large	684 (72)	7.74 (6.91)	660 (94)	4.8 (5.91)	736 (108)	4.36 (5.43)
Small	Small	665 (100)	5.92 (7.05)	658 (84)	5.88 (6)	722 (99)	3.35 (4.19)

RT ANALYSIS

ANOVA's were carried out on the average RTs for each language separately with N density in Dutch (small or large) and English (small or large) as within-participant factors. Analysis by participants and by items of Dutch performance showed that there was no effect of Dutch N density ($F1\ (1,29) < 1$, $p = .419$, $F2(1,76) < 1$, $p = .522$) and for English N density only in the by participants analysis ($F1(1,29) = 10.22$, $p = .003$, $F2(1,76) < 1$, $p = .323$). The interaction did not reach significance either ($F1(1,29) < 1$, $p = .424$, $F2(1,76) < 1$, $p = .987$).

For English reading, there was no effect of Dutch N density ($F1(1,29) < 1$, $p = .849$, $F2(1,76) < 1$, $p = .476$). The English N density effect was significant for the analysis by participants ($F1(1,29) = 13.58$, $p < .001$), but not by items ($F2(1,76) = 1.797$, $p = .184$). The interaction was not significant ($F1(1,29) < 1$, $p = .509$, $F2(1,76) < 1$, $p = .786$).

ERROR PERCENTAGE ANALYSIS

For Dutch reading, the Dutch N density effect was only significant for participants ($F1(1,29) = 5.514$, $p = .026$, $F2(1,76) = 1.007$, $p = .319$). There was no effect of English N density ($F1(1,29) < 1$, $p = .81$, $F2(1,76) < 1$, $p = .951$). The interaction was only significant for the participants analysis ($F1 (1,29) = 25.47$, $p < .001$, $F2(1,76) = 1.86$, $p = .177$).

For English reading, the Dutch N density effect was again only significant for participants ($F1(1,29) = 9.644$, $p = .004$, $F2(1,76) = 1.475$, $p = .228$). There was also only an effect of English N density for the participant analysis but not for the items ($F1(1,29) = 5.758$, $p = .023$, $F2(1,76) < 1$, $p = .468$). The interaction was not significant ($F1(1,29) < 1$, $p < .7$, $F2(1,76) < 1$, $p = .804$).

DISCUSSION ANOVA RESULTS

For the L1 words, there was no effect of within- or cross-language N density, both in RTs and Error rates. For the L2 words, we found a within-language effect of N density, but only for RT analyses by participants. There was no within-language effect for the Error rates and the cross-language effect N density effect was not significant for neither RTs or Error rates. So, using the same stimuli and analysis, with participants of the same bilingual population we could did not replicate any of van Heuven et al.'s (1998) findings. We should note however that, in comparison with the study of van Heuven et al. (1998), the current RTs are in general slower (on average 76.5ms) and the error rates somewhat lower (2.74% less errors). However, before drawing definitive conclusions from these data, we wanted to investigate these data in an alternative manner. We had several reasons for doing so.

Our data was analyzed using (a) new measures of neighborhood density and frequency, (b) linear mixed effects modeling, and (c) continuous covariates.

First, it has become clear that N densities are inconsistently identified in the literature (Marian et al., 2012), meaning that researchers use different language databases to determine how large the neighborhoods of their stimuli are. This makes it very difficult to compare results across experiments. To overcome this problem, Marian et al. (2012) developed the CLEARPOND database (Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities), which provides N densities and also allows comparing N densities across languages by including comparable corpora of multiple languages. When using this database to calculate N densities, the dichotomous neighborhood density classification that van Heuven et al. (1998) made does not apply anymore. In the current study, we used CLEARPOND (Marian et al., 2012) to determine a more accurate N density and frequency value, including within and cross-language substitution, addition and deletion neighbors.

Second, we favored the use of a linear mixed model over the classical by participants and by items analyses. First of all the correct minF' statistic proposed by Clark (1973) is hardly ever applied in psycholinguistic research and is often difficult to compute when the data contain many missing values (e.g. during eye tracking when there is no fixation time due to word skipping). Second, Baayen, Davidson, and Bates (2008) pointed out that linear mixed models are superior to the by participants and by items analyses for analyzing psycholinguistic data. Using this technique, both participants and items can be put in the statistical model as random factors at once. Furthermore, there is no theoretical motivation to choose arbitrary numbers as small and large N densities, as van Heuven et al. (1998) did, or to determine that the difference between those two numbers is large enough. With linear mixed effect models, variables can be investigated continuously, so there is no need for an arbitrary categorization of variables such as neighborhood density. Covariates and interactions can easily be included in the model, thus providing us with a more fine-grained overview of N density effects.

Third, we included several covariates in the analysis. The most important covariate was included because of its indisputable role in (bilingual) visual word recognition: word frequency (Baayen et al., 2006; Keuleers et al., 2010). Bigram frequency was added because there was a lot of variation for this variable between conditions.[18] We also added corrected Levenshtein distance (Schepens, Dijkstra, & Grootjen, 2012) to the analysis (See Appendix A for the formula). Although no cognates were included in the stimulus set, we thought it important to include this covariate because Van Assche et al., (2011) showed that an increased amount of cross-lingual overlap causes a continuous facilitatory effect in word recognition. See Table 1 for a summary of word characteristics on all of these variables.

LINEAR MIXED MODEL ANALYSIS

All analyses were performed in R (R Core Team, 2014). Models were fitted using the lme4 (version 1.1-7) and the lmerTest package (version 2.2-20) of R (version 3.1.2) (Bates, Mächler, Bolker, & Walker, 2014; R Core Team, 2014). For all analyses, RTs, word frequencies and average bigram frequencies were log transformed with base 10 to normalize their distribution. All continuous variables were centered to reduce collinearity. For the analysis, stimuli that did not reach 70% accuracy were excluded (5.31% of the data). Furthermore, responses that were more than 2.5 standard deviations above or below participants' mean RT (4.99% of the data) were excluded. Additionally, for the RTs analysis incorrect responses were excluded (4.96% of the data). Separate analyses were carried out for each language (Dutch and English) and for the nonwords, both for RTs and error rates. The fixed factors in the models were Dutch N density (continuous), English N density (continuous), Dutch N Frequency ("Yes" indicated that the word had a higher frequent neighbor in Dutch, "No" if it

[18] As can be seen in Table 1 and Table 3, the average bigram frequency is a lot higher for the Dutch than it is for the English materials. Apparently, the frequency distribution of bigrams is very different in the Dutch language from the distribution in the English language. In effect, when we calculate the average bigram frequency for all words included in the CELEX database we see that for Dutch it is 8 095 and for English 1 950 after controlling for corpus size and word length.

did not), English N Frequency ("Yes" indicated that the word had a higher frequent neighbor in English, "No" if it did not), word frequency (continuous), average bigram frequency (continuous) and corrected Levenshtein distance (continuous). We included a random intercept per subject in all initial models. This ensured that differences between subjects concerning genetic, developmental or social factors were modeled. We also included a random intercept per word, to be able to generalize to other nouns, because our stimuli sample is not an exhaustive list of all nouns in a language. First a full model, including the two random clusters and all of the 2-way interactions between the neighborhood variables and word frequency, word length and bigram frequency, was fitted. The optimal model was discovered by backward fitting of the fixed effects, then forward fitting of the random effects and finally again backward fitting of the fixed effects (Barr, Levy, Scheepers, & Tily, 2013).

We report the analysis of the Dutch and English words below. Wherever interactions reached significance, we determined the region of significance with simple effect estimates using linear contrasts. The analysis of the nonwords is reported in Table S.2 of the supplementary materials.

RESULTS DUTCH WORDS

Results of the analysis of RTs and error rates are presented in Table D.1 and D.2 of Appendix D. We did not find any main effects of within- or cross-lingual neighborhood density or frequency on reaction times or error rates.

However, for error rates the interaction between cross-lingual N density and word frequency was marginally significant (β= 0.13, SE=0.07, t=1.93, p=0.053, see Figure D.1 in Appendix D). Linear contrasts revealed that there were fewer errors for low frequent words (<1.73 log word frequency, χ^2=3.84, df=1, $p < 0.05$) and more errors for high frequent words (>4.19 log word frequency, χ^2=2.71, df=1, $p < 0.1$) with increasing cross-lingual N density.

RESULTS ENGLISH WORDS

Results of the analysis of RTs and error rates are presented in Table D.3 and D.4 of Appendix D.

For reaction times, again no main effect of any neighborhood variable was found. Nevertheless, there was a significant interaction between cross-lingual N density and bigram frequency (β=-0.013, se=0.0066, t=-2.04, $p < 0.05$, see Figure 2). Linear contrasts revealed slower reaction times for words with a low bigram frequency (<2.953 log average bigram frequency, χ^2=3.85, df=1, $p < 0.05$) with increasing Dutch N density.

The interaction between the presence of a higher frequent English neighbor and word frequency was significant (β=-0.040, se=0.016, t=-2.51, $p < 0.05$). Reaction times were slower for low frequent words (<3.29 log word frequency, χ^2=3.84, df=1, $p < 0.05$) and a faster for high frequent words (>3.87 log word frequency, χ^2=2.71, df=1, $p < 0.1$) when the target word had an English neighbor of higher frequency. The contrasts of the marginally significant interactions between English N density and word frequency (β=0.0044, se=0.0025, t=1.75, p = .086) and English N frequency and bigram frequency (β=-0.056, se=0.032, t=-1.79, p = .078) did not yield significant effects.

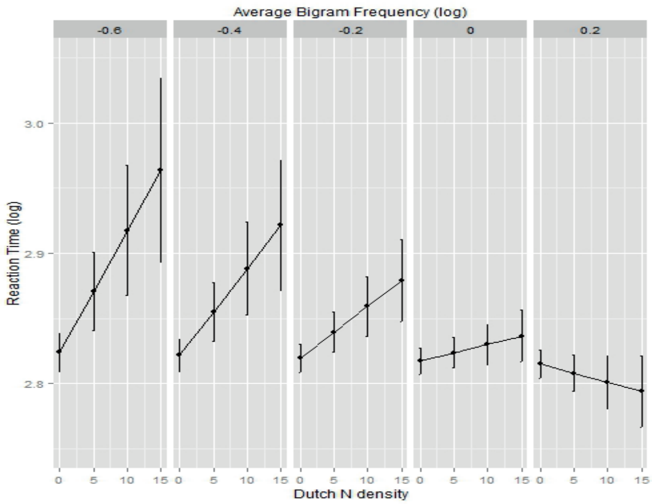


Figure 2. Reaction times (log transformed on the y-axis) for English words by Dutch N density (on the x-axis) and bigram frequency of the word (panels) for a generalized lexical decision task.

For error rates, the main effect of cross-lingual N density was significant (β=0.10, se=0.040, t=2.32, $p < .05$, See Figure D.2 in appendix D). More errors were made when the English noun had more Dutch neighbors. No other main effects of neighborhood were significant.

The marginal interaction between English N density and bigram frequency (β=-0.29, se=0.17, t=-1.65, $p = .0099$) showed that for low bigram frequency words (>3.1055 log average bigram frequency) there were fewer errors with increasing English N density (χ^2=3.84, df=1, $p < 0.05$). The linear contrasts for the marginally significant interaction between English N frequency and word frequency (β=-1.01, se=0.55, t=-1.84, $p = .065$) showed that there were more errors for low frequent words (<3.665 log frequency) when the word had a higher frequent English neighbor (χ^2=3.84, df=1, $p < 0.05$).

DISCUSSION LMER RESULTS

In the present experiment, we attempted to replicate van Heuven et al.'s (1998) findings of cross-lingual effects of N density in a generalized lexical decision task for English words. In the first analysis, that van Heuven et al.

also performed, we failed to find any of the earlier reported effects. A more detailed pattern of neighborhood effects was however discovered by analyzing the data by means of linear mixed models.

CROSS-LINGUAL NEIGHBORHOOD EFFECTS

L1 LEXICAL DECISION. For L1 (Dutch) reading we found a lower error rate with an increased cross-lingual N density for low frequent words, but a reversed pattern for high frequent words. This means that the effect of cross-lingual N was facilitatory for low frequent words and inhibitory for high frequent words.

L2 LEXICAL DECISION. In L2 (English) reading, the cross-lingual N effects were all inhibitory: we found slower reaction times for low bigram frequency words and more errors for all L2 words when cross-lingual N density increased.

WITHIN-LANGUAGE NEIGHBORHOOD EFFECTS

L1 LEXICAL DECISION. We found no L1 (Dutch) within-language effect of N density or frequency in the current study.

L2 LEXICAL DECISION. When within-language L2 (English) N density increased, less errors were made towards words with a low bigram frequency. We also found slower reaction times and more errors for low frequents words when the noun had a within-language higher frequent neighbor.

As mentioned before, our RTs were in general slower than those of van Heuven et al. (1998). Instruction format can make a difference in lexical decision tasks when dealing with N density effects (Sears et al., 2006), but we emphasized both speed and accuracy (as Van Heuven et al. did) so this is an unlikely cause of the slower reaction times. Furthermore, the language proficiency of our participants could be different from those of van Heuven et al. (1998), causing the difference in RTs. Unfortunately, van Heuven et al. (1998) did not provide proficiency scores for their participants so we cannot make a comparison. There were however some procedural differences between our generalized lexical decision task and van Heuven et al.'s

(1998). For example, our participants were allowed more time to answer. This might have slowed down the responses of our participants.

Van Heuven et al. (1998) found an inhibitory effect of L1 N density for reaction times to L2 words in a generalized lexical decision task. For L1 words the effect of L2 N density did not reach significance. In our lmer analyses of the replication, we found similar results for L2 words: inhibition with increasing L1 N density for reaction times (for words with a low bigram frequency) and error rates. For L1 words, the effect of L2 N density again did not reach significance for reaction times, but it did for error rates. The fact that the direction of the effect is determined by word frequency is not that surprising, but the way it is determining it is. The low frequent words are responded to more accurately whereas the high frequent words are responded to less accurately when they have a lot of cross-lingual neighbors. We might expect low frequent words to be more liable to lateral inhibition from other lexical candidates, because the lower resting activation needs more time to reach threshold. It seems that in L1 recognition of low frequent words is helped by activated neighbors from the other language.

The dominant finding in the monolingual literature is facilitation of N density in lexical decision tasks (e.g., Andrews, 1989,1992; Carreiras et al., 1997; Pollatsek et al., 1999). We observed this in L2 but not in L1. The inhibitory within-language effect of a higher frequent neighbor (e.g., Carreiras et al., 1997; Davis & Taft, 2005; Grainger & Jacobs, 1996; Perea & Pollatsek, 1998) usually found in monolingual lexical decision was also only present in L2. Taking in account all of these results, we can conclude that the results for the English L2 words are fairly comparable to the existing literature on neighborhood effects. The discrepancy for L1 words could be explained by the fact that a generalized lexical decision task might force participants in a bilingual context, which possibly leads to results different from a normal unilingual lexical decision task (e.g. van Heuven et al.'s (1998) English lexical decision task yielded no L2 within-language effect for bilingual participants, whereas this effect was present in the generalized lexical decision task).

EXPERIMENT 2

In experiment 2, we investigated N density and N frequency effects in a large database of natural reading (Cop et al., in press). We analyzed the eye movements of late unbalanced Dutch-English bilinguals when reading L1 and L2 nouns. Because of the problems with lexical decision as a marker for lexical access, it is of capital importance that cross-lingual N effects are replicated in text reading. Because cross-lingual neighborhood effects have only been investigated in lexical decision (van Heuven et al. (1998) and our experiment 1), if we find cross-lingual neighborhood effects in these analyses, this would provide the first direct evidence in a completely unilingual context for the existence of activation of non-target language lexical representations.

Given the fact that the BIA+ model could explain facilitatory and inhibitory effects of cross-lingual neighborhood density (Dijkstra & van Heuven, 2002), it is not clear whether we expect facilitation or inhibition from either neighborhood density or frequency. Some of the previous monolingual reading research has corroborated the idea that inhibition from neighbors might arise later in the reading process than facilitation (Perea & Pollatsek, 1998; Pollatesek et al., 1999), although the BIA+ model does not necessarily predict this.

Because of the low correlations between reaction times on lexical decision times and eye movements (Kuperman et al., 2013) and because it has been shown that neighborhood effects are very task dependent (e.g., Andrews, 1997; Carreiras et al., 1997) it is also difficult to make predictions based upon the results of the cross-lingual N effects found in lexical decision tasks.

We do expect that cross-lingual neighborhood effects should parallel within-language neighborhood effects, because according to the BIA+ model there is no inherent difference between lexical items from the target or the non-target language and because top down inhibition from the language nodes is absent (Dijkstra & van Heuven, 2002). We also expect that for Dutch L1 reading, the cross-lingual effects will be smaller than for English L2 reading. This because for unbalanced bilinguals, most L2 representations are expected to be of lower frequency, thus having lower resting level activation. This is also the reason why we expect that cross-lingual neighborhood effects might be stronger than within language effects in L2 reading.

Because our materials are a whole text, we analyze words varying in both neighborhood density and frequency. This means that we can examine the two effects at the same time. This will be very informative about the net effect of the neighborhood variables in bilingual natural language reading. The fact that our materials are not selected on certain lexical variables, also means that we will investigate a database of nouns with a full range of word frequency, word length and bigram frequency. Since some results have shown that such lexical variables can modulate the neighborhood effects, we do expect to identify the precise conditions in which neighborhood effects are stronger.

METHOD

PARTICIPANTS AND MATERIALS.

We selected all nouns (1 745 unique English and 1 777 unique Dutch nouns) from the eye-tracking corpus of (Cop et al., in press). This corpus consists of eye movements recorded from nineteen unbalanced Dutch-English bilingual (seventeen female, M_{age} = 21.2 [2.2]) and thirteen English monolingual undergraduates (seven female, M_{age}=21.8 [5.6]) who read the entire novel “The mysterious affair at Styles” by Agatha Christie (Title in Dutch: “De zaak Styles”). Participants’ proficiency was tested with a proficiency battery including the LexTALE (Lemhöfer & Broersma, 2012), a lexical decision task and spelling tests (GL&SCHR for Dutch, De Pessemier & Andries (2009); WRAT4 for English, Wilkinson & Robertson (2006)). See Table B.2 in Appendix B for detailed proficiency scores. All nouns that had an identical cognate in the other language were excluded from the dataset (8% for Dutch, 9.1% for English). Due to paucity of data for the highest values for N density, we trimmed the range to within 2.5 standard deviations of the mean. The unrestricted range for Dutch was [0-29] for Dutch N and [0-37] for English N. We excluded nouns with more than 19 Dutch neighbors or more than 14 English neighbors in the Dutch dataset (6.4%). The unrestricted range for English was [0-39] for English N and [0-29] for Dutch N. We excluded nouns with more than 25 English neighbors or more than 15 Dutch neighbors in the English dataset (7.4%). The final dataset consisted of 1 503 unique Dutch and 1 496 unique English nouns. See Table 3 for characteristics of these nouns.

Table 3. *Descriptive Statistics for the nouns analyzed in Experiment 2, averaged over stimuli per language (standard deviations between parentheses).*

	Neighborhood density[a]		Neighborhood Frequency[b]		Word Frequency[c]	Average Bigram Frequency[d]	Average Word Length	CLD[e]
	Dutch	English	Dutch	English				
Dutch	3.42 (4.16)	1.33 (2.37)	.27 (.44)	.15 (.35)	3.18 (0.96)	3308.00 (1509.16)	6.71 (2.29)	0.32 (0.26)
English	1.80 (2.91)	5.50 (5.69)	.23 (.42)	.52 (.50)	3.35 (0.93)	1766.00 (682.68)	6.00 (1.99)	0.33 (0.27)

[a]Total CLEARPOND N densities (Marian et al., 2012); [b]Log10 Subtlex frequencies: Subtlex-NL for Dutch words (Keuleers et al., 2010), Subtlex-US for English words (Brysbaert & New, 2009); [b]The proportion of words with a higher frequent Neighbor; [c]Summated bigram frequencies (calculated using WordGen, (Duyck et al., 2004) were normalized for the respective corpus size and then divided by word length to obtain average bigram frequencies. [e]Corrected Levenshtein distance was calculated with the formula in Appendix A by manually comparing the word with its closest translation.

PROCEDURE.

Each participant read the entire novel silently over four separate sessions. The monolinguals read the entire novel in English. The bilinguals read half of the novel in Dutch, the other half in English. The order was counterbalanced. Paragraphs were presented on the screen. Participants pressed a button when they were ready to progress. After each chapter, multiple-choice questions were asked to check whether participants were reading for comprehension. For further details on the procedure, see Cop et al. (in press).

ANALYSES EYE MOVEMENTS

We analyzed two eye movement measures that reflect early language processes: Probability of first pass skipping of a word and single fixation duration, the fixation duration on a word that is fixated exactly once. We analyzed a measure reflecting intermediate language processing: Gaze duration, the sum of all fixation durations during first passage before the eyes move out of the word. Finally, we analyzed two measures that reflect later, higher-order, language processes such as semantic integration: total reading time (the sum of all fixation durations on the target word, including refixations) and finally regression probability, the probability of making a regression back towards the target word.

Reading time measures and skipping probabilities were fitted in (general) linear mixed models using the lme4 (version 1.1-7) and the lmerTest package (version 2.2-20) of R (version 3.1.2) (Bates et al., 2014; R Core Team, 2014). All of the initial models contained the fixed factors of English N Density (continuous), English N Frequency (Yes or No), Dutch N Density (continuous) and Dutch N Frequency (Yes or No). As in experiment 1, Word Frequency (continuous), Bigram Frequency (continuous) and Orthographic Overlap (continuous) were included as covariates. Here, also Word Length (continuous) was included because this variable was not constant, as it was in experiment 1. All factors and covariates were calculated the same way as in experiment 1. Model fitting and simple effect estimates were done in the same way as in experiment 1.

RESULTS

We will first present the eye-tracking results of the bilingual L1 and L2 reading in detail. After this we will give a short summary of the monolingual eye-tracking results for purpose of validating the cross-lingual effects in the bilingual results.

L1 READING (DUTCH)

EARLY MEASURES. The outcome of the final model for skipping probabilities and single fixation durations is presented in Table E.1 and E.2 in Appendix E. For skipping probability, a logistic linear mixed model was fitted. For the single fixation analyses, only the nouns that received one fixation were selected (56.1%). Single fixation durations that differed more than 2.5 standard deviations from the subject means were excluded (2.20%).

CROSS-LINGUAL N EFFECTS. We found no main effects of cross-lingual neighborhood density or frequency for the early measures. The interaction between English N frequency and word frequency was significant (β=0.11, se=0.046, t=2.46, p < .05). The probability of skipping a word was higher when this noun had a higher frequent English neighbor, but only when the noun was high frequent (>4.00 log word frequency, χ^2=3.95, df=1, p-value < 0.05). When the noun was low frequent (<1.53 log word frequency, χ^2=3.89, df=1, p-value < 0.05), it was skipped less often (see Figure 3). For single fixation durations we found no cross-lingual neighborhood effects.

Within-language N effects. For skipping rates, we found significant interactions between Dutch neighborhood density and word frequency (β=-0.017, se=0.0046, t=-3.71, p <. 001) and word length (β=-0.012, se=0.0034, t=-3.53, p < .001). The linear contrasts showed that when nouns were very low frequent (<1.89 log word frequency, χ^2=3.85, df=1, p-value<0.05) or short (5 characters or less, χ^2=10.89, df=1, p-value < 0.01), a larger amount of Dutch neighbors makes it more likely that the noun is skipped. For high frequent nouns (>4.06 log word frequency, χ^2=3.88, df=1, p-value < 0.05) and longer words (9 characters or more, χ^2=4.16, df=1, p-value < 0.05) a larger neighborhood density makes it less likely the noun was skipped.

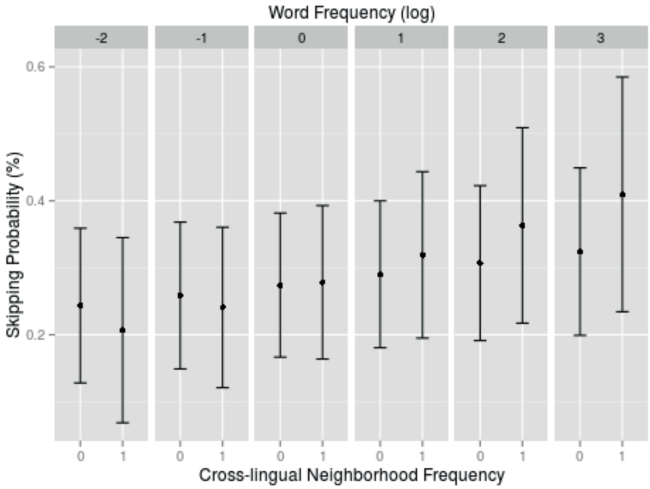


Figure 3. Skipping probability for nouns with a higher frequent neighbor (1 on the x-axis) and without (0 on the x-axis) dependent on target word frequency (panels) for Dutch L1 reading.

The interaction between Dutch neighbor frequency and word length was marginally significant (β=0.038, se=0.022, t=1.73, p = .0807). For words 4 characters long or shorter, there was marginal inhibition for having a higher frequent Dutch neighbor.

For single fixation durations, we found a similar interaction effect of Dutch neighborhood density with word frequency (β=0.0015, se=0.00034, t=4.30, p <. 001). As the number of Dutch neighbors increased, single fixations became shorter for low frequent nouns (<2.88 log word frequency, χ^2=3.96, df=1, p-value < 0.05) and longer for high frequent nouns (>4.17 log word frequency, χ^2=4.18, df=1, p-value < 0.05).

To sum up, in L1 reading we observed cross-lingual N effects in skipping rates, an indicator of early language processing. The presence of a higher frequent cross-lingual L2 neighbor yielded skipping of high frequent L1 nouns, but the reverse for low frequent L1 nouns. There was also within-language N density facilitation for low frequent, short words and inhibition for high frequent, long words early in the word recognition process.

INTERMEDIATE MEASURES. The outcome of the final model for gaze durations is presented in Table E.3 in Appendix E. Gaze durations that differed more than 2.5 standard deviations from the subject means were excluded (2.55%).

CROSS-LINGUAL N EFFECTS. We found no main effects of any of the cross–lingual neighborhood variables in the gaze duration data. None of the interaction effects including cross-lingual neighborhood variables reached significance.

WITHIN-LANGUAGE N EFFECTS. We found no main effects of the within-language N density or frequency. Again, the interaction between Dutch N density and word frequency was significant (β=0.0023, se=0.00038, t=6.12, p <. 001). Linear contrasts showed that for high frequent nouns (>3.89 log word frequency, χ^2=4.00, df=1, p-value < 0.05), the effect was inhibitory whereas for low frequent nouns the effect was facilitating (<3.03 log word frequency, χ^2=3.86, df=1, p-value < 0.05).

LATE MEASURES. The outcome of the final model for total reading times and regression rates is presented in Table E.4 and E.5 in Appendix E. Total reading times that differed more than 2.5 standard deviations from the subject means were excluded (2.90%). For regression rate a logistic linear mixed model was fitted.

CROSS-LINGUAL N EFFECTS. We found no significant effect of cross-lingual N density or frequency for total reading times. For regression rates, having a higher frequent English neighbor had a marginal facilitatory effect (β=-0.18, se=0.099, t=-1.80, p =. 0718). This means that a noun with a cross-lingual higher frequent neighbor was regressed to less often. There was a marginally significant interaction-effect between English N frequency and bigram frequency (β=-0.55, se=0.33, t=-1.68, p =. 0924, see Figure 4). Linear contrasts showed significant facilitating effects for nouns with 3.51 log bigram frequency or higher (χ^2=4.01, df=1, p-value < 0.05).

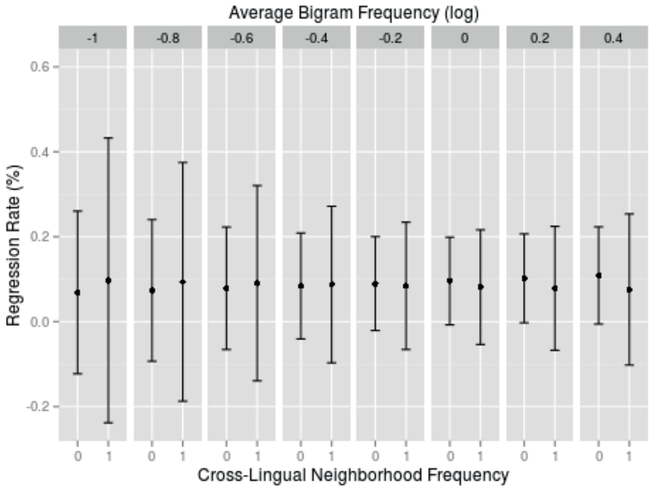


Figure 4. Regression Rates for nouns with a higher frequent neighbor (1 on the x-axis) and without (0 on the x-axis) dependent on average bigram frequency of the noun (panels) for Dutch L1 reading.

WITHIN-LANGUAGE N EFFECTS. For total reading times, again the interaction between Dutch N density and word frequency was significant (β=0.0028, se=0.00048, t=5.87, p <. 001). Dutch N density had a facilitating effect for low frequent nouns (<2.73 log word frequency, χ^2=3.90, df=1, p-value < 0.05) and an inhibitory effect for high frequent nouns (>3.64 log word frequency, χ^2=3.89, df=1, p-value < 0.05). We also found that participants were less likely to regress back to nouns with more Dutch neighbors (β=-0.015, se=0.0056, t=-2.60, p <. 01), but only when words were 9 characters long or longer (χ^2=3.86, df=1, p-value < 0.05). It was marginally less likely to make a regression back towards the noun when this noun had a higher frequent Dutch neighbor (β=0.51, se=0.27, t=1.89, p = .0591). This was only the case when the nouns bigram frequency was low (<2.62 log bigram frequency, χ^2=3.47, df=1, p-value =0.0625).

In sum, for L1 reading, having a higher frequent L2 neighbor makes it less likely that a regression will be made to the target word, at least when it had a high bigram frequency.Again, we found a facilitating effect of within-language N density for low frequent words and an inhibitory effect for high frequent words. This effect was also inhibitory for long words. Nouns with low bigram frequency that had a higher frequent Dutch neighbor were regressed to less often than when they didn't have a higher frequent neighbor.

L2 READING (ENGLISH).

EARLY MEASURES. The outcome of the final model for skipping probabilities and single fixation durations is presented in Table E.6 and E.7 in Appendix E. We fitted a logistic linear mixed model for skipping probability**.** For the single fixation analyses, only the nouns that received one fixation were selected (53.7%). Single fixation durations that differed more than 2.5 standard deviations from the subject means were excluded (2.14%).

CROSS-LINGUAL N EFFECTS. There was a main facilitating effect of cross-lingual N density for skipping probabilities (β=0.014, se=0.0064, t=2.14, p <. 05). More Dutch neighbors made it more likely that the target nouns would be skipped (see Figure 5). The main effect of cross-lingual N density was also significant for single fixation durations (β=-0.0026, se=0.0011, t=-2.33, p <. 05). The interaction of Dutch neighborhood density and word length was also significant (β=-0.0011, se=0.00052, t=-2.18, p <. 05, see Figure 8). This interaction showed that the facilitation was only significant for words 5 characters long or longer (χ^2=3.93, df=1, p-value < 0.05).

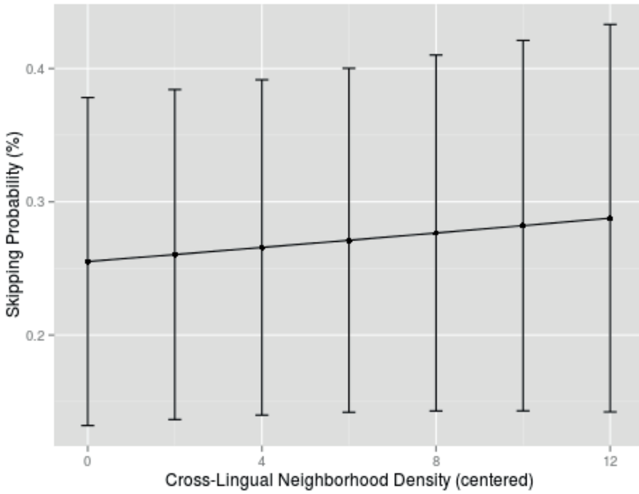


Figure 5. Skipping probability (y-axis) for nouns dependent on cross-lingual neighborhood density (x-axis) for English L2 reading.

WITHIN-LANGUAGE N EFFECTS. The main effect of within-language N density was significant for skipping rates (β=0.0089, se=0.0039, t=2.27, $p <. 05$). Targets with more neighbors were more to be skipped. The interaction of English N density and bigram frequency was significant (β=0.033, se=0.013, t=2.51, $p <. 05$). Linear contrast showed that the facilitation was only present for nouns with a high bigram frequency (>3.18 log bigram frequency, χ^2=4.13, df=1, p-value < 0.05). There was no effect of within-language neighborhood measures for single fixation durations.

In sum for L2 reading, we found facilitating effects of cross-lingual L1 N density on early language processing, particularly for long words. Within-lingual neighborhood density also had a facilitating effect for words with a high bigram frequency.

INTERMEDIATE MEASURES. The outcome of the final model for gaze durations is presented in Table E.8. Gaze durations that differed more than 2.5 standard deviations from the subject means were excluded (2.55%).

CROSS-LINGUAL N EFFECTS. We found a marginal facilitating effect of cross-lingual N density on gaze durations (β=-0.0029, se=0.0015, t=-1.95, p =

.0513). This measure interacted significantly with word length (β=-0.0016, se=0.00067, t=-2.39, p < .05, see Figure 6). For long nouns (6 characters or more) there was facilitation (χ^2=3.85, df=1, p-value < 0.05).

WITHIN-LANGUAGE N EFFECTS. There were no significant effects of within-language N measures for gaze durations.

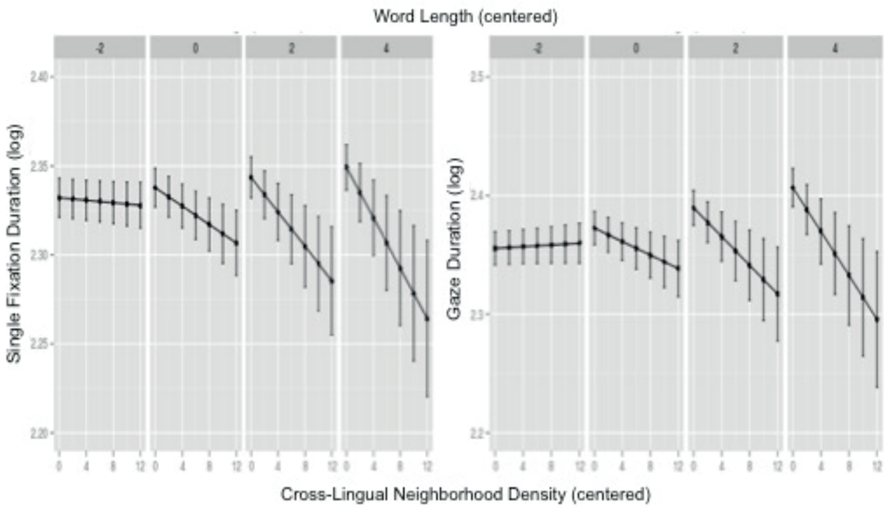


Figure 6. Single fixation durations (y-axis on first plot) and Gaze durations (y-axis on second plot) for nouns by cross-lingual neighborhood density (x-axis) and Word Length (panels) for English L2 reading.

LATE MEASURES. The outcome of the final model for total reading times and regression rates is presented in Table E.9 and E.10 in Appendix E. Total reading times that differed more than 2.5 standard deviations from the subject means were excluded (2.84%). For regression rate a logistic linear mixed model was fitted.

CROSS-LINGUAL N EFFECTS. We found a marginally significant facilitating effect of cross-lingual N density on total reading times (β=-0.0015, se=0.00091, t=-1.65, p = .0993). This variable interacted significantly with bigram frequency (β=-0.0059, se=0.0030, t=-1.99, p < .05, see Figure 7). Linear contrasts showed that the effect of cross-lingual neighborhood density was significantly facilitatory when bigram frequency was high (>3.3 log bigram frequency). We failed to find any effects of cross-lingual neighborhood measures on regression rates.

WITHIN-LANGUAGE N EFFECTS. We found a significant interaction effect of within-language N density and word frequency on total reading times (β=0.00096, se=0.00039, t=2.42, $p < .05$).This interaction indicated that the net result of having more within-language neighbors was inhibitory for reading high frequent English nouns (>3.48 log word frequency, χ^2=3.94, df=1, p-value < 0.05). There were no significant effects of any within-language N variables for L2 regression rates.

In sum, for L2 reading, we found L1 N density facilitation for high bigram frequent words in late recognition processes, whereas there was an inhibitory effect of L2 N density for high frequent nouns.

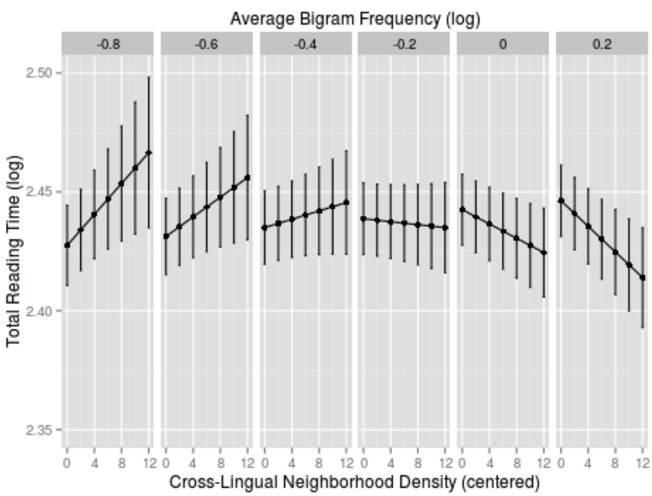


Figure 7. Total Reading Times (y-axis) for nouns dependent on cross-lingual neighborhood density (x-axis) and average bigram frequency (panels) for English L2 Reading.

MONOLINGUAL READING (ENGLISH).

To validate our neighborhood variables, we analyzed the eye movement towards nouns of monolinguals reading the same novel. These monolinguals were specifically selected as having no knowledge of any other language than English. None of the eye movement measures showed significant or marginally significant main effects of Dutch neighborhood density or frequency. Neither did any of the interactions between these measures and word frequency, word length or bigram frequency. We did find early facilitating effects of English neighborhood density and a late effect of English neighborhood frequency on regression rates. For full analyses see Appendix E.

DISCUSSION EXPERIMENT 2

CROSS-LINGUAL NEIGHBORHOOD EFFECTS

L1 READING. For L1 (Dutch) reading of high frequent nouns, the bilinguals showed an early facilitatory effect of cross-lingual N frequency. For low frequent nouns this effect reversed and became inhibitory. High frequent nouns with a larger L2 N frequency were skipped more often but low frequent nouns were skipped less often. Also, later in the reading process there were fewer regressions towards nouns with a higher frequent L2 neighbor. This facilitatory effect was larger for words with a high bigram frequency. In sum, for L1 reading, even in later stages of lexical identification and in post-lexical processes it seems that a higher frequent L2 neighbor facilitated reading. A higher frequent cross-lingual neighbor only had an inhibitory effect for low frequent target nouns very early in the reading process (fewer skips).

L2 READING. For L2 (English) reading, we found early facilitating effects of cross-lingual N density: nouns were skipped more often with increasing L1 N size and when fixated only once, these fixations were shorter. This facilitating effect was also found for gaze durations. The fact that the effects on single fixation duration and gaze duration were stronger for long words, might be an indication that lexical access was indeed facilitated by feedback from activated neighbors to letter representations, thus speeding up the identification especially for longer words. For total reading times we find marginal facilitation for nouns with increasing L1 N density. This facilitation was only significant for nouns with a high bigram frequency. Again this could be due to feedback towards letters and/or bigrams.

In summary, the cross-lingual effects found in our bilingual reading data were predominantly facilitatory. The only significant inhibitory effect of cross-lingual N was the effect of L2 N frequency found in the skipping rates when reading low frequent L1 nouns: A higher frequent L2 neighbor made it less likely that a target word would be skipped. There were a few trends towards inhibitory effects for nouns with a low bigram frequency. As we see in figure 4, reading such L1 nouns attracted more regressions when they had a higher frequent L2 neighbor. In L2 reading, nouns with a low bigram frequency showed longer total reading times with increasing L1 neighborhood density (see Figure 7). These effects did not reach significance, partly because of larger variance in the dependent variable for nouns with a lower bigram frequency.

WITHIN-LANGUAGE NEIGHBORHOOD EFFECTS

L1 READING. For L1 reading we found within language effects of N density for early (skipping rates and single fixation durations), intermediate (gaze durations) and late (total reading times) eye movement measures. The direction of these effects was largely determined by the word frequency of the target noun. For low frequent words, a larger N density seemed to facilitate the processing of that word. For high frequent words the opposite was the case: an increasing neighborhood density slowed down the reading of the target word. Also, short words were skipped more often with increasing N density, whereas long words were skipped less. These longer words were also regressed to less often when they had a larger neighborhood. Words with a higher frequent neighbor and a low bigram frequency were skipped slightly less and regressed to slightly less.

L2 READING. For English L2 reading we found an early facilitating effect and a late inhibiting effect of N density. More nouns were skipped when they had a larger neighborhood, but only when bigram frequency was high. The total reading times were longer for high frequent nouns when N density was large. There were no effects of within-language N frequency.

MONOLINGUAL READING. The analysis of English monolingual reading showed facilitatory effects of N density for early measures (skipping probability and single fixation durations) and inhibitory effects of N frequency for late measures (regression rate).

Our results for within-language neighborhood density are largely consistent with the results reported by Pollatsek et al. (1999). After controlling for the number of higher frequent neighbors, they found early facilitating effects of neighborhood density. Most of our early effects of neighborhood density were facilitating, except the effects for the L1 reading of high frequent, long words. Pollatsek et al.'s target nouns had a rather low word frequency (2.60 average log word frequency) and were rather short (average 4.5 characters). For the nouns with similar characteristics we also found early facilitating effects in our data.

We did not replicate the late inhibitory effects of neighborhood frequency found in Davis et al. (2009), Perea and Pollatsek (1998), or Slattery (2009) in our bilingual reading data. On the contrary, in the Dutch L1 reading data words with a higher frequent neighbor were regressed to less often than words without a higher frequent neighbor. This might be partially explained by the fact that in our data, participants read a coherent story in paragraphs, whereas in the other experiments participants read isolated sentences. This might make occasional misidentifications of the target word as a higher frequent neighbor less detrimental for the reading process, because the rich semantic context helps resolve possible errors without a second reading. For our English monolinguals, we did detect this inhibitory effect of neighbor frequency for regression rates. These monolingual English data contest the hypothesis, brought forward by Andrews (1997) and Sears et al. (2006), that there would be no inhibition from neighborhood frequency for English thus separating it from other alphabetic languages, like Spanish and Dutch. There is a possibility that the inhibitory effect of neighbor frequency is only detected in monolinguals, because bilinguals might have adapted their visual word recognition process slightly more to the larger lexical candidate sets. Further research into this specific finding should prove interesting.

A lot of the N effects are situated in the skipping rates. Facilitating effects in skipping rates of neighborhood density or frequency have been explained by misidentification of the target word with its higher frequent neighbor (Pollatsek et al., 1999; Slattery, 2009) instead of as a real reflection of faster lexical access. When we look at our Dutch L1 reading results, we observe similar effects in single fixation durations, gaze durations and reading times. We indeed find a higher correlation between skips and regressions ($r=0.55$) for nouns with a higher frequent neighbor than we do for nouns without one ($r=0.45$; $z=11.16$, $p<0.001$). But we find no positive correlation between the skipping rate for nouns with a higher frequent neighbor and the total reading time for these nouns ($r=-0.043$, $t=-4.12$, $df=9252$, p-value=1). These results show that it might be the case that a fraction of nouns was misidentified but these misidentifications do not have a significant effect on the total time spent on nouns with a high frequent neighbor.

In general our bilingual and our monolingual within-language reading data show, in accordance with Pollatsek et al. (1999) that there might be early facilitation from activation of letters/bigrams of lexical candidates and late inhibition in the selection phase.

GENERAL DISCUSSION

In this paper we investigated the effects of cross-lingual orthographic neighbors on bilingual language processing in two experiments. In experiment 1, the performance of Dutch-English bilinguals on a generalized lexical decision task was investigated, replicating Van Heuven et al. (1998). In experiment 2 a large database of eye movements during natural reading of a similar group (Cop et al., 2015) was analyzed.

For the data of experiment 1, the generalized lexical decision task, we conducted two different analyses. The first was an ANOVA analysis to make the comparison with the results reported in van Heuven et al. (1998). This analysis did not reveal any significant effects of cross-lingual neighborhood effects. In the second analysis, using LMER models and better measures for neighborhood density and frequency (Marian et al., 2008), we did find longer reaction times and more errors for L2 (English) words with increasing cross-lingual neighborhood density. For L1 (Dutch) words, error rates were higher for low frequent words, but lower for high frequent words with increasing cross-lingual neighborhood density. We can conclude that only with this second, more refined, analysis we replicated the most important result of van Heuven et al. (1998), namely the cross-lingual effect of neighbor density in a generalized lexical decision task for L2 words. In contrast to van Heuven et al., we found an effect of cross-lingual N density for L1 words in the error rates. This indicates that activation of cross-lingual lexical candidates is not confined to the processing of L2 words.

Although these cross-lingual effects are present in the reaction times and error rates of the generalized lexical decision task, it is not clear whether these effects would be found in a more unilingual context, because such a context might provide a cue to restrict lexical search and access to the target language (e.g. Van Assche et al., 2012). Also, lexical decision tasks have been criticized as good reflections of lexical access (e.g. Balota & Chumbley, 1984; Paap & Johansen, 1994). This is why we extended these findings to (a) a completely unilingual language context and (b) a more natural language context. Eye tracking during natural reading is perfectly suited for these goals. In experiment 2, a large database of bilingual eye movements (Cop et al., 2015) was analyzed to find evidence for activation of cross-lingual representations. The eye movements showed effects of cross-lingual neighborhood in early and late eye movement measures both for L1 and L2 natural reading. The pattern of results provides strong evidence that during natural reading, both in the early phase of lexical access as in the later language processes, written words activate not only orthographically similar words belonging to the target language but also those belonging to the non-target language of the bilingual. The absence of any cross-lingual neighborhood effects for English monolinguals strongly suggests that it was indeed the knowledge of a second language that produced these cross-lingual neighborhood effects in the bilingual participants, not some unknown lexical variable we failed to control.

In summary, both the results of experiment 1 and 2 provide evidence for parallel activation of lexical representations in bilingual word recognition and add strength to the argument of the existence of an integrated bilingual lexicon with language independent lexical access implemented in the BIA+ model (Dijkstra & van Heuven, 2002). We expected to see an asymmetry in cross-lingual effects (stronger effects in L2 than in L1 processing) because within the framework of the BIA+ model, L2 words should have a lower resting level of activation on average than L1 words for our population of unbalanced bilinguals. This means that L2 words need more time to be activated, thus making them more sensitive to influences of other activated lexical candidates (Dijkstra & van Heuven, 2002). And indeed, both in the lexical decision and the eye movement results, the cross-lingual N effects were more pervasive in L2. In the lexical decision task, L1 words only showed a cross-lingual effect in error rates, whereas L2 words showed effects in both error rates and reaction times. For the eye movements we see that for L1 reading cross-lingual N influenced only skipping rates and regression rates, whereas for L2 reading cross-lingual N effects were also present in the fixation duration data.

Within the BIA+ architecture, orthographic neighbors, both of the target and the non-target language, should influence lexical access to the target word by a complex interplay between inhibitory and excitatory connections at the word and letter level (Dijkstra & van Heuven, 2002). Simulations of the BIA+ have confirmed that cross-lingual neighborhood density effects could be inhibitory by means of lateral inhibition (Dijkstra et al., 1998). Within the BIA+ framework lateral inhibition from neighbors might be hidden by excitatory activation between representations for letters and words (as shown for the IA framework, Coltheart & Rastle, 1994; Pollatsek et al., 1999). Our data indeed shows that both inhibitory and facilitatory effects from neighbors are at play at the same time during word recognition.

In experiment 1, the cross-lingual N effects were mostly inhibitory: for high frequent L1 words error rates were higher and for L2 words reaction times were slower and error rates were higher with increasing cross-lingual N density. Only for the error rates for low frequent L1 words this effect was facilitatory. In contrast, the cross-lingual N effects in the reading data were mostly facilitatory, even in late language processing. Only for skipping rates for L1 low frequent words, we found an inhibitory effect of a higher frequent cross-lingual neighbor. This indicates that while performing a generalized lexical decision task, the cross-lingual activation generated by activated non-target language neighbors, slows performance, whereas in general, natural language reading benefits from this cross-lingual activation.

When interpreting this difference we have to keep in mind that the lexical decision task entails a decision component that might provoke different kinds of strategies in participants, leading to the masking of the real nature of the effect on lexical access (e.g. Paap & Johansen, 1994; Rayner & Liversedge, 2011). Lexical decision results have thus been shown to be very sensitive to blocking manipulations (e.g. van Heuven et al. 1998) and the selection of nonword stimuli (e.g. Lupker & Pexman, 2010; Stone & Van Orden, 1993). The fact that we found mostly inhibitory effects of neighborhood, whereas the results of most lexical decision experiments have found facilitating effects of N density for monolingual participants illustrates this sensitivity. For all of the reasons given above, we lend more importance to the effects found in the database of eye movements during natural reading.

Supporting the fact that the results of the generalized lexical decision task might not reflect lexical access alone, we found almost no overlap between findings of experiment 1 and 2, neither for within-language nor for cross-lingual neighborhood effects. For example, in the lexical decision task for Dutch words, we found no within-language N effects, whereas there was a cross-lingual N effect in error rates. For English words in the lexical decision task, we found an inhibitory effect of cross-lingual N density for error rates while this was facilitatory for within-language N density. A language system with an integrated lexicon, such as the BIA+ model, does not make a qualitative distinction between L1 and L2 lexical representations (Dijkstra & van Heuven, 2002). The partly activated neighbors from the target and non-target language should then have similar effects on target language word recognition or reading. In line with these expectations, for natural reading most of the cross-lingual effects resemble closely, although not exactly, the effects of within-language neighborhood effects. Where the patterns do diverge we see that this difference is driven by word frequency. We see here again a confirmation that natural reading might be a better approximation of lexical access than lexical decision.

The BIA+ architecture predicts effects of the frequency of the target word and the frequency of the neighbor words. Because the subjective frequency of representations determines the resting activation of these representations, this could change the complex interactions between excitatory and inhibitory effects of activated neighbors. In our analyses of L1 and L2 language processing, we used corpus word frequencies that are supposed to reflect the frequency of exposure to words for monolinguals (Kuperman & Van Dyke, 2013). In experiment 2 especially, because we investigated natural reading including a large range of nouns, we expect to find big effects of frequency variables. In classic experiment designs where stimuli are matched on these variables per condition, it is impossible to investigate these effects.

Indeed, in both experiments we find that the frequency and the bigram frequency of the target word modulate the neighborhood effects. In experiment 1, the effect of cross-lingual N density on L1 error rates and the effect of within-language N frequency on L2 reaction times were modulated by word frequency. In experiment 2, the effects of cross-lingual N frequency on skipping probabilities and within-language N density on early language processes in L1 reading are modulated by word frequency. In experiment 1 and 2 the effect of increasing N density was facilitatory for low frequent words and inhibitory for high frequent words. In both experiments, the effect of a more frequent neighbor was inhibitory for low frequent words and facilitatory for high frequent words. Apparently more low frequent neighbors speed up low frequent word processing, while they slow down high frequent word processing. The presence of a higher frequent neighbor is more inhibitory for low frequent words than for high frequent words. This has also been found in the monolingual reading studies of Davis et al. (2009), Perea and Pollatsek (1998) and Slattery (2009).

In experiment 1, the effect of cross-lingual N density on L2 reaction times and the effect of within-language N density on L2 error rates are modulated by bigram frequency.

In experiment 2, the effect of cross-lingual and within-language N frequency on regression rates in L1 reading, the effect of cross-lingual N density on total reading times and the effect of within-language N density on skipping probabilities in L2 reading were modulated by bigram frequency. In experiment 1, reaction times to L2 words were slower with increasing L1 N density, but only when the word had a low bigram frequency. In experiment 2, regression rates were lower for L1 nouns with a cross-lingual higher frequent neighbor and total reading times for L2 nouns were faster with increasing cross-lingual N density but only when the target word had a high bigram frequency. It seems that bigram frequency is important in determining the direction of cross-lingual N effects in word recognition. Lexical access to words with high frequent bigrams seem to be helped by having cross-lingual neighbors, whereas lexical access to words with low frequent bigrams is hindered. This might indicate that the high frequent bigrams activated in the neighbor words are feeding activation to the correct target word.

This is the first study investigating the effect of a higher frequent cross-lingual neighbor on word recognition. In experiment 1 we found no effect of cross-lingual N frequency above and beyond N density. In experiment 2, we found effects of cross-lingual neighborhood frequency in our L1 reading data, in the absence of any effects of cross-lingual N density. In L2 reading we found the opposite: There were only effects of cross-lingual neighborhood density, not N frequency. For Dutch L1 reading, the L2 neighbors seem to have to be of higher objective frequency than the target word before they are even known to our unbalanced participants. For English L2 reading the neighbors do not have to be of high frequency to have an effect, since the L1 neighbors will already be on average of higher frequency than the L2 target words (Dijkstra & van Heuven, 2002). It is unlikely that this difference is the result of a language difference because we do find an effect of neighborhood frequency for English monolingual reading.

Considering our own findings as well as other studies finding effects of addition, deletion and transposition neighbors (e.g. Blythe, Johnson, Liversedge, & Rayner, 2014; Davis et al.,2009), we believe it important that the BIA+ model should be modified to accommodate a more flexible letter position coding mechanism[19]. A mechanism lending itself for this purpose is the one proposed in the *overlap model* of Gomez et al. (2008). This model proposes that the representation of a letter is distributed across ordinal positions in the letter string. Every letter position has a specific standard deviation as free parameter in the model. This model expressively only models the letter coding mechanism, not any other higher order word recognition processes. This makes the overlap model easy to implement in other models, such as the BIA+ model. The effects of average bigram frequency in our data might also suggest that some kind of *open bigram coding* (Grainger & van Heuven, 2003), also implemented in the SERIOL model (Whitney, 2001), might be a good fit for these effects. Here words are coded by all of the ordered letter pairs that occur in that word. For example the word *hand* would be determined by the bigrams [ha, hn, hd, an, ad, nd].

In our opinion, the main architectural elements of the BIA+ model have promise in accommodating our most important results, namely the cross-lingual neighborhood effects found in natural reading, as long as a more flexible letter coding mechanism is implemented.

In conclusion, our lexical decision and natural reading data both provide convincing evidence for the existence of cross-lingual activation of lexical candidates during bilingual visual word recognition. Further research should

---

[19] Note that we did not report separate analyses for the effects of addition and deletion neighbors in our result section for reasons of brevity. This leaves the possibility that only substitution neighbors are responsible for the N effects. We did however conduct separate analyses with an N measure, only including addition and deletion neighbors, without substitution neighbors. In these analyses significant cross-lingual and within-language N effects were still found. For example in the bilingual L2 eye-tracking data a significant facilitating effect of Cross-lingual N density was found for skipping rates ($\beta=0.027$, se=0.013, t=2.08, $p < .05$).

focus on the lexical variables that modulate the size or the direction of these effects, such as the word frequency, both of the target word and its neighbors.

ACKNOWLEDGEMENTS

This research was supported by a grant from the FWO (Fonds voor Wetenschappelijk Onderzoek) and by a concerted research action project awarded by the special research fund of Ghent University.

REFERENCES

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition*, 24, 477–492.

Andrews, S. (1989). Frequency and Neighborhood effects on Lexical Access: Activation or Search? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(5), 802–814.

Andrews, S. (1992). Frequency and Neighborhood Effects on Lexical Access or Orthographic Redundancy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(2), 234–254. doi:10.1037//0278-7393.18.2.234

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review*, *4*(4), 439–461.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. doi:10.1016/j.jml.2007.12.005

Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX lexical data base on CD_ROM*. Philadelphia: Linguistic Data Consortium.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access: The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 340–357.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting Linear Mixed-Effects Models using lme4. *arXiv:1406.5823 [stat]*. Retrieved from http://arxiv.org/abs/1406.5823

Blythe, H. I., Johnson, R. L., Liversedge, S. P., & Rayner, K. (2014). Reading transposed text: effects of transposed letter distance and consonant-vowel status on eye movements. *Attention, Perception, & Psychophysics*, *76*(8), 2424–2440. doi:10.3758/s13414-014-0707-2

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41* (4), 977-990.

Bultena, S., Dijkstra, T., & van Hell, J. G. (2013). Cognate and word class ambiguity effects in noun and verb processing. *Language and Cognitive Processes*, *28*(9), 1350–1377. doi:10.1080/01690965.2012.718353

Caramazza, A., & Brones, I. (1979). Lexical Access in Bilinguals. *Bulletin of the Psychonomic Society*, *13*(4), 212–214.

Carreiras, M., Perea, M., & Grainger, J. (1997). Effects of orthographic neighborhood in visual word recognition: Cross-task comparisons. *Journal of Experimental Psychology: Learning Memory and Cognition*, *23*(4), 857–871. doi:10.1037/0278-7393.23.4.857

Clark, H. H. (1973). The Language-as-Fixed-Effect Fallacy : A Critique of Language Statistics in Psychological Research. *Journal of Verbal Learning and Verbal Behavior*, *12*, 335–359. doi:10.1016/s0022-5371(73)80014-3

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and Performance VI* (pp. 535–555). London: Academic Press.

Coltheart, M., & Rastle, K. (1994). Serial Processing in Reading Aloud - Evidence for Dual-Route Models of Reading. *Journal of Experimental Psychology:Human Perception and Performance*, *20*(6), 1197–1211. doi:10.1037/0096-1523.20.6.1197

Cop, U., Keuleers, E., Drieghe, D., & Duyck, W. (in press). Frequency Effects in Monolingual and Bilingual Natural Reading. *Psychonomic Bulletin and Review*, *20*, 963–972.

Davis, C. J., & Bowers, J. S. (2004). What Do Letter Migration Errors Reveal About Letter Position Coding in Visual Word Recognition? *Journal of Experimental Psychology: Human Perception and Performance*, *30*(5), 923–941. doi:10.1037/0096-1523.30.5.923

Davis, C. J., & Bowers, J. S. (2006). Contrasting five different theories of letter position coding: Evidence from orthographic similarity effects. *Journal of Experimental Psychology:Human Perception and Performance*, *32*(3), 535–557. doi:10.1037/0096-1523.32.3.535

Davis, C. J., Perea, M., & Acha, J. (2009). Re(de)fining the orthographic neighborhood: The role of addition and deletion neighbors in lexical decision and reading. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(5), 1550–1570. doi:10.1037/a0014253

Davis, C. J., & Taft, M. (2005). More words in the neighborhood: Interference in lexical decision due to deletion neighbors. *Psychonomic Bulletin & Review*, *12*(5), 904–910.

De Pessemier, P., & Andries, C. (2009). *GL&SCHR*. Leuven/Apeldoorn: Garant.

Dijkstra, T. (2007). The multilingual lexicon. In M. Gaskell (Ed.), *Handbook of Psycholinguistics* (pp. 251–265). Oxford University Press.

Dijkstra, T., Grainger, J., & van Heuven, W. (1999). Recognition of Cognates and Interlingual Homographs: The Neglected Role of Phonology. *Journal of Memory and Language*, *41*(4), 496–518.

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language*, *62*(3), 284–301. doi:10.1016/j.jml.2009.12.003

Dijkstra, T., & van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. Jacobs (Eds.), *Localist Connectionist Approaches to Human Cognition* (pp. 189–225). New York: Mahwah.

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(03). doi:10.1017/S1366728902003012

Dijkstra, T., van Heuven, W. J. B., & Grainger, J. (1998). Simulating cross-language competition with the bilingual interactive activation model. *Psychologica Belgica*, *38*(3-4), 177–196.

Duyck, W., Desmet, T., Verbeke, L. P. C., & Brysbaert, M. (2004). WordGen: A tool for word selection and nonword generation in Dutch, English, German, and French. *Behavior Research Methods Instruments & Computers*, *36*(3), 488–499. doi:10.3758/BF03195595

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 663–679. http://doi.org/10.1037/0278-7393.33.4.663

Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. Walker (Eds.), *New Approaches to Language Mechanisms* (pp. 257–287). Amsterdam.

Forster, K. I., & Shen, D. (1996). No enemies in the neighborhood: Absence of inhibitory neighborhood effects in lexical decision and semantic categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(3), 696.

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review*, *115*(3), 577–600. doi:10.1037/a0012667

Grainger, J. (1990). Word-Frequency and Neighborhood Frequency-Effects in Lexical Decision and Naming. *Journal of Memory and Language*, *29*(2), 228–244. doi:10.1016/0749-596X(90)90074-A

Grainger, J., & Jacobs, A. M. (1996). Orthographic Processing in Visual Word Recognition: A Multiple Read-out Model. *Psychological Review*, *103*(3), 518–565. doi:10.1037/0033-295X.103.3.518

Grainger, J., Oregan, J., Jacobs, A., & Segui, J. (1992). Neighborhood Frequency-Effects and Letter Visibility in Visual Word Recognition. *Perception & Psychophysics*, *51*(1), 49–56. doi:10.3758/BF03205073

Grainger, J., O'regan, J. K., Jacobs, A. M., & Segui, J. (1989). On the role of competing word units in visual word recognition: The neighborhood frequency effect. *Perception & Psychophysics*, *45*(3), 189–195.

Grainger, J., & Segui, J. (1990). Neighborhood Frequency-Effects in Visual Word Recognition - a Comparison. *Perception & Psychophysics*, *47*(2), 191–198. doi:10.3758/BF03205983

Grainger, J., & van Heuven, W. J. B. (2003). Modeling letter position coding in printed word perception. In P. Bonin (Ed.), *Mental lexicon: Some words to talk about words* (pp. 1–23). Hauppauge, NY: NOVA Science Publishers.

Grossi, G., Savill, N., Thomas, E., & Thierry, G. (2012). Electrophysiological Cross-Language Neighborhood Density Effects in Late and Early English-Welsh Bilinguals. *Frontiers in Psychology*, *3*. doi:10.3389/fpsyg.2012.00408

Guasch, M., Boada, R., Ferré, P., & Sánchez-Casas, R. (2013). NIM: A Web-based Swiss army knife to select stimuli for psycholinguistic studies. *Behavior Research Methods*, *45*(3), 765–71. doi:10.3758/s13428-012-0296-8

Holcomb, P., Grainger, J., & O'rourke, T. (2002). An electrophysiological study of the effects of orthographic neighborhood size on printed word perception. *Cognitive Neuroscience, Journal of*, *14*(6), 938–950.

Huntsman, L. A., & Lima, S. D. (1996). Orthographic neighborhood structure and lexical access. *Journal of Psycholinguistic Research*, *25*(3), 417–429. doi:10.1007/BF01727000

Huntsman, L. A., & Lima, S. D. (2002). Orthographic neighbors and visual word recognition. *Journal of Psycholinguistic Research*, *31*(3), 289–306. doi:10.1023/A:1015544213366

Jacobs, A., & Grainger, J. (1992). Testing a semistochastic variant o the interactive activation model in different word recognition experiments. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 1174–1188.

Johnson, N., & Pugh, K. (1994). A Cohort Model of Visual Word Recognition. *Cognitive Psychology*, *26*(3), 240–346. doi:10.1006/cogp.1994.1008

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: a new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, *42*(3), 643–50. doi:10.3758/BRM.42.3.643

Kuperman, V., Drieghe, D., Keuleers, E., & Brysbaert, M. (2013). How strongly do word reading times and lexical decision times correlate? Combining data from eye movement corpora and megastudies. *Quarterly Journal of Experimental Psychology, 66*, 563-580.

Kuperman, V. & Van Dyke, J.A. (2013). Reassessing word frequency as a Determinant of word recognition for skilled and unskilled readers. *Journal of Experimental Psychology, 39* (3), 802-823.

Laxon, V., Coltheart, V., & Keating, C. (1988). Children Find Friendly Words Friendly Too - Words with Many Orthographic Neighbors Are Easier to Read and Spell. *British Journal of Educational Psychology*, *58*, 103–119.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325–343. doi:10.3758/s13428-011-0146

Lupker, S.J., & Pexman, P.M. (2010). Making things difficult in lexical decision: The impact of pseudohomophones and transposed-letter nonwords on frequency and semantic priming effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35* (5), 1267-1289. doi: 10.1037/a0020125

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities. *PLOS ONE*, *7*(8), e43230. doi:10.1371/journal.pone.0043230

McClelland, J., & Rumelhart, D. (1981). An interactive activation model of context effects in letter perception. An account of basic findings. *Psychological Review*, *88*, 375–407.

Midgley, K. J., Holcomb, P. J., van Heuven, W. J. B., & Grainger, J. (2008). An electrophysiological investigation of cross-language effects of orthographic neighborhood. *Brain Research*, *1246*, 123–135. doi:10.1016/j.brainres.2008.09.078

Paap, K., & Johansen, L. (1994). The Case of the Vanishing Frequency Effect - a Retest of the Verification Model. *Journal of Experimental Psychology-Human Perception and Performance*, *20*(6), 1129–1157. doi:10.1037/0096-1523.20.6.1129

Paap, K., Newsome, S., McDonald, J., & Schvaneveldt, R. (1982). An Activation Verification Model for Letter and Word Recognition - The Word-Superiority Effect. *Psychological Review*, *89*(5), 573–594. doi:10.1037/0033-295X.89.5.573

Peereman, R., & Content, A. (1995). Neighborhood Size Effect in Naming - Lexical Activation or Sublexical Correspondences. *Journal of Experimental Psychology-Learning Memory and Cognition*, *21*(2), 409–421. doi:10.1037/0278-7393.21.2.409

Peeters, D., Dijkstra, T., & Grainger, J. (2013). The representation and processing of identical cognates by late bilinguals: RT and ERP effects. *Journal of Memory and Language*, *68*(4), 315–332. doi:10.1016/j.jml.2012.12.003

Perea, M., & Pollatsek, A. (1998). The effects of neighborhood frequency in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 767.

Pollatsek, A., Perea, M., & Binder, K. S. (1999). The Effects of "Neighborhood Size" in Reading and Lexical Decision. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(4), 1142–1158.

Rayner, K., & Liversedge, S. P. (2011). Linguistic and cognitive influences on eye movements during reading. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye movements* (pp. 751–766). UK: Oxford University Press.

Rayner, K., & Pollatsek, A. (1989). The psychology of reading. New York, NY: Prentice-Hall.

R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Scarborough, D., Gerard, L., & Cortese, C. (1984). Independence of Lexical Access in Bilingual Word Recognition. *Journal of Verbal Learning and Verbal Behavior*, *23*(1), 84–99. doi:10.1016/S0022-5371(84)90519-X

Schepens, J., Dijkstra, T., & Grootjen, F. (2012). Distributions of cognates in Europe as based on Levenshtein distance. *Bilingualism: Language and Cognition*, *15*(01), 157–166. doi:10.1017/S1366728910000623

Schneider, W., Eschman, A., & Zuccolotto, A. (2012). *E-Prime User's Guide*. Pittsburgh: Pschology Software Tools, Inc.

Sears, C. R., Campbell, C. R., & Lupker, S. J. (2006). Is there a neighborhood frequency effect in English? Evidence from reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(4), 1040–1062. doi:10.1037/0096-1523.32.4.1040

Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighborhood size and neighborhood frequency effects in word recognition. *Journal of*

Experimental Psychology: Human Perception and Performance, *21*(4), 876.

Slattery, T. J. (2009). Word misperception, the neighbor frequency effect, and the role of sentence context: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1969–1975. doi:10.1037/a0016894

Snodgrass, J. G., & Mintzer, M. (1993). Neigborhood effects in visual word recognition: Facilitatory or inhibitory? *Memory & Cognition*, *21*, 247–266.

Stone, G.O, & Van Orden, G.C. (1993). Strategic control of processing in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 19* (4), 744-774.

Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *Journal of Memory and Language*, *64*(1), 88–107. doi:10.1016/j.jml.2010.08.006

Van Assche, E., Duyck, W., & Brysbaert, M. (2013). Verb Processing by Bilinguals in Sentence Contexts. *Studies in Second Language Acquisition*, *35*(02), 237–259. http://doi.org/10.1017/S0272263112000873

Van Assche, E., Duyck W., & Hartsuiker, R.J. (2012). Bilingual word recognition in a sentence context. *Frontiers in Psychology,3 ,*174, doi:10.3389/fpsyg.2012.00174

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. *Psychological Science*, *20*(8), 923–927.

Van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic Neigborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language*, *39*, 458–483.

Van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2013). SUBTLEX-UK: a new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. doi:10.1080/17470218.2013.850521

Wilkinson, G. S., & Robertson, G. . (2006). *Wide Range Achievement Test 4 professional manual.* Lutz, FL: Psychological Assessment Resources.

Williams, C. C., Perea, M., Pollatsek, A., & Rayner, K. (2006). Previewing the neighborhood: The role of orthographic neighbors as parafoveal previews in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(4), 1072–1082. doi:10.1037/0096-1523.32.4.1072

APPENDIX A: FORMULA FOR ORTHOGRAPHIC OVERLAP

(The formula for the Corrected Levenshtein Distance (taken from Schepens, Dijkstra, & Grootjen, 2012).

$$Orthographic\ Overlap = 1 - \frac{Distance}{Length}$$

Distance = min (number of insertions, deletions and substitutions needed to edit target word into translation word)

Length = max (length of target word, length of translation word)

APPENDIX B: PROFICIENCY SCORES

Table B.1. *Average percentage scores (standard deviations between brackets and range between square brackets) on the LexTALE. Average rating on the self-report questionnaire (standard deviations between brackets).*

		Dutch	English	t-value L1-L2
LexTALE-score (%)		87.58 (7.03) [70.00-96.25]	73.04 (9.08) [57.50-88.75]	6.519***
Self Report				
	Listening	4.9 (0.4)	4 (0.58)	5.141***
	Speaking	4.87 (0.34)	3.5 (0.612	7.628***
	Reading	4.9 (0.3)	3.93 (0.63)	5.604***
	Writing	4.8 (0.48)	3.43 (0.72)	6.899***
	Average	4.87 (0.29)	3.72 (0.47)	7.523***

* p<0.05, **p<0.01, ***p<0.001

Participants had on average a higher proficiency for Dutch then English, both on the LexTALE, $t(29) = 7.518$, $p < .001$, and the average self-proficiency ratings, $t(29) = 10.891$, $p < .001$.

Table B.2. *Average percentage scores (standard deviations between brackets and range between square brackets) on the LexTALE, Spelling test and Lexical Decision task for the bilingual and monolingual group in experiment 2.*

	Monolinguals	Bilinguals L1	Bilinguals L2	t-value L1-L2	t-value L1-mono
LexTALE- score (%)	91.07(8.92) [71.25-100]	92.43 (6.34) [73.75-100]	75.63(12.87) [51.25-98.75]	7.59 ***	0.49
Spelling score (%)	80.78 (7.26) [73.81-90.48]	83.16(7.80) [67.00-93.00]	69.92 (8.74) [52.00-83.00]	8.15 ***	0.99
Lexical Decision score (%)	77.89 (12.01) [54.61-95.23]	80.47 (5.45) [68.87-88.76]	56.75 (11.01) [38.46-75.86]	9.87 ***	0.67

* p<0.05, **p<0.01, ***p<0.001

The Dutch (L1) proficiency of the bilinguals was matched with the English proficiency of the monolinguals indicating that both groups were equally proficient in their first language. Neither the LexTALE (t=0.488, df=22.254, p=0.630), the spelling test (t=0.989, df=29.282, p=0.331), nor the lexical decision tasks (t=0.667, df=17.092, p=0.514) yielded significant differences for these two groups performing in L1. The bilingual L2 LexTALE scores were significantly lower than their L1 scores (t=7.587, df=18, p<0.001). The bilingual L2 Spelling scores were lower than the L1 scores (t=8.154, df=18, p<0.001). The performance of the bilinguals on the classic lexical decision task was significantly better in L1 (t=9.873, df=18, p<0.001) than in L2.

APPENDIX C: STIMULUS MATERIALS EXPERIMENT 1

DUTCH WORDS

Large Dutch N, Large English N. Bons, borg, bril, dolk, hiel, klam, knie, oord, plek, rund, sein, spar, takt, tolk, vork, wolk, worp, woud, wrak, zalf

Large Dutch N, Small English N. Berg, beul, bouw, deun, dief, eter, fuik, kelk, kies, knal, kous, rede, snik, teug, touw, twee, unie, vals, verf, vies

Small Dutch N, Large English N. Brug, bult, draf, drie, fris, galg, hemd, heup, lach, meid, melk, munt, nota, pret, prik, smid, stug, vete, welp, wilg

Small Dutch N, Small English N. Akte, ambt, blad, erwt, ezel, gesp, gids, gips, inkt, joch, muts, ober, pech, pion, rots, snor, stro, toga, trui, veld

ENGLISH WORDS

Large Dutch N, Large English N. Aunt, blue, farm, hawk, knit, left, loan, loud, maid, monk, moon, path, quit, shoe, suit, tool, verb, weak, wrap, zero

Large Dutch N, Small English N. Army, atom, bias, bird, diet, edge, germ, huge, butt, jerk, keen, knee, liar, lion, myth, noon, nude, obey, poem, poor

Small Dutch N, Large English N. Bath, bomb, busy, clue, coin, desk, dial, dirt, dish, firm, grey, hurt, iron, joke, lamb, limb, loss, milk, prey, rude

Small Dutch N, Small English N. Deny, duty, earl, envy, evil, folk, frog, guts, idol, kiss, okay, oral, oval, soup, true, twin, ugly, used, vein, view

Nonwords

Large Dutch N, Large English N. Aril, aunk, blag, boul, boup, braf, bret, dris, duef, elap, fram, frip, furk, gonk, heud, jeef, knat, knub, koup, loem, meem, merd, mots, oram, peit, pern, piot, pral, pred, rama, sluf, sluk, snus, sols, stui, tess, trum, tult, vene, zork

Large Dutch N, Small English N. Alof, besp, bito, bouf, daus, drot, epoe, etel, feik, goep, grul, heut, irok, jees, jeul, jund, jurf, kalp, kelf, kerd, keun, loga, morp, muig, mups, nazz, noge, nont, noto, obel, oune, pris, puif, reug, reun, slen, smir, viem, woup, zuls

Small Dutch N, Large English N. Aute, bele, bulf, ceot, chah, cham, clet, dolo, drid, dulp, feul, foug, fran, genk, girs, jant, jero, jert, liry, lurd, lurp, lusp, naul, nirk, nudo, orim, pani, prad, prog, puet, raut, reud, rion, ruze, seto, snam, tirk, tran, vich, vorn

Small Dutch N, Small English N. Aler, anas, arns, aurd, baun, cafa, chof, deim, dilm, drio, durs, enip, fenk, feup, frig, frus, giep, heif, hilp, jalp, jofe, kach, kiot, knaf, luet, maup, moug, nige, omil, paby, ridi, siom, taur, torp, tuni, twol, unar, vota, zous, zuke

APPENDIX D: LMER RESULTS OF EXPERIMENT 1

Table D.1. *Estimates, standard errors, t-values and p-values for the fixed and random effects of the final general linear mixed effect model for reaction times for Dutch words.*

Dutch Words					
	Estimate	SE	t-value	p-value	
Fixed Effects					
Intercept	2.806	0.0156	179.90	<.001	***
Dutch N density	-0.00006	0.0014	-0.04	.969	
English N density	0.00001	0.0012	0.01	.991	
Dutch N Frequency	-0.0019	0.0119	-0.16	.876	
Higher frequent English N	-0.0003	0.0125	-0.02	.981	
Word frequency	-0.0391	0.0069	-5.66	<.001	***
Levenshtein distance	-0.0073	0.0167	-0.44	.664	
Average bigram frequency	0.0163	0.0176	0.93	.357	
	Variance	SD			
Random Effects					
Word					
(Intercept)	0.001	0.032			
Subject					
(Intercept)	0.003	0.051			
p<0.1 . p<0.05 * p<0.01 ** p<0.001***					

Table D.2. *Estimates, standard errors, z-values and p-values for the fixed and random effects of the final general linear mixed effect model for error rate for Dutch words.*

Dutch Words					
	Estimate	SE	z-value	p-value	
Fixed Effects					
Intercept	-3.30	0.47	-7	< 0.001	***
Dutch N density	0.05	0.05	1.19	0.233	
English N density	-0.01	0.04	-0.19	0.849	
Dutch N Frequency	-0.22	0.39	-0.58	0.56	
English N Frequency	0.06	0.39	-0.58	0.88	
Word frequency	-1.22	0.24	-5.14	< 0.001	***
Levenshtein distance	-0.15	0.54	-0.28	0.781	
Average bigram frequency	-0.5	0.55	-0.93	0.355	
English N density * Word frequency	0.13	0.07	1.93	0.053	.
	Variance	SD			
Random effects					

Word			
(Intercept)	0.443	0.666	
Subject			
(Intercept)	0.389	0.624	
p<0.1 . p<0.05 * p<0.01 ** p<0.001***			

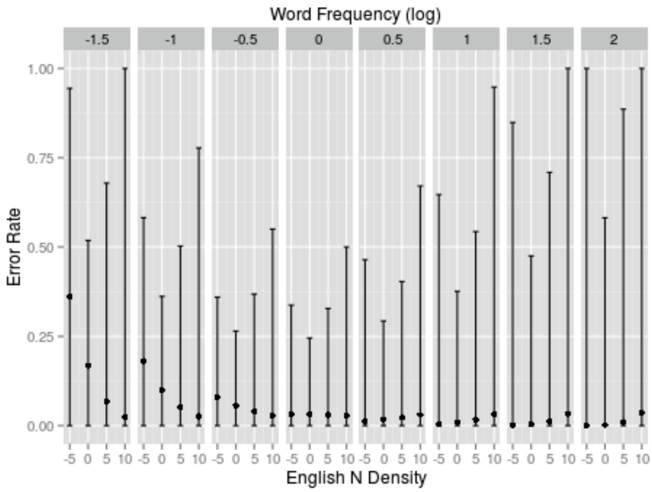


Figure D.1. Error rate (on the y-axis) for Dutch words dependent on English N density (on the x-axis) dependent on word frequency of the word (panels) for a generalized lexical decision task.

Table D.3. *Estimates, standard errors, t-values and p-values for the fixed and random effects of the final general linear mixed effect model for reaction times for English words.*

English Words					
	Estimate	SE	t-value	p-value	
Fixed Effects					
Intercept	2.807	0.0128	218.66	<.001	***
Dutch N density	0.0013	0.0011	1.15	.254	
English N density	0.00007	0.0011	0.07	.946	
Dutch N Frequency	0.0087	0.0079	1.10	.277	
English N Frequency	0.0081	0.087	0.94	.353	
Word frequency	-0.0076	0.0129	-0.59	.556	
Levenshtein distance	-0.0117	0.0124	-0.95	.349	
Average bigram frequency	0.0250	0.0258	0.97	.337	

	Dutch N density * Average bigram frequency	-0.0134	0.0066	-2.04	.046	*
	English N density * Word Frequency	0.0044	0.0025	1.75	.086	.
	English N Frequency * Word Frequency	-0.0402	0.016	-2.51	.015	*
	English N Frequency * Average bigram frequency	-0.0564	0.0315	-1.79	.078	.
		Variance	SD			
Random Effects						
	Word					
	(Intercept)	0.0005	0.021			
	Subject					
	(Intercept)	0.0027	0.052			
p<0.1 . p<0.05 * p<0.01 ** p<0.001***						

Table D.4. Estimates, standard errors, z-values and p-values for the fixed and random effects of the final general linear mixed effect model for error rates for English words.

English Words						
		Estimate	SE	z-value	p-value	
Fixed Effects						
	Intercept	-3.50	0.40	-8.71	< .001	***
	Dutch N density	0.10	0.04	2.32	.021	*
	English N density	-0.07	0.05	-1.35	.177	
	Dutch N Frequency	-0.06	0.33	-0.18	.857	
	English N Frequency	0.228	0.40	0.57	.566	
	Word frequency	-0.48	0.43	-1.10	.270	
	Levenshtein distance	0.05	0.50	0.11	.914	
	Average bigram frequency	-1.36	0.67	-2.01	.044	*
	English N density * Average bigram frequency	-0.29	0.17	-1.65	0.099	.
	English N Frequency * word frequency	-1.01	0.55	-1.84	0.065	.
		Variance	SD			
Random Effects						

Word				
(Intercept)	0.542	0.737		
Subject				
(Intercept	0.487	0.698		
p<0.1 . p<0.05 * p<0.01 ** p<0.001***				

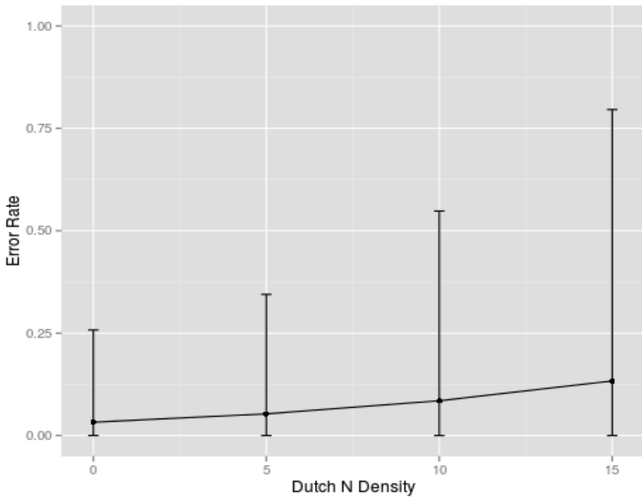


Figure D.2. Error rates (on the y-axis) for English words dependent on Dutch N density (on the x-axis) in a generalized lexical decision task.

APPENDIX E: LMER RESULTS OF EXPERIMENT 2

Table E.1. *Estimates, standard errors, z-values and p-values for the fixed and random effects of the final general linear mixed effect model for Skipping Rate for bilingual L1 reading*.

Bilingual L1					
	Estimate	SE	z-value	p-value	
Fixed Effects					
(Intercept)	-0.97	0.11	-8.89	<.001	***
Dutch N Density	-0.0027	0.0079	-0.34	.736	
Dutch N Frequency	0.0033	0.042	0.079	.937	
English N Density	0.0033	0.0081	0.41	.684	
English N Frequency	0.010	0.050	0.21	.836	
Word Frequency	0.080	0.021	3.76	<.001	***

Word Length	-0.27	0.014	-19.06	<.001	**
Average Bigram Frequency	0.086	0.070	1.22	.222	
Dutch N Density * Word Frequency	-0.017	0.0046	-3.71	<.001	***
Dutch N Density * Word Length	-0.012	0.0034	-3.53	<.001	***
Dutch N Frequency * Word Length	0.038	0.022	1.73	.0807	.
English N Frequency * Word Frequency	0.11	0.046	2.46	.0140	*
Control variables					
Orthographic Overlap	0.011	0.054	0.21	.832	

	Variance	SD
Random Effects		
Word		
(Intercept)	0.029	0.17
Subject		
(Intercept)	0.20	0.45
p<0.1 . p<0.05 * p<0.01 ** p<0.001***		

Table E.2. *Estimates, standard errors, t-value and p-values for the fixed and random effects of the final linear mixed effect model for Single Fixation Durations for bilingual L1 reading.*

Bilingual L1					
	Estimate	SE	t-value	p-value	
Fixed Effects					
(Intercept)	2.31	0.010	220.38	<.001	***
Dutch N Density	-0.00049	0.00043	-1.14	.255	
Dutch N Frequency	-0.00011	0.0031	-0.036	.971	
English N Density	-0.000022	0.00068	0.033	.974	
English N Frequency	0.0041	0.0042	0.97	.331	
Word Frequency	-0.0092	0.0015	-6.11	<.001	***
Word Length	0.0029	0.00072	4.07	<.001	***
Average Bigram Frequency	0.0067	0.0056	1.20	.230	
Dutch N Density * Word Frequency	0.0015	0.00034	4.30	<.001	***
Control variables					
Orthographic Overlap	-0.0067	0.0042	-1.58	.114	

	Variance	SD
Random Effects		
Word		
(Intercept)	0.00038	0.019
Subject		

(Intercept)	0.0019	0.044
p<0.1 . p<0.05 * p<0.01 ** p<0.001***		

Table E.3. *Estimates, standard errors, t-value and p-values for the fixed and random effects of the final linear mixed effect model for Gaze Durations for bilingual L1 reading.*

Bilingual L1					
	Estimate	SE	t-value	p-value	
Fixed Effects					
(Intercept)	2.31	0.025	91.58	<.001	***
Dutch N Density	-0.00064	0.00049	-1.31	.190	
Dutch N Frequency	0.000066	0.0034	0.019	.985	
English N Density	0.000017	0.00079	0.022	.982	
English N Frequency	0.0026	0.0048	0.54	.588	
Word Frequency	-0.012	0.0017	-7.45	<.001	***
Word Length	0.0069	0.00078	8.90	<.001	***
Average Bigram Frequency	0.0050	0.0063	0.80	.424	
Dutch N Density * Word Frequency	0.0023	0.00038	6.12	<.001	***
Control variables					
Orthographic Overlap	-0.0067	0.0047	-1.41	.158	

	Variance	SD
Random Effects		
Word		
(Intercept)	0.00061	0.025
Subject		
(Intercept)	0.0028	0.053
p<0.1 . p<0.05 * p<0.01 ** p<0.001***		

Table E.4. *Estimates, standard errors, t-value and p-values for the fixed and random effects of the final linear mixed effect model for Total Reading Times for bilingual L1 reading.*

Bilingual L1					
	Estimate	SE	t-value	p-value	
Fixed Effects					
(Intercept)	2.38	0.014	175.29	<.001	***
Dutch N Density	-0.000099	0.00061	-0.16	.870	
Dutch N Frequency	0.00082	0.0043	0.19	.849	
English N Density	-0.00058	0.00099	-0.058	.953	
English N Frequency	-0.0020	0.061	-0.32	.749	
Word Frequency	-0.017	0.0021	-8.16	<.001	***
Word Length	0.010	0.00097	10.63	<.001	***

Average Bigram Frequency	0.012	0.0079	1.50	.133	
Dutch N Density * Word Frequency	0.0028	0.00048	5.87	<.001	***
Control variables					
Orthographic Overlap	-0.0030	0.0059	-0.51	.609	

	Variance	SD
Random Effects		
Word		
(Intercept)	0.0011	0.033
Subject		
(Intercept)	0.0032	0.057
p<0.1 . p<0.05 * p<0.01 ** p<0.001***		

Table E.5. *Estimates, standard errors, z-values and p-values for the fixed and random effects of the final general linear mixed effect model for Regressions for bilingual L1 reading*

Bilingual L1					
	Estimate	SE	z-value	p-value	
Fixed Effects					
(Intercept)	-2.22	0.11	-20.77	<.001	***
Dutch N Density	-0.014	0.015	-0.89	.374	
Dutch N Frequency	-0.013	0.072	-0.18	.859	
English N Density	-0.020	0.017	-1.17	.244	
English N Frequency	-0.18	0.099	-1.80	.0718	.
Word Frequency	-0.086	0.033	-2.60	.00921	**
Word Length	-0.13	0.025	-5.50	<.001	***
Average Bigram Frequency	0.23	0.17	1.37	.171	
Dutch N Density * Word Length	-0.015	0.0056	-2.60	.00936	**
Dutch N Frequency * Average Bigram Frequency	0.51	0.27	1.89	.0591	.
English N Frequency * Average Bigram Frequency	-0.55	0.33	-1.68	.0924	.
Control variables					
Orthographic Overlap	0.019	0.10	0.19	.849	

	Variance	SD
Random Effects		
Word		
(Intercept)	0.23	0.48
Subject		
(Intercept)	0.16	0.39
p<0.1 . p<0.05 * p<0.01 ** p<0.001***		

Table E.6. *Estimates, standard errors, z-values and p-values for the fixed and random effects of the final general linear mixed effect model for Skipping Rate for bilingual L2 reading.*

Bilingual L2					
	Estimate	SE	z-value	p-value	
Fixed Effects					
(Intercept)	-1.088	0.12	-8.75	<.001	***
English N Density	0.0089	0.0039	2.27	.0235	*
English N Frequency	0.030	0.031	0.97	.332	
Dutch N Density	0.014	0.0064	2.14	.0324	*
Dutch N Frequency	-0.0097	0.036	-0.27	.787	
Word Frequency	0.14	0.018	8.04	<.001	***
Word Length	-0.20	0.013	-15.23	<.001	***
Average Bigram Frequency	-0.15	0.077	-1.92	.0548	.
English N Density * Average Bigram Frequency	0.033	0.013	2.51	.0121	*
Control variables					
Orthographic Overlap	0.12	0.050	2.49	.0129	*

	Variance	SD
Random Effects		
Word		
(Intercept)	0.026	0.16
Subject		
(Intercept)	0.28	0.53
p<0.1 . p<0.05 * p<0.01 ** p<0.001***		

Table E.7. *Estimates, standard errors, t-value and p-values for the fixed and random effects of the final linear mixed effect model for Single Fixation Durations for bilingual L2 reading.*

Bilingual L2					
	Estimate	SE	t-value	p-value	
Fixed Effects					
(Intercept)	2.34	0.011	213.25	<.001	***
English N Density	0.00016	0.00034	0.46	.647	
English N Frequency	0.0029	0.0025	1.14	.254	
Dutch N Density	-0.0026	0.0011	-2.33	.0205	*
Dutch N Frequency	0.0036	0.0031	1.15	.251	
Word Frequency	-0.018	0.0014	-12.28	<.001	***
Word Length	0.0029	0.0014	2.10	.0363	*
Average Bigram Frequency	0.016	0.0066	2.46	.0143	*
Dutch N Density * Word Length	-0.0011	0.00052	-2.18	.0302	*
Control variables					
Orthographic Overlap	-0.0044	0.0040	-1.10	.273	

	Variance	SD
Random Effects		
Word		
(Intercept)	0.00027	0.017
Subject		
(Intercept)	0.0022	0.047
p<0.1 . p<0.05 * p<0.01 ** p<0.001***		

Table E.8. *Estimates, standard errors, t-value and p-values for the fixed and random effects of the final linear mixed effect model for Gaze Durations for bilingual L2 reading.*

Bilingual L2					
	Estimate	SE	t-value	p-value	
Fixed Effects					
(Intercept)	2.33	0.030	76.49	<.001	***
English N Density	0.00033	0.00045	0.744	.457	
English N Frequency	0.0019	0.0032	0.61	.542	
Dutch N Density	-0.0029	0.0015	-1.95	.0513	.
Dutch N Frequency	0.0018	0.0040	0.46	.649	
Word Frequency	-0.017	0.0017	-9.87	<.001	***
Word Length	0.0086	0.0017	5.02	<.001	***
Average Bigram Frequency	0.013	0.0083	1.60	.11	
Dutch N Density * Word Length	-0.0016	0.00067	-2.39	.0171	*
Control variables					
Orthographic Overlap	-0.0030	0.0048	-0.62	.534	

	Variance	SD
Random Effects		
Word		
(Intercept)	0.00089	0.030
Subject		
(Intercept)	0.0036	0.060
p<0.1 . p<0.05 * p<0.01 ** p<0.001***		

Table E.9. *Estimates, standard errors, t-value and p-values for the fixed and random effects of the final linear mixed effect model for Total Reading Times for bilingual L2 reading.*

Bilingual L2					
	Estimate	SE	t-value	p-value	
Fixed Effects					
(Intercept)	2.44	0.015	162.04	<.001	***
English N Density	0.00077	0.00052	1.49	.138	
English N Frequency	0.0033	0.0037	0.89	.376	

Dutch N Density	-0.0015	0.00091	-1.65	.0993	.
Dutch N Frequency	0.0031	0.0047	0.66	.509	
Word Frequency	-0.030	0.0022	-13.95	<.001	***
Word Length	0.014	0.0014	10.17	<.001	***
Average Bigram Frequency	0.019	0.0096	1.96	.0504	.
English N Density * Word Frequency	0.00096	0.00039	2.42	.0157	*
Dutch N Density * Average Bigram Frequency	-0.0059	0.0030	-1.99	.0467	*
Control variables					
Orthographic Overlap	-0.0094	0.0056	-1.67	.0960	.

	Variance	SD
Random Effects		
Word		
(Intercept)	0.0012	0.035
Subject		
(Intercept)	0.0041	0.064
p<0.1 . p<0.05 * p<0.01 ** p<0.001***		

Table E.10. *Estimates, standard errors, z-values and p-values for the fixed and random effects of the final general linear mixed effect model for Regressions for bilingual L2 reading*.

Bilingual L2					
	Estimate	SE	z-value	p-value	
Fixed Effects					
(Intercept)	-2.18	0.11	-19.17	<.001	***
English N Density	0.0063	0.0071	0.88	.378	
English N Frequency	0.039	0.053	0.74	.462	
Dutch N Density	0.0071	0.012	0.59	.556	
Dutch N Frequency	0.019	0.065	0.29	.776	
Word Frequency	-0.12	0.029	-4.24	<.001	***
Word Length	-0.071	0.019	-3.74	<.001	***
Average Bigram Frequency	0.14	0.13	1.03	.303	
Control variables					
Orthographic Overlap	-0.10	0.083	-1.22	.224	

	Variance	SD
Random Effects		
Word		
(Intercept)	0.17	0.41
Subject		
(Intercept)	0.22	0.47
p<0.1 . p<0.05 * p<0.01 ** p<0.001***		

CHAPTER 7
GENERAL DISCUSSION

In this dissertation we presented the Ghent Eye tracking Corpus (GECO): a large database of eye movements of bilinguals and monolinguals reading an entire novel. We provided a large and comprehensive account of how bilinguals (and monolinguals) read an expansive narrative text. We also disclosed the corpus to other researchers to use it for their specific research purposes. This corpus might be used in an exploratory way, yielding interesting research questions and insights into bilingual reading. Because of the large amount of text material included in the novel, this corpus can also be used for more specific hypothesis testing. Also, we hope this dataset can help evaluate the generalizability of models of reading and language processing and can ultimately boost theorizing about bilingual language processing. In the final section of this chapter, we highlight the most interesting research possibilities of the GECO.

The main empirical goal of this dissertation was the investigation of bilingual lexical access in contextual language comprehension. The eye tracking method is particularly suited for this goal, because it is the most naturalistic method available for studying the time-course of lexical access: there are no irrelevant cognitive processes involved, unlike other often-used methods such as naming and lexical decision (Grainger & Jacobs, 1996). Below we provide an overview of the empirical findings.

OVERVIEW OF THE FINDINGS

In CHAPTER 3 we conducted two important comparisons of the global sentence-level eye movement patterns found in the GECO. The first one was the comparison of the eye movement pattern of unbalanced Dutch-English bilinguals reading in their first language (L1) and their second language (L2). Because of the lower amount of language exposure to a less used and later acquired second language, it has been hypothesized that lexical entrenchment might be lower for the second than in the first language of a bilingual person and that this might lead to slower lexical access (e.g., Diependaele, Lemhöfer, & Brysbaert, 2013). In the GECO, we indeed found that Bilinguals were slower when reading L2 sentences than L1 sentences. They made more fixations, which each lasted longer, and skipped fewer words when reading in their L2 than when reading in their L1. The durations of the fixations that these readers made were also more sensitive to the length of the words in L2 than in L1. These changes from L1 reading to L2 reading largely resembled the changes from adult to child-like reading reported is several eye tracking studies (e.g., Blythe, Liversedge, Joseph, White, & Rayner, 2009; Rayner, 1986). An exception to this overlap was the regression rate. It is usually found that children regress backwards in the text more often than adults do (e.g., Blythe, Häikiö, Bertam, Liversedge, & Hyönä, 2011; Buswell, 1922), this was not found for the bilinguals' L2 reading in the GECO. This difference could be explained by the fact that children have worse working memory (e.g., Cowan, 1998; Cowan, Morey, AuBuchon, Zwilling, & Gilchrist, 2010) and have to rely on a less extensive semantic network for text comprehension than adults do. This might cause more text integration difficulties for children than for bilingual adults, leading to a higher regression rate in the eye movement record of children. We showed in CHAPTER 3 that the global changes for L2 reading could be interpreted as a slowing of the process of lexical access, by framing our findings within the E-Z reader model.

The second important comparison was the one between L1 reading for bilinguals versus monolinguals. According to the weaker links account, bilinguals necessarily divide the frequency of language use between their two known languages (Gollan et al., 2011; Gollan, Montoya, Cera, & Sandoval, 2008). This would lead to weaker links between the lexical and semantic representations for both languages of the bilingual person compared to links in the language system of the monolingual. Contrary to predictions made by this account, bilinguals showed a largely similar (and equally fast) global eye movement pattern relative to monolinguals as reported in CHAPTER 3. This implies that the ‘weakening’ of links between lexical and semantic representations for all bilinguals is not as universal as implied by some researchers (Gollan et al., 2008).

In CHAPTER 4, we investigated the effect of word frequency on the single fixation durations for content words in the GECO, both for bilinguals reading in L1 and L2 and for monolinguals reading in L1. Word frequency accounts for a large portion of the variance in word recognition behavior (Brysbaert et al., 2011), and it is thus perhaps the primary lexical variable that influences lexical access, also for bilingual reading. Repeated exposure to a certain word raises the baseline activation of the relevant representation, so that lexical selection of that particular word becomes faster during reading (e.g., Monsell, 1991). The maximal speed of lexical access is limited, so once a word has received a certain amount of exposure, no more facilitation will be expected when there is additional exposure to that particular item (Morton, 1970). This means that lower language exposure would predict a larger difference between the reading time of low and high frequent words. In the tradition of the weaker links account, it has been hypothesized that bilingual frequency effects should be larger than monolingual frequency effects because of a smaller language exposure to both languages of the bilingual person (Gollan et al., 2011).

Word frequencies are usually defined by a count of the occurrence of the target word over a large amount of texts (Keuleers, Brysbaert, & New, 2010; van Heuven, Mandera, Keuleers, & Brysbaert, 2014). It has been shown that for monolinguals with a small vocabulary, these corpus word frequencies are inaccurate, especially for low frequent words (Kuperman & Van Dyke, 2011). We generalized this idea to bilinguals, who have smaller vocabularies in L2 than in L1. Our analysis of the GECO demonstrated that our group of bilinguals had larger frequency effects in L2 than in L1. Because we log-transformed the corpus word frequencies for analyses, this effect was not due to an absolute lower exposure to all words, but rather to an overestimation of the relative frequency for low frequent words by corpus word frequencies. Monolingual and bilingual frequency effects were equally large in L1, implying similarly sized L1 vocabularies, which was confirmed by the results of the LexTALE test. The combination of similarly sized frequency effects and vocabulary for bilinguals and monolinguals in L1 show that the weaker links account (Gollan et al., 2008) does not hold for this group of bilinguals either because the assumption of lower exposure is false or because this lower exposure does not necessarily lead to weaker links in the language system. We will discuss this issue in more depth in the next section of this chapter.

Both for bilingual L1 and L2 and monolingual reading, the size of the frequency effect was influenced by L1 vocabulary size: a larger vocabulary reduced the slope of the frequency effect. Again, this showed that participants with a smaller vocabulary had a lower relative frequency for low frequent words, leading to slower lexical access to these words. These results showed that the same functional relationship exists between lexical access and language exposure for all language users

In CHAPTER 5 and CHAPTER 6, we aimed to find stronger evidence for language independent lexical access by searching for markers of cross-lingual activation in the eye movement patterns. We have studied the recognition of identical and non-identical cognate nouns by bilingual participants, both in the L1 and L2 GECO in CHAPTER 5. Cognates are words that show (some) orthographic overlap with their translation equivalent word in the other language. An example of a non-identical cognate is *appel* in Dutch, meaning *apple* in English. An example of an identical cognate is the Dutch word *piano*. These words are investigated often in the context of evidence for parallel lexical access because if these words are processed differently than control words it means that the knowledge of another languages affects reading in the target language. A lot of studies have shown faster processing of identical cognates compared to control words (e.g., Duyck, Van Assche, Drieghe, & Hartsuiker, 2007; Lemhöfer & Dijkstra, 2004), and some even show facilitation with increasing orthographic overlap (e.g., Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010; Van Assche, Drieghe, Duyck, Welvaert, & Hartsuiker, 2011; Van Assche, Duyck, Hartsuiker, & Diependaele, 2009). This means that bilingual participants process the word ‘apple’ (appel in L1) faster than the word ‘’lie’ (leugen in L1), solely because the word ‘apple’ is more orthographically similar to its translation. Furthermore, Dijkstra et al. (2010) showed that identical cognates had a discontinuous processing advantage compared to non-identical cognates in a lexical decision task. They viewed this as evidence for one shared orthographic representation for identical cognates across languages.

In our analyses of L2 reading, first fixations were shorter for words with more cross-lingual orthographic overlap with their translation equivalent word. Also, nouns were more likely to be skipped with rising cross-lingual orthographic overlap. Additionally, identical cognates were fixated longer in first pass reading and total reading and were regressed to more often (indicated by longer go past times) than other words. In L1 reading, first fixations were shorter for words with more cross-lingual orthographic overlap, but only when these words were more than 9 letters long. Also, total reading times of identical cognates were shorter, but only when the cognates were highly frequent. This is the first demonstration of cognate facilitation in L1 reading in a naturalistic narrative text, i.e. without the use of deliberately constructed sentences that pose low semantic constraints on the cognate nouns. We did not find the discontinuous advantage for identical cognates as Dijkstra et al. (2010) did. In order to explain the early linear effects of orthographic overlap in our results, we hypothesized excitatory weighted links between translation equivalent orthographic representations in different languages. The late facilitation for identical cognates might be explained by stronger semantic feedback of these words to the correct orthographic representation. Of course, these results strongly support the claim of language independent lexical access of the BIA+ model (Dijkstra & van Heuven, 2002), even in a unilingual L1 context.

In CHAPTER 6 we used a similar approach as in CHAPTER 5 and examined more directly whether lexical access is limited to lexical candidates from the target language when reading isolated words and text. According to the most prominent models of visual word recognition, such as the Interactive Activation (IA) model (Mcclelland & Rumelhart, 1981) and the Multiple Readout model (MROM) (Grainger & Jacobs, 1996), during reading the reader has to choose the correct lexical representation from a pool of orthographically similar representations. These similar words are referred to as *neighbors* (Coltheart, Davelaar, Jonasson, & Besner, 1977). The number of neighbors of a target word and the frequency of the neighbor words affect target word reading times (e.g., Perea & Pollatsek, 1998; Pollatsek, Perea, & Binder, 1999). The IA model explains this effect by assuming lateral inhibition among the lexical candidates. An account of bilingual parallel lexical access assumes that for bilinguals this pool of lexical candidates would extend to the non-target language (Dijkstra & van Heuven, 2002; van Heuven & Dijkstra, 1998).

To directly test this hypothesis, we used a generalized lexical decision task and the GECO. For L2 word recognition, we found slower reaction times in the generalized lexical decision task for words with more cross-lingual neighbors. In contrast, in the eye movement record, we found that L2 nouns were read faster and skipped more often when they had more cross-lingual neighbors. This shows that when processing L2 words, in or out of a sentence context, lexical candidates of L1 are active and either inhibit or facilitate lexical access towards the target word. For L1 isolated word recognition, we found higher error rates for high frequent and lower error rates for low frequent words with more cross-lingual neighbors in the generalized lexical decision task. In text reading, high frequent nouns with a higher cross-lingual neighborhood frequency were skipped more often but low frequent nouns were skipped less often. Also, there were fewer regressions towards words with a higher cross-lingual neighborhood frequency. The GECO showed considerable cross-lingual effects of neighborhood density and frequency, which were largely facilitatory, whereas results from the generalized lexical decision task point towards inhibitory effects of cross-lingual neighbors. These findings strongly suggest that lexical candidates are selected based on orthographic similarity, not language membership, and that in extensive text reading in most cases these lexical candidates from the non-target language actually help speed up lexical access to the target word.

Although these results provide evidence for bilingual language independent lexical access as proposed by the BIA+ (Dijkstra & van Heuven, 2002), our results are not compatible with the model's assumption of fixed letter coding. In the BIA+ for each letter of a word, the position is perfectly encoded. This assumption about location-specific letter processing was copied from the IA model (Mcclelland & Rumelhart, 1981). As letter position determines which words are orthographically similar to each other it determines the pool of lexical candidates, or what words are considered as neighbors. In a letter coding system as proposed by the BIA+ model, only substitution neighbors would affect lexical access to the target word. However, we showed that deletion and addition neighbors affect reading times for target words. This means that the BIA+ model might have to be modified to accommodate a more flexible letter coding mechanism such as the one proposed in the overlap model by Gomez, Ratcliff, and Perea (2008).

To summarize, the results of the empirical chapters of this dissertation imply that unbalanced bilinguals showed a cost when reading in an L2 compared to when reading in an L1, which is most outspoken for low frequent words, i.e. words that have received few exposures. The results also show that there is no cost of being a bilingual for L1 reading: lexical access in reading narrative text is not necessarily slower for bilinguals than for monolinguals. Together, these findings imply that the slowing of lexical access in L2, reported in CHAPTER 3 and 4, and the slowing of bilingual L1 word recognition found in other studies (Lehtonen et al., 2012; Ransdell & Fischler, 1987) but not here, could be linked to individual differences in language representations, not to group differences in these representations. We propose in this final chapter that language proficiency/exposure largely determines an individual's speed of lexical access, not whether a person is bilingual or monolingual. Our results of CHAPTER 5 and 6 also clearly show that lexical access for bilinguals is fundamentally language-independent in a naturalistic reading setting, both in L1 and L2. This is remarkable because both of the reading contexts (L1 and L2) were completely unilingual. The text that was read by the participants consisted of a coherent set of sentence contexts, which were representative for daily extensive narrative reading.

This shows that in a set of unilingual sentence contexts, ranging in length and complexity, lexical access is still language independent.

BILINGUAL READING IN CONTEXT: INTEGRATION OF EMPIRICAL FINDINGS

In CHAPTERS 3 to 6 of this dissertation, different aspects of bilingual lexical access in a narrative reading context were investigated. In the next section we will take a critical look at these findings across these chapters, which used different analytic methods on the same data, and discuss some of the striking findings emerging from this research effort.

ABSENCE OF A MONOLINGUAL ADVANTAGE

For a long time, the metaphor of the brain as a finite storage room or limited-capacity system has been dominant in ideas about learning. Operating within this framework, the notion that whatever cognitive skill or knowledge a person gains in one area, that person must lose in another, has lingered in our thinking about human performance. This is why with the rise of a more bilingual society in the beginning of the 20th century, many researchers focused on detecting the possible disadvantages that being a bilingual could entail, such as smaller L1 vocabularies (Grabo, 1931; Harris, 1948) and worse scores on general intelligence tests (Darcy, 1953). Epstein (1905) even referred to multilingualism as a 'social plague'. Although we cannot deny that there are some limits on processes in the brain, recent evidence about the scope of our brain capacity undercuts the notion of finite memory storage entirely: the capacity of the human brain to store knowledge and learn new skills is enormous (e.g., Caine & Caine, 1991) and an estimate by Wang, Liu, and Wang (2003) of this capacity resulted in a computational solution which is a *magnitude that is very much higher than the total memory capacity of all computers ever available in the world (Wang et al., 2003, page 197)*. The idea of a restricted capacity for language competence seems therefore rather obsolete.

However, some language theories do conjure up similar ideas in respect to multilingualism. For example, the weaker links account (Gollan et al., 2008) proposes that what we gain in the second language (i.e. language exposure) we must lose in our first language. This leads to weaker links between representations in the bilingual language system. Admittedly, the authors do not claim that the links are weaker because the brain's capacity is limited, but because language exposure is limited and will on average be equal over participants, but is therefore distributed across languages if multiple languages are present. We proposed in CHAPTER 3 and 4 that this is not necessarily the case: There might be bilingual groups who, because of their bilingualism, have larger language exposure than others. Also, the weaker links account does not discuss the possibility that connections between L1 representations, once sufficiently strong, may not need a sustained level of exposure, to remain equally strong.

It must be said that the authors of the weaker links account are mostly inspired by research using a different kind of bilinguals than those tested in this dissertation. The bilinguals in this dissertation are unbalanced, i.e. they are dominant in their L1 and they started learning an L2 at an average age of eleven. Those used by the weaker links account were bilinguals whose proficiency is more balanced across the languages, and it makes sense that their language use is divided about equally across languages. Also, most of these bilinguals learned their second language rather early. For example, Gollan and Acenas (2004) tested Spanish-English bilinguals living in San Diego. Their mother tongue was Spanish, but they learned English at an average age of three. The first difference is that the late bilinguals in our GECO might be characterized by a rise in total language exposure due to their late second language learning. The second difference is that these bilinguals had already acquired a rich vocabulary in their L1 and therefore have had a large amount of L1 exposure before learning a second language. This might make these representations more robust and more resistant against a drop in language exposure, than the L1 representations of early bilinguals, on which the weaker links data is based.

When we look at the overall results presented in this dissertation, it is clear that all participants, including the bilinguals, performed rather satisfactory in the reading of a complete novel. In fact, the scores on the multiple-choice questions were the same for the bilinguals reading in L1 (79.6%), reading in L2 (79.0%) and the monolinguals (78.3%). This indicates that all participants understood the story equally well. When comparing the scores of a broad range of language proficiency indicators, such as a spelling and a vocabulary test, bilinguals scored equally well in L1 as the monolinguals did. Bilinguals did show a disadvantage in these measures when tested in L2. Seeing that vocabulary scores can be directly related to language exposure (see CHAPTER 4 for a more detailed explanation), this could be interpreted as an indication that bilinguals do not necessarily *lose* L1 language exposure to an L2. Another interpretation is that they do lose some L1 language exposure in favor of L2 exposure, but because this happened after the age of eleven it did not affect vocabulary learning as such, because language exposure is much more important for vocabulary expansion in the early phases of language acquisition (see figure 1 in CHAPTER 4).

When looking at the eye movement record, we did not find any clear sentence-level differences between monolingual and bilingual L1 reading (CHAPTER 3) and we found frequency effects of equal size for these two groups (CHAPTER 4). This does not mean that bilingual L1 language processing does not entail different mechanisms than monolingual language processing does. As illustrated in CHAPTER 5 and 6 of this dissertation, cross-lingual activation is present in bilingual L1 reading, whereas it is not in monolingual reading, because the monolingual lexicon does not hold any lexical representations from another language than the target language. Interestingly, the results reported in CHAPTER 3 and 4 show that these cross-lingual activations did not have a detrimental effect on global reading speed or the size of the frequency effect, despite the fact that the lexicon that has to be scanned for word recognition is almost twice as large for bilinguals.

To summarize, late bilinguals read equally fluent in their L1 as monolinguals do. Although there is clearly activation of non-target lexical candidates during the process of lexical access of the correct target word, this does not cause a net disadvantage for bilingual reading. These findings indicate that for our group of bilinguals, either L1 language exposure does not suffer under L2 language exposure or L1 language exposure does not need to be at the same level to sustain the strength of already formed connections between L1 semantic and orthographic representations. It would be interesting to investigate how other groups of bilinguals, such as early or more balanced bilinguals, would perform in an extensive reading task in their L1. Although they have acquired their second language at a later age, our bilingual group had no problem in comprehending an L2 text equally well as an L1 text, although their eye movements showed that lexical access might be slightly slower. This is a very positive and encouraging finding in a society that values multilingualism and encourages young people to learn foreign languages.

IMPORTANCE OF LANGUAGE EXPOSURE

In the process of first language development, children need extensive exposure to language in order to become a proficient language user. Similarly, in order to acquire a second language, an adult or child must be exposed to this second language repeatedly. This means that language exposure might be of great importance in explaining individual differences in reading. To get an idea of how much exposure our participants have had in their L1 and their L2, we measured language proficiency with a spelling test, a vocabulary test (LexTALE), a lexical decision task and a self report questionnaire (see Table 1 in CHAPTER 3) for all participants. Our bilinguals were less proficient in their second language than in their mother tongue for all of these different measures. The same tests revealed no difference between the L1 proficiency for the bilingual and the monolingual participants.

In the analysis of global eye movements (CHAPTER 3), we found a significant effect of L2 proficiency on the number of fixations per sentence in bilingual reading, both in L1 and L2: a bilingual made fewer fixations when their L2 proficiency score was higher. For saccade length, we detected a nearly significant effect of L1 proficiency for both bilingual and monolingual L1 reading: bilinguals made longer saccades when their L1 proficiency score was higher. It is difficult to interpret these results, because we did not include both L1 and L2 proficiency in the analysis. Because of the high, significant correlation between these two variables, we do not know whether it is the score on L1 or the score on L2 proficiency, which drove the effects in this analysis. Also, in this analysis we used a composite proficiency score of all of the proficiency measures, making it unclear which of the proficiency components were most important in determining these eye movements. What is clear is that the effects of proficiency were not modulated by the factor of language or bilingualism, implying that the relationship between language exposure and reading is similar for the two groups of participants and the two languages of the bilingual participants.

In CHAPTER 4 we analyzed all non-cognate content words of the novel. We detected a significant modulation of the size of the frequency effect by L1 proficiency: With increasing L1 proficiency, the frequency effect was reduced. This modulation was the same for bilinguals reading in L1 and L2, and for monolinguals. L2 proficiency influenced bilingual reading times: A higher score on L2 proficiency yielded shorter fixations for L2 reading and longer fixations for L1 reading. In these analyses, L1 and L2 proficiency were both included in the statistical model, making it possible to dissociate their effects. Here, L1 and L2 proficiency was operationalized by the respective LexTALE scores, a measure of vocabulary size, making it possible to draw conclusions about the relationship between language exposure and frequency effects.

In CHAPTER 5, we analyzed all nouns, including identical cognates. We detected a main effect of L1 and L2 proficiency on bilingual skipping rates. A bilingual participant was more likely to skip a noun when this participant scored high on L1 proficiency and less likely to skip a noun when he scores high on L2 proficiency. Again these effects were not dependent on the reading language.

In all of these different analyses, L1 proficiency facilitated word and sentence processing. This relationship between L1 proficiency and language processing remains the same across groups of participants (bilinguals and monolinguals) and even across languages (bilingual L1 and L2 reading). Because we never detected any statistically significant effect of the group factor of bilingualism, we propose that variation in L1 proficiency might be sufficient to explain individual differences in the speed of lexical access in L1 reading within the bilingual and monolingual group. This means that for our population, we believe quantitative differences rather than qualitative differences govern L1 reading patterns. These results underline the necessity of the measurement of L1 language proficiency in any investigation of group differences in visual word recognition, especially in the bilingual field. This does not happen often; very few monolingual studies use L1 proficiency measures in the investigation or analysis of frequency or neighborhood effects and in most bilingual studies of visual word recognition only L2 proficiency or exposure is measured (e.g., Gollan et al., 2011; Whitford & Titone, 2012) As L2 and L1 proficiency are usually correlated (.69 for the LexTALE scores in our sample), the effects are not distinguishable when only one of these is included in the statistical model. Also, we encourage researchers to measure objective instead of subjective language proficiency, for reasons of accuracy and comparative usefulness.

L2 proficiency seems to have more diverse effects on language processing. On a sentence level we see facilitatory effects of L2 proficiency on fixation counts in both L1 and L2 reading, although, as said above, this effect could partly be caused by L1 proficiency, because this variable was not included in this analysis. On a word level, for content words we see that a higher score on L2 vocabulary was beneficial for L2 reading times, but had the reverse effect for L1 reading times, whereas for nouns, L2 proficiency lengthened fixation times in both L1 and L2. It is fascinating to see that L2 proficiency can have both a negative as well as a positive effect on L1 and L2 reading. Future research, controlling both L1 and L2 proficiency, could help to disentangle these effects

It is important to note that L1 and L2 proficiency measures did not account for all variability between L1 and L2 bilingual reading. This is in line with findings by Whitford and Titone (2012). The age of acquisition of the second language might account for the remaining differences, but it is a challenge for future researchers to identify the variables that could fully explain these differences.

LANGUAGE INDEPENDENT LEXICAL ACCESS

The only implemented model of lexical access during bilingual reading, the BIA+ model (Dijkstra & van Heuven, 2002) assumes that lexical access is language independent. Language nodes are included in the architecture of the model, but these cannot be directly influenced by non-linguistic activation outside of the lexicon (e.g. expectation of a certain language), and they cannot send top down activation backwards to the lexical representations in order to tune lexical search to a specific language (Dijkstra & van Heuven, 2002). This means that while reading, all of the orthographically similar lexical representations in the bilingual person's lexicon become activated, regardless of to which language they belong. The last two empirical chapters of this dissertation focused on whether bilinguals' lexical access is still language-independent while reading under naturalistic conditions, that is in a completely unilingual and coherent semantic context. The answer to this question is a convincing 'yes'. This is an exciting finding, because in this situation, readers could use the language of the sentence or the semantic constraints of the sentence as a cue to restrict lexical access to one language or a part of the lexicon. Our results highly suggest that this is not the case: a) In CHAPTER 5 we showed that identical and non-identical cognates are processed differently than words without orthographic overlap with their translation equivalent in bilingual L1 and L2 reading, b) The results of CHAPTER 6 show even more strongly that both target language and non-target language lexical candidates, often referred to as orthographic neighbors, are activated in bilingual L1 and L2 reading.

As the results of both CHAPTER 5 and 6 demonstrate, the activation of non-target representations mostly *facilitates* lexical access to the correct target representation. The facilitating effects reported in these chapters were larger in L2 than in L1. It is not clear whether this can be attributed to the dominance of the language, i.e. the fact that this language is the most proficient one, or to the global age of acquisition. This is because in our population these two factors are the same: all bilinguals were dominant in their first acquired language, Dutch. Even though the effects were larger in L2 reading, the effects were still reliable in L1 reading. This is striking because, as said before, for our unbalanced group of bilinguals this language is acquired only at age eleven and is a lot less proficient than their mother tongue. Nevertheless, the representations of this late L2 affect the activation of the L1 dominant representations. This implies that the activation of non-target lexical candidates should be very strong if they are able to alter and sometimes even slow the activation of strong L1 representations.

As described above, we found mostly facilitating effects of cross-lingual activation in the GECO. There are exceptions. In CHAPTER 5 we found that during L1 reading, total reading times were longer for identical cognates, but only for low frequent cognates. Similarly in CHAPTER 6, in L1 reading we detected an inhibitory effect for low frequent words: these words were skipped less often when they had a higher frequent cross-lingual neighbor. Apparently, lexical access to low frequent L1 words could be hampered by the activation of L2 lexical representations. This was a rather unexpected finding. If there would be inhibition, we expected it to arise in L2 reading. This is because L2 words would be less frequently encountered than L1 words, making the threshold for activation higher for L2 words, giving the L1 representations more change to interfere with L2 recognition. L1 neighbors would thus provide strong lateral inhibition for the processing of the correct L2 word. It would be interesting to pursue this line of thinking in some more controlled experimental design.

In conclusion, learning a second language has a profound impact on L1 language processing, but this influence is not necessarily detrimental. As shown, most cross-lingual effects make it easier to process words that show overlap with non-target language, with the exception of low frequent L1 words., and at least for the two related (German) languages Dutch and English. Reading of the acquired L2 is also impacted by the knowledge of L1. This influence is surprisingly, mostly beneficial for reading speed.

THE B(ILINGUAL) - E-Z READER MODEL

We will attempt to frame some of our empirical findings in the best known model of eye movements during reading, the E-Z Reader model (Reichle, Pollatsek, Fisher, & Rayner, 1998). The most important features of the E-Z reader model are that it a) assumes that attention is allocated serially, b) the programming of saccades is decoupled from attention, c) the process of lexical access has two stages; a first familiarity check which triggers the programming of a saccade and a second completion of lexical access which shifts the attention towards the next word.

First, we look at the global eye movement pattern of L2 reading reported in CHAPTER 3. According to the E-Z reader model, when the programming of an intra-word saccade is completed faster than the familiarity check of the fixated word, an intra-word fixation is made (Reichle, 2011). A higher number of re-fixations of words when the bilingual participants read in L2 compared to L1, could thus be related to a slower familiarity check, the first phase of lexical access. According to the E-Z reader model, if the familiarity check takes longer in L2, the eyes should rest longer at the same location. Indeed, the average fixation duration was also longer for L2 reading compared to L1 reading. Another prediction of the E-Z reader model is that when the familiarity check of a parafoveal word is completed before the saccade programming to that word is completed, this word will be skipped (Reichle, 2011). In line with the idea that the familiarity check can be completed faster when reading in the mother tongue than when reading in L2, we observed that the skipping probability is lower when reading in L2 than in L1.

In conclusion, all of the findings reported in CHAPTER 3 are consistent with a more effortful familiarity check and slower overall lexical processing for bilinguals reading in L2. This familiarity check is dependent on word frequency and exposure, which is lower for the L2 of our unbalanced population of bilinguals. In CHAPTER 4 we indeed show that for low frequent words, the L2 disadvantage in reading time is larger than it is for high frequent words. These results combined, show that the bilingual L2 disadvantage in visual language processing might be reduced to a quantitative difference of exposure to the lexical items in the lexicon. Reichle et al. (2013) already showed that the eye movement pattern of children could be modeled by simply reducing the rate of lexical processing. Given that we established a close parallel between patterns of eye movement in children and L2 readers in CHAPTER 3, we hypothesize the same, although smaller, adjustment to the E-Z reader model parameters could possibly also model the L2 reading pattern of unbalanced bilinguals.

FUTURE USES OF THE EYE TRACKING CORPUS

The GECO can be used to study many aspects of bilingual language behavior in a naturalistic reading context. We outline the research ideas we find most interesting below.

FURTHER INVESTIGATION OF THE BILINGUAL LEXICON

Many elements in the organization of the bilingual lexicon, are yet unclear. We discuss two ways in which the GECO can be used to clarify certain representational issues in the bilingual domain.

CROSS-LINGUAL HOMOGRAPHS

Cross-lingual homographs are unique in that they have the exact same orthography (like identical cognates) but do not share meaning (like orthographic neighbors). An example of a cross-lingual homograph is the orthographic word form ‘glad’ meaning slippery in Dutch and contented in English. In the study of the selectivity of bilingual lexical access, multiple studies have investigated the processing of cross-lingual homographs in semantic priming tasks (Beauvillain & Grainger, 1987), lexical decision tasks (de Groot, Delmaar, & Lupker, 2000; Dijkstra, Grainger, & van Heuven, 1999; Dijkstra, Van Jaarsveld, & ten Brinke, 1998; Gerard & Scarborough, 1989; Lemhöfer & Dijkstra, 2004), progressive demasking (Dijkstra et al., 1999), translation recognition (de Groot et al., 2000) and reading (Libben & Titone, 2009). Most of these studies detected a difference in the processing of homographs compared to control words, implying that the non-target language had an influence in the processing of a target language, thus providing evidence for language independent access. This is a similar approach as we have used in CHAPTER 5 and 6 of this dissertation. It is not yet clear whether homographs have shared or separate orthographic representations across languages (Dijkstra et al., 1999), but most evidence points to the latter option (Lemhöfer & Dijkstra, 2004). Also, Libben and Titone (2009) found very early inhibitory effects of homographs in eye movements, showing that the investigation of these words in the GECO could yield some interesting results. Investigating the processing of these words in conjunction with cognates could shed light on the exact way in which the different levels of representations are linked.

WORD LEVEL AGE OF ACQUISITION

Many bilingual studies report global Age of Acquisition (AoA), typically defined as the age at which a bilingual is first exposed to their second language. For example in our corpus study, the average AoA was eleven for our bilingual participants. Some studies have linked this variable to the speed of lexical access in L2 word recognition (e.g., Canseco-Gonzalez et al., 2010; Silverberg & Samuel, 2004).

There is however another use of this concept. The AoA of individual words in the domain of monolingual language processing refers to the age at which a specific word has been learned. Values for AoA have generally been collected by large subjective rating studies with adult participants who are asked when they learned a certain word (Spanish: Angeles Alonso, Fernandez, & Diez, 2015; Dutch: Brysbaert, Stevens, De Deyne, Voorspoels, & Storms, 2014; English: Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012). Although this seems like a rather unreliable method, these subjective values show a high inter-rater reliability and correlate highly with more objective measures of when certain words are learned (e.g., Carroll & White, 1973; Morrison, Chappell, & Ellis, 1997), which implies that these ratings are a valid measure of AoA. Sometimes more objective measures of AoA are gathered either from child behavior (e.g., Zevin & Seidenberg,

2002) or from parental reports on child vocabulary use (e.g., Kittredge, Dell, Verkuilen, & Schwartz, 2008).

Carroll and White (1973) were the first to show that word-level AoA had a direct effect on lexical access during a picture naming task, independent of the effect of word frequency. Later, this finding was confirmed in word naming tasks (Barry, Hirsh, Johnston, & Williams, 2001; Brown & Watson, 1987; Brysbaert, Lange, & Wijnendaele, 2000; Morrison & Ellis, 1995) and lexical decision tasks (e.g., Butler & Hains, 1979; Gerhand & Barry, 1998; Morrison & Ellis, 1995). Some authors have attributed these effects to the fact that lexical representations are of better quality when learned within the critical age period. Other authors presented findings which are not compatible with such an account: Yamazaki, Ellis, Morrison, and Ralph (1997) showed effects of AoA after the critical period. Brysbaert, Van Wijnendaele, and De Deyne (2000) hypothesized that the order of word acquisition is an important organizing factor in the language system, because word meaning depends on previously acquired meanings. According to this account, the AoA effect would originate from the semantic representations, but this is still a matter of debate (e.g. Ellis & Lambon Ralph, 2000; Zevin & Seidenberg, 2002)

A way to investigate the origins of AoA effects is to look at these effects in the L2 of late bilinguals, because these bilinguals acquired their L2 after the critical period of language acquisition. Also, the semantic account by Brysbaert, Van Wijnendaele, et al. (2000) would predict that when acquiring L2 representations, which likely share semantics with L1 representations, the AoA characteristics associated with these representations, would generalize to the newly acquired L2 word forms. This means that L1 AoA would heavily influence L2 performance. In the bilingual domain, few studies have investigated the role of word level AoA on L2 word recognition. The first study, by Izura and Ellis (2002), tested late Spanish-English bilinguals on a picture naming and a lexical decision task. They used subjective ratings of 132 words, rated by 28 participants with the same global AoA and language history. After this, the same authors, using the same AoA ratings and the same bilingual population, tested participants with a translation judgement task (Izura & Ellis, 2004). The last study was conducted by Montrul and Foote (2014). They investigated the interaction between global and word-level AoA for English-Spanish bilinguals in a lexical decision and a translation decision task. They tested two groups of English-Spanish bilinguals differing in their global L2 AoA. In all of these experiments, L2 performance on a lexical decision task was influenced by L2 AoA, not by L1 AoA, showing that it was unlikely that AoA effects arise from the semantic representations (in contrast to the account by Brysbaert, Van Wijnendaele, et al., 2000). Their results were compatible with a mapping hypothesis (Ellis & Lambon Ralph, 2000), in which early acquired words play a more important role than later acquired words in the organization of the lexicon.

The effects of word-level AoA have never been investigated using eye movements during L2 reading. As eye tracking is a powerful tool to disentangle early and late language processes, we believe it would shed light on the exact locus of AoA effects. Our research lab has already started gathering AoA ratings for all nouns in the GECO, and we hope to present these results soon.

SENTENCE CONTEXT EFFECTS

Achieving lexical access is not the only task an accomplished reader must tackle. A reader has to use the information that is currently available to predict upcoming words and integrate the lexical representation in the syntactic structure of a sentence, in order to truly understand text. Below we outline two promising research projects concerning bilingual context effects.

PREDICTION IN NATURALISTIC L2 READING

In sentence reading, the upcoming word and the features of that word are predicted in function of what is currently being processed (e.g., DeLong, Urbach, & Kutas, 2005; Otten & Van Berkum, 2008), and this anticipation process is fundamental for the fluency of language comprehension (Pickering & Garrod, 2007). The predictability of a word given the sentence context has an influence on the processing time of that word, and therefore also on fixation times and skipping rates recorded with eye tracking (e.g., Drieghe, Brysbaert, & Desmet, 2005; Frisson, Rayner, & Pickering, 2005; Rayner & Well, 1996). It is not yet clear however, whether this process is equally important for bilinguals reading in L2 as it is for people reading in L1. The sentence-level reading disadvantage for L2, compared to L1, reported in CHAPTER 3, as well as other evidence that L2 processing is delayed compared to L1 processing (Frenck-Mestre, German, & Foucart, 2014), might be partly due to the fact that the ability of bilinguals to predict words and their features is reduced when processing an L2.

Many studies have shown that L2 processing is sensitive to sentence context (Altarriba, Kroll, Sholl, & Rayner, 1996; Duyck et al., 2007; Libben & Titone, 2009; Van Assche et al., 2011; van Hell & de Groot, 2008). In these experiments, L2 readers show easier processing of target words in a high-constraint context. Because these studies looked at reading times, fixation and gaze durations of the target words of the sentences, not to eye movements before the actual viewing of the target word, Foucart, Martin, Moreno, and Costa (2014) and Martin et al. (2013) proposed that these effects might not actually say something about real prediction processes, but might reflect integration processes rather than actual anticipation processes.

With the GECO corpus it would be possible to compute the forward transitional predictability for every word (McDonald & Shillcock, 2003) to investigate bilingual predictability effects. Using eye-tracking measures on all words in the sentences would serve to disentangle integration processes from actual predictive processes and further the understanding of bilingual anticipation in L2. Rayner, Ashby, Pollatsek, and Reichle (2004) showed that the E-Z reader model (Reichle et al., 1998) could be used as a framework to investigate prediction in L1 reading. According to this model the familiarity check and the lexical access are both influenced by target word predictability. A word is skipped when the familiarity check of word n+1, the word in the parafovea, is completed early enough. This means that the predictability of a word determines among other things the skipping probability of that word. A combination of the huge amount of eye movement data of the GECO and a useful theoretical framework like the E-Z reader model could lead to a better understanding of prediction processes in bilingual reading.

Syntactic Parsing and Processing Complexity

Researchers have asked the question whether L2 syntactic parsing is qualitatively similar to L1 parsing. A way to investigate this question is to look at ambiguity resolution in bilingual reading. Some studies provided evidence that bilinguals do not use structure based principles in L2 parsing like they do in L1 parsing (Felser, Roberts, Marinis, & Gross, 2003; Papadopoulou & Clahsen, 2003), other studies do show an influence of L2 syntactic structures on the resolution of ambiguities in L1 reading (Dussias & Sagarra, 2007), showing that these L2 structures are represented in the same way as the L1 structures. Dussias (2010) concludes that when eye tracking is used (e.g., Dussias & Cramer Scaltz, 2008; Dussias & Sagarra, 2007; Felser, Sato, & Bertenshaw, 2009; Roberts, Gullberg, & Indefrey, 2008), the evidence points in favor of structure based parsing in L2. Because we have a large database of eye movements in L1 and L2 reading, the preconditions for this structure based parsing could be investigated in detail.

In the process of syntactic parsing of a sentence, ambiguous structures are not the only challenge. In some instances, the parsing process is more difficult than in others. Demberg and Keller (2008) investigated why some sentences have a higher processing complexity than others using data from the Dundee corpus (Kennedy & Pynte, 2005). They concluded that surprisal (Hale, 2001), defined by probabilistic grammar, and dependency locality theory (DLT; Gibson, 1998), focusing more on the distance based integration cost, must both be taken into account to explain processing complexity. An interesting question is whether for L2 processing the probabilistic L1 grammar has an influence on the processing complexity in L2 or whether the integration cost has more weight in the processing complexity in L2, since probabilistic dependencies are not yet established fully for L2. The GECO would be able to test these hypotheses and this might be the first step towards generalizing Demberg and Keller's theory to L2 sentence processing.

CONCLUSION

The research in this dissertation was motivated by an attempt to understand how a growing population of multi-linguals tackles the immensely complex process of reading text. Like in most research efforts, we had to (reluctantly) limit the scope of our inquiry to a more restricted and well-delineated question: how do late bilinguals attain lexical access? We are pleased to say that this dissertation contributes to the bilingual research field in three major ways. One, The Ghent Eye tracking Corpus (GECO), which was constructed in function of this dissertation, is freely available to other researchers, making hypothesis testing and model evaluation concerning bilingual (and monolingual) reading more cost-efficient. We hope this database of eye movements can help make research into naturalistic reading more pervasive in language research and might spur other researchers to gather corpora of eye movements from the same novel in other languages, advancing cross-lingual comparative research. Two, in the first two empirical chapters we have shown that language exposure, as measured by vocabulary size or proficiency, is a central concept in understanding reading differences between L1 and L2 bilingual reading, as well as between L1 bilingual and monolingual reading. Future research might be able to delineate the influence of this factor from other factors influencing these differences. Third, by analyzing eye movements towards the nouns of the novel, we provided strong evidence for language independent lexical access. This is the first time cognate facilitation has been detected in L1 reading of a naturalistic text, which includes high semantic constraint sentences. Also, it is the first time that effects of cross-lingual neighborhood size and frequency are found in either L1 or L2 text reading. These findings support the assumptions of the BIA+ model (Dijkstra & van Heuven, 2002), namely that language does not serve as a cue that can restrict lexical access to the target language.

REFERENCES

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition*, *24*(4), 477–492.

Angeles Alonso, M., Fernandez, A., & Diez, E. (2015). Subjective age-of-acquisition norms for 7,039 Spanish words. *Behavior Research Methods*, *47*(1), 268–274. http://doi.org/10.3758/s13428-014-0454-2

Barry, C., Hirsh, K. W., Johnston, R. A., & Williams, C. L. (2001). Age of Acquisition, Word Frequency, and the Locus of Repetition Priming of Picture Naming. *Journal of Memory and Language*, *44*(3), 350–375. http://doi.org/10.1006/jmla.2000.2743

Beauvillain, C., & Grainger, J. (1987). Accessing Interlexical Homographs: Some Limitations of a Language-Selective Access. *Journal of Memory and Language*, *26*, 658–672.

Blythe, H. I., Häikiö, T., Bertam, R., Liversedge, S. P., & Hyönä, J. (2011). Reading disappearing text: Why do children refixate words? *Vision Research*, *51*(1), 84–92. http://doi.org/10.1016/j.visres.2010.10.003

Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2009). Visual information capture during fixations in reading for children and adults. *Vision Research*, *49*(12), 1583–1591. http://doi.org/10.1016/j.visres.2009.03.015

Brown, G. D. A., & Watson, F. L. (1987). First in, first out: Word learning age and spoken word frequency as predictors of word familiarity and word naming latency. *Memory & Cognition*, *15*(3), 208–216. http://doi.org/10.3758/BF03197718

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: a review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology*, *58*(5), 412–424. http://doi.org/10.1027/1618-3169/a000123

Brysbaert, M., Lange, M., & Wijnendaele, I. V. (2000). The effects of age-of-acquisition and frequency-of-occurrence in visual word recognition: Further evidence from the Dutch language. *European Journal of Cognitive Psychology*, *12*(1), 65–85.

Brysbaert, M., Stevens, M., De Deyne, S., Voorspoels, W., & Storms, G. (2014). Norms of age of acquisition and concreteness for 30,000 Dutch words. *Acta Psychologica*, *150*, 80–84. http://doi.org/10.1016/j.actpsy.2014.04.010

Brysbaert, M., Van Wijnendaele, I., & De Deyne, S. (2000). Age-of-acquisition effects in semantic processing tasks. *Acta Psychologica*, *104*(2), 215–226. http://doi.org/10.1016/S0001-6918(00)00021-4

Buswell, G. T. (1922). *Fundamental reading habits: a study of their development*. University of California Libraries.

Butler, B., & Hains, S. (1979). Individual differences in word recognition latency. *Memory & Cognition*, *7*(2), 68–76. http://doi.org/10.3758/BF03197587

Caine, R. N., & Caine, G. (1991). Making Connections: Teaching and the Human Brain. Retrieved from http://eric.ed.gov/?id=ED335141

Canseco-Gonzalez, E., Brehm, L., Brick, C. A., Brown-Schmidt, S., Fischer, K., & Wagner, K. (2010). Carpet or Cárcel: The effect of age of acquisition and language mode on bilingual lexical access. *Language and Cognitive Processes*, *25*(5), 669–705. http://doi.org/10.1080/01690960903474912

Carroll, J. B., & White, M. N. (1973). Word frequency and age of acquisition as determiners of picture-naming latency. *Quarterly Journal of Experimental Psychology*, *25*(1), 85–95. http://doi.org/10.1080/14640747308400325

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and Performance VI* (pp. 535–555). London: Academic Press.

Cowan, N. (1998). *The Development of Memory in Childhood*. Psychology Press.

Cowan, N., Morey, C. C., AuBuchon, A. M., Zwilling, C. E., & Gilchrist, A. L. (2010). Seven-year-olds Allocate Attention Like Adults Unless Working Memory is Overloaded. *Developmental Science*, *13*(1), 120. http://doi.org/10.1111/j.1467-7687.2009.00864.x

Darcy, N. T. (1953). A review of the literature on the effects of bilingualism upon the measurement of intelligence. *Journal of Genetic Psychology*, *82*, 21–57.

De Groot, A. M., Delmaar, P., & Lupker, S. J. (2000). The processing of interlexical homographs in translation recognition and lexical decision: support for non-selective access to bilingual memory. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *53*(2), 397–428. http://doi.org/10.1080/713755891

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*(8), 1117–1121. http://doi.org/10.1038/nn1504

Demberg, V., & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, *109*(2), 193–210. http://doi.org/10.1016/j.cognition.2008.07.008

Diependaele, K., Lemhöfer, K., & Brysbaert, M. (2013). The word frequency effect in first-and second-language word recognition: a lexical entrenchment account. *The Quarterly Journal of Experimental Psychology*, *66*(5), 843–863.

Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of Cognates and Interlingual Homographs: The Neglected Role of Phonology. *Journal of Memory and Language*, *41*(4), 496–518.

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language*, *62*(3), 284–301. http://doi.org/10.1016/j.jml.2009.12.003

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(03). http://doi.org/10.1017/S1366728902003012

Dijkstra, T., Van Jaarsveld, H., & ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition*, *1*(1), 51–66. http://doi.org/10.1017/S1366728998000121

Drieghe, D., Brysbaert, M., & Desmet, T. (2005). Parafoveal-on-foveal effects on eye movements in text reading: does an extra space make a difference? *Vision Research*, *45*(13), 1693–1706. http://doi.org/10.1016/j.visres.2005.01.010

Dussias, P. E. (2010). Uses of Eye-Tracking Data in Second Language Sentence Processing Research. *Annual Review of Applied*

Linguistics, *30*, 149–166.
http://doi.org/10.1017/S026719051000005X

Dussias, P. E., & Cramer Scaltz, T. R. (2008). Spanish–English L2 speakers' use of subcategorization bias information in the resolution of temporary ambiguity during second language reading. *Acta Psychologica*, *128*(3), 501–513.
http://doi.org/10.1016/j.actpsy.2007.09.004

Dussias, P. E., & Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish-English bilinguals. *Bilingualism: Language and Cognition*, *10*(01), 101–116.
http://doi.org/10.1017/S1366728906002847

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 663–679.
http://doi.org/10.1037/0278-7393.33.4.663

Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: insights from connectionist networks. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *26*(5), 1103–1123.

Epstein, I. (1905). *La pensee et la poligloise.* Lausanne: Libraire Payot.

Felser, C., Roberts, L., Marinis, T., & Gross, R. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics*, *24*(03), 453–489.

Felser, C., Sato, M., & Bertenshaw, N. (2009). The on-line application of binding Principle A in English as a second language. *Bilingualism: Language and Cognition*, *12*(04), 485–502.
http://doi.org/10.1017/S1366728909990228

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(5), 1461–1469. http://doi.org/10.1037/a0036756

Frenck-Mestre, C., German, E. S., & Foucart, A. (2014). Qualitative Differences In Native And Nonnative Semantic Processing as Revealed by ERPs. In R. R. Heredia & J. Altarriba (Eds.), *Foundations of Bilingual Memory* (pp. 237–255). New York, NY:

Springer New York. Retrieved from http://link.springer.com/10.1007/978-1-4614-9218-4_12

Frisson, S., Rayner, K., & Pickering, M. J. (2005). Effects of Contextual Predictability and Transitional Probability on Eye Movements During Reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(5), 862–877. http://doi.org/10.1037/0278-7393.31.5.862

Gerard, L. D., & Scarborough, D. L. (1989). Language-specific lexical access of homographs by bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(2), 305–315. http://doi.org/10.1037/0278-7393.15.2.305

Gerhand, S., & Barry, C. (1998). Word frequency effects in oral reading are not merely age-of-acquisition effects in disguise. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(2), 267–283. http://doi.org/10.1037/0278-7393.24.2.267

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*(1), 1–76.

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), 246–269. http://doi.org/10.1037/0278-7393.30.1.246

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*(3), 787–814. http://doi.org/10.1016/j.jml.2007.07.001

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General*, *140*(2), 186–209. http://doi.org/10.1037/a0022256

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review*, *115*(3), 577–600. http://doi.org/10.1037/a0012667

Grabo, R. P. (1931). *A study of comparative vocabularies of junior high school pupils from English and Italian speaking homes* (No. Bulletin No. 13). Washington, D.C.: U.S. Office of Education.

Grainger, J., & Jacobs, A. M. (1996). Orthographic Processing in Visual Word Recognition: A Multiple Read-out Model. *Psychological Review*, *103*(3), 518–565. http://doi.org/10.1037/0033-295X.103.3.518

Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of the second meeting of the North American Chapter of the Association for Computational Linguistics on Language technologies* (pp. 1–8). Association for Computational Linguistics. Retrieved from http://dl.acm.org/citation.cfm?id=1073357

Harris, C. W. (1948). An exploration of language skill patterns. *Journal of Educational Psychology*, *32*, 351–364.

Izura, C., & Ellis, A. W. (2002). Age of acquisition effects in word recognition and production in first and second languages. *Psicológica: Revista de Metodología Y Psicología Experimental*, *23*(2), 245–282.

Izura, C., & Ellis, A. W. (2004). Age of acquisition effects in translation judgement tasks. *Journal of Memory and Language*, *50*(2), 165–181. http://doi.org/10.1016/j.jml.2003.09.004

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research*, *45*(2), 153–168. http://doi.org/10.1016/j.visres.2004.07.037

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, *42*(3), 643–650. http://doi.org/10.3758/BRM.42.3.643

Kittredge, A. K., Dell, G. S., Verkuilen, J., & Schwartz, M. F. (2008). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cognitive Neuropsychology*, *25*(4), 463–492. http://doi.org/10.1080/02643290701674851

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978–990. http://doi.org/10.3758/s13428-012-0210-4

Kuperman, V., & Van Dyke, J. A. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language*, *65*(1), 42–73. http://doi.org/10.1016/j.jml.2011.03.002

Lehtonen, M., Hultén, A., Rodríguez-Fornells, A., Cunillera, T., Tuomainen, J., & Laine, M. (2012). Differences in word recognition between early bilinguals and monolinguals: behavioral and ERP evidence. *Neuropsychologia*, *50*(7), 1362–1371. http://doi.org/10.1016/j.neuropsychologia.2012.02.021

Lemhöfer, K., & Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: Effects of code similarity in language-specific and generalized lexical decision. *Memory & Cognition*, *32*(4), 533–550.

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(2), 381–390. http://doi.org/10.1037/a0014875

Mcclelland, J., & Rumelhart, D. (1981). An Interactive Activation Model of Context Effects in Letter Perception .1. an Account of Basic Findings. *Psychological Review*, *88*(5), 375–407. http://doi.org/10.1037/0033-295X.88.5.375

McDonald, S. A., & Shillcock, R. C. (2003). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science*, *14*(6), 648–652.

Monsell, S. (1991). The Nature and Locus of Word Frequency Effects in Reading. In D. Besner & G. W. Humphreys (Eds.), *Basic Processes in Reading: Visual Word Recognition* (pp. 148–197). Hillsdale, NJ, England: Erlbaum.

Montrul, S., & Foote, R. (2014). Age of acquisition interactions in bilingual lexical access: A study of the weaker language of L2 learners and heritage speakers. *International Journal of Bilingualism*, *18*(3), 274–303. http://doi.org/10.1177/1367006912443431

Morrison, C. M., Chappell, T. D., & Ellis, A. W. (1997). Age of Acquisition Norms for a Large Set of Object Names and Their Relation to Adult Estimates and Other Variables. *The Quarterly*

Journal of Experimental Psychology Section A, *50*(3), 528–559. http://doi.org/10.1080/027249897392017

Morrison, C. M., & Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 116–133. http://doi.org/10.1037/0278-7393.21.1.116

Morton, J. (1970). A Functional Model for Memory. In D. A. Norman (Ed.), *Models of Human Memory* (pp. 203–254). Elsevier.

Otten, M., & Van Berkum, J. J. A. (2008). Discourse-Based Word Anticipation During Language Processing: Prediction or Priming? *Discourse Processes*, *45*(6), 464–496. http://doi.org/10.1080/01638530802356463

Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing. *Studies in Second Language Acquisition*, *25*(04), 501–528.

Perea, M., & Pollatsek, A. (1998). The effects of neighborhood frequency in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 767.

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, *11*(3), 105–110. http://doi.org/10.1016/j.tics.2006.12.002

Pollatsek, A., Perea, M., & Binder, K. S. (1999). The Effects of "Neighborhood Size" in Reading and Lexical Decision. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(4), 1142–1158.

Ransdell, S. E., & Fischler, I. (1987). Memory in a monolingual mode: When are bilinguals at a disadvantage? *Journal of Memory and Language*, *26*(4), 392–405.

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology*, *41*(2), 211–236.

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The Effects of Frequency and Predictability on Eye Fixations in Reading: Implications for the E-Z Reader Model. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(4), 720–732. http://doi.org/10.1037/0096-1523.30.4.720

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, *3*(4), 504–509. http://doi.org/10.3758/BF03214555

Reichle, E. D. (2011). Serial Attention Models of Reading. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.), *Oxford Handbook On Eye Movements* (pp. 767–786). Oxford, England: Oxford University Press.

Reichle, E. D., Liversedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2013). Using E-Z Reader to examine the concurrent development of eye-movement control and reading skill. *Developmental Review*, *33*(2), 110–149. http://doi.org/10.1016/j.dr.2013.03.001

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*(1), 125.

Roberts, L., Gullberg, M., & Indefrey, P. (2008). Online pronoun resolution in L2 discourse: L1 influence and general learner effects. *Studies in Second Language Acquisition*, *30*(03), 333–357.

Silverberg, S., & Samuel, A. G. (2004). The effect of age of second language acquisition on the representation and processing of second language words. *Journal of Memory and Language*, *51*(3), 381–398. http://doi.org/10.1016/j.jml.2004.05.003

Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *Journal of Memory and Language*, *64*(1), 88–107. http://doi.org/10.1016/j.jml.2010.08.006

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. *Psychological Science*, *20*(8), 923–927.

Van Hell, J. G., & de Groot, A. M. B. (2008). Sentence context modulates visual word recognition and translation in bilinguals. *Acta Psychologica*, *128*(3), 431–451. http://doi.org/10.1016/j.actpsy.2008.03.010

Van Heuven, W. J. B., & Dijkstra, T. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language*, *39*, 458–483.

Van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190.

Wang, Y., Liu, D., & Wang, Y. (2003). Discovering the Capacity of Human Memory. *Brain and Mind*, *4*(2), 189–198. http://doi.org/10.1023/A:1025405628479

Whitford, V., & Titone, D. (2012). Second-language experience modulates first- and second-language word frequency effects: Evidence from eye movement measures of natural paragraph reading. *Psychonomic Bulletin & Review*, *19*(1), 73–80. http://doi.org/10.3758/s13423-011-0179-5

Yamazaki, M., Ellis, A. W., Morrison, C. M., & Ralph, M. A. L. (1997). Two age of acquisition effects in the reading of Japanese Kanji. *British Journal of Psychology*, *88*(3), 407–421. http://doi.org/10.1111/j.2044-8295.1997.tb02648.x

Zevin, J. D., & Seidenberg, M. S. (2002). Age of Acquisition Effects in Word Reading and Other Tasks. Retrieved from http://citeseer.uark.edu:8080/citeseerx/viewdoc/summary?doi=10.1.1.13.9805

NEDERLANDSE SAMENVATTING

INLEIDING

Meertaligheid is allesbehalve zeldzaam. Het is de norm in de meeste moderne samenlevingen: meer dan de helft van de wereldpopulatie heeft kennis van meer dan 1 taal (Grosjean, 2010). Vaak wordt het voorkomen van twee-of meertaligheid onderschat. Een van de redenen voor deze onderschatting, is dat mensen vaak denken dat om als meertalige te worden beschouwd, ze een gelijke mate van vaardigheid in beide talen moeten hebben. Dit is echter niet hoe tweetaligheid meestal wordt gedefinieerd in een politieke of wetenschappelijke context (voor voorbeelden zie Eurobarometer, 2006; Grosjean, 1992). Een andere reden is dat eentaligheid vaak word verondersteld de norm te zijn in linguïstische en taalontwikkelingstheorieën (Ellis, 2006), terwijl dit niet het geval is.

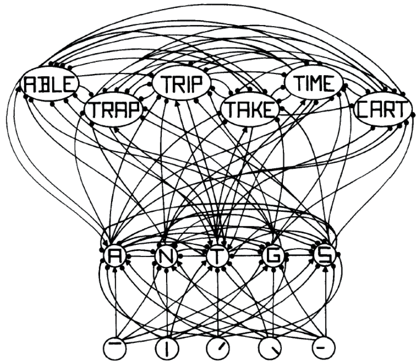
Het is niet altijd duidelijk wat er exact bedoeld wordt als men in de wetenschap over een tweetalige persoon spreekt (Hoffmann, 2014). De meest populaire definitie van tweetaligheid is die van Grosjean (1982), waarbij enige kennis en dagelijks gebruik van een andere taal voldoende is om als tweetalige te kwalificeren. Aangezien onderzoek heeft aangetoond dat onder andere de taalvaardigheid en de leeftijd van het aanleren van een tweede taal een groot effect hebben op taalgebruik van tweetaligen, is het erg belangrijk om zowel objectieve maten van taalvaardigheid als zoveel mogelijk informatie over de taalachtergrond van de participanten te verzamelen.

De evolutie van een eentalige naar een meertalige samenleving creëert een waaier aan onderzoeksmogelijkheden, die inherent gerelateerd zijn aan het cognitieve, linguïstisch en niet-linguïstisch, functioneren van een meertalige persoon. In dit proefschrift ligt de focus op een taak die erg centraal staat in onze cultuur, namelijk lezen. Welke cognitieve processen spelen een rol bij tweetalig lezen? Hoe gaat een tweetalig persoon van letters op een pagina naar een woord met vorm en betekenis? Verschilt het proces van lezen in een tweede taal van lezen in een eerste taal?

Het is duidelijk dat mensen worden geboren met taalontwikkelingspotentieel (bv., Lucy, 1996, Haun, 2007): baby's leren al erg vroeg brabbelen zonder expliciete instructies. Ondanks de centraliteit van geschreven taal in onze samenleving moet geletterdheid expliciet aangeleerd worden (Pinker, 1997). Dit heeft als oorzaak dat schrijven en lezen evolutionair gezien relatief nieuwe menselijke vaardigheden zijn (Deacon, 1998): Het zijn culturele uitvindingen van slechts een paar duizend jaar oud. Dit betekent dat het menselijke brein nieuwe connecties moet aanmaken tussen verschillende hersenregio's, die instaan voor meer basische cognitieve processen, om te leren lezen (Dehaene, Le Clec'H, Poline, Le Bihan, & Cohen, 2002). Het is dus de plasticiteit van de menselijke hersenen (Neville & Bavelier, 2000) en de leercapaciteit van de mens die er voor zorgt dat een kind een vaardige lezer wordt.

VISUELE WOORDHERKENNING

Lezers moeten een enorme cognitieve inspanning leveren: aandachts- en geheugenprocessen, visuele en linguïstische processen moeten worden aangewend om te slagen in deze taak. Door deze complexiteit leggen de meeste modellen van lezen de nadruk op slechts een klein deel van het leesproces. Dit proefschrift behandeld eveneens slechts een deel van het leesproces. We stellen de vraag hoe een tweetalige lezer letters omzet in woorden en toegang krijgt tot de informatie die bij dit woord hoort. Dit proces wordt vaak beschreven als visuele woordherkenning of lexicale toegang. Het belangrijkste eentalige model van visuele woordherkenning is het interactieve activatie (IA) model van McClelland en Rumelhart (1981) (zie Figuur 1). In dit model zijn er 3 niveaus van representaties. Het laagste is het visuele kenmerk niveau, op dit niveau worden de lijnen op het papier gecodeerd. Het volgende niveau zijn de letter representaties. Het laatste is het woord niveau. Belangrijk is het bestaan van laterale inhibitie tussen de representaties binnen niveaus. Dit zorgt voor mogelijke competitie tussen letters en woorden tijdens lexicale toegang.



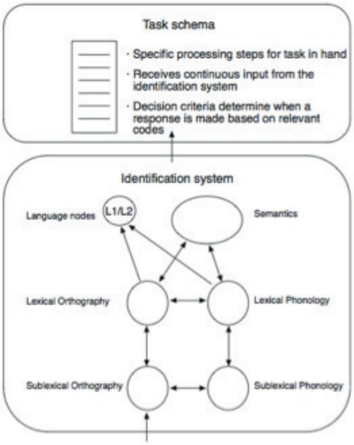
Figuur 1. Een voorbeeld van een geactiveerd network gedurende de herkenning van de letter 'T' volgens de architectuur van het IA model. Pijltjes zijn excitatorische connecties, bolletjes zijn inhibitorische connecties. Figuur overgenomen uit McClelland and Rumelhart (1981), pagina 830.

ONDERZOEK NAAR CROSS-LINGUALE INTERACTIES EN HET BIA+MODEL

Tijdens het proces van visuele woordherkenning moet een lezer de correcte representatie selecteren. De meerderheid van het tweetalige onderzoek heeft de vraag gesteld of deze lexicale toegang voor tweetaligen taalafhankelijk of onafhankelijk is. Meer bepaald, kunnen woorden van een andere taal dan de taal die verwerkt wordt, geactiveerd worden tijdens het proces van lexicale toegang. Om deze hypothese te onderzoeken worden vaak woorden gebruikt die overlap vertonen met de andere taal. Veelgebruikte woorden in dit soort onderzoek zijn cognaten, cross-linguale homografen en buurwoorden. Cognaten zijn woorden die een (bijna) gelijke vorm en betekenis hebben over talen heen. Een voorbeeld is het Engelse woord *apple,* dat *appel* betekent. Cross-linguale homografen zijn woorden die een gelijke vorm, maar een andere betekenis over talen heen hebben. Een voorbeeld is het woord *glad* dat *glibberig* betekent, maar *tevreden* in het Engels. Meerdere experimenten tonen aan dat cognaten sneller worden verwerkt dan andere woorden (bv., Caramazza & Brones, 1979; Duyck, Van Assche, Drieghe, & Hartsuiker, 2007; Gollan, Forster, & Frost, 1997; Lemhöfer, Dijkstra, & Michel, 2004; Van Hell & Dijkstra, 2002). Homografen worden soms sneller, soms trager, verwerkt (bv., de Groot, Delmaar, & Lupker, 2000; Dijkstra, Grainger, & van Heuven, 1999).

Buurwoorden van een bepaald woord worden gedefinieerd als alle woorden die erg gelijkend zijn aan dat woord. Zo heeft het woord *hond* 11 buurwoorden in het Nederlands, waaronder *hand* en *mond* en 14 buurwoorden in het Engels, waaronder *hold* en *hind*. Tot nu toe heeft 1 studie aangetoond dat het aantal buren in de andere taal een effect heeft op de verwerkingssnelheid van woorden (van Heuven & Dijkstra, 1998). Deze resultaten tonen allemaal aan dat lexicale toegang in de ene taal effecten vertoont van lexicale representaties in de andere taal.

Geïnspireerd door deze bevindingen werd het IA model succesvol aangepast aan tweetalig lezen door (Dijkstra & van Heuven, 2002) naar het Bilingual Interactive Activation (BIA+) model (zie figuur 2).



Figuur 2. De architectuur van het BIA+ model. Figuur overgenomen uit Dijkstra en van Heuven (2002), pagina 183

Het belangrijkste kenmerk van het BIA+ model (Dijkstra & van Heuven, 2002) is een parallelle, taalonafhankelijke lexicale toegang tot een geïntegreerd tweetalige lexicon. Dit lexicon bevat alle gekende woorden van alle gekende talen van een tweetalige persoon. Dit model heeft *language nodes*. Deze geven de taal aan waartoe een lexicale representatie behoort, maar kunnen geen top-down activatie naar deze representaties sturen. Dit betekent dat de lexicale toegang niet kan worden beperkt tot 1 taal via deze *language nodes*. Tot nu toe is dit het enige geïmplementeerde model van tweetalig lezen.

TWEETALIG ONDERZOEK IN EEN ZINSCONTEXT

Het overgrote deel van tweetalig onderzoek onderzocht lexicale toegang tot woorden die in isolatie op een scherm werden gepresenteerd. Proefpersonen moesten dan ofwel het woord benoemen (benoemtaak) , beslissing of een bepaalde lettercombinatie een bestaand woord vormde of niet (lexicale decisie taak), of zo snel mogelijk het woord herkennen. Voor deze taken is altijd een respons van de proefpersonen nodig. Dit betekent dat er andere cognitieve processen aan de grondslag kunnen liggen van de effecten die in dit soort onderzoek worden gevonden (Grainger & Jacobs, 1996). Ook is het aannemelijk dat een zincontext een grote invloed op woordherkenning heeft. Aangezien mensen bijna altijd woorden lezen die ingebed zijn in een betekenisvolle zin, moet er onderzoek komen die kijkt naar hoe lexicale toegang gebeurt in een meer natuurlijke context.

OOGBEWEGINGEN TIJDENS LEZEN

Een ander belangrijk onderdeel van lezen zijn de oogbewegingen die gemaakt worden. Tijdens lezen lijkt het of de ogen vlot over de tekst glijden, maar dit is een illusie. Eigenlijk maakt het oog snelle bewegingen, dit zijn saccades, en blijft het voor korte periodes stil staan, dit zijn fixaties. De typische saccade is 8 letters lang en de typische fixatie duurt 200-250 ms (Rayner, 2009).

Een groot aantal studies relateert deze bewegingen van het oog aan het cognitieve functioneren van de mens (bv., Just & Carpenter, 1980; Reichle, Pollatsek, Fisher, & Rayner, 1998). Dit betekent dat de oogbewegingen tijdens het lezen van tekst kunnen gebruikt worden om te onderzoeken welke cognitieve processen gebruikt worden tijdens het lezen. Het registreren van oogbewegingen wordt *eye-tracking* genoemd. Met deze methode wordt de positie van het oog elke milliseconde (of dubbel zo vaak) geregistreerd met een hoge spatiale nauwkeurigheid. Op deze manier kan een erg rijke dataset bekomen worden. Dit is dus een ideale manier om lexicale toegang te onderzoeken in een natuurlijke leescontext zonder tussenkomst van andere cognitieve processen. Meerdere onderzoekers hebben inderdaad tweetalige lexicale toegang bestudeert met behulp van eye-tracking (e.g., Altarriba, Kroll, Sholl, & Rayner, 1996; Bultena, Dijkstra, & van Hell, 2014; Dussias & Cramer Scaltz, 2008; Duyck et al., 2007). De meeste van deze studies kijken naar oogbewegingen gericht op specifieke woorden in geïsoleerde zinnen die een lage semantische restrictie opleggen aan die woorden. Wij zijn geïnteresseerd in tweetalige lexicale toegang tijdens een meer uitgebreide narratieve context die ook zinnen bevat die wel hoge restricties opleggen aan woorden.

De veelheid aan informatie die eye-tracking biedt, kan nog worden vergroot door erg veel proefpersonen te testen of erg veel tekst aan te bieden. Deze zogenaamde *mega* studies of *corpus* studies (bv., Kennedy & Pynte, 2005; Kliegl, Nuthmann, & Engbert, 2006) zijn in het verleden al succesvol aangewend om specifieke hypotheses te toetsen en taalmodellen te ontwikkelen. Een dergelijk eye-tracking corpus ontbreekt volledig voor tweetalig lezen.

Het belangrijkste eentalig model van oogbewegingen is het E-Z reader model (Pollatsek, Reichle, & Rayner, 2006; Rayner, Ashby, Pollatsek, & Reichle, 2004; Reichle et al., 1998; Reichle, Rayner, & Pollatsek, 1999; Reichle, Warren, & McConnell, 2009). Dit model is een serieel model. Dit betekent dat de aandacht slechts op 1 woord tegelijk kan worden gericht. In dit model worden aandacht en oogbewegingen losgekoppeld van elkaar, deze processen gebeuren dus niet noodzakelijk synchroon. Als er wordt gefixeerd op een bepaald woord begint de eerste fase van lexicale toegang, de *familiarity check.* Als deze compleet is wordt het programmeren van een oogbeweging naar het volgende woord geïnitieerd. Het beëindigen van het tweede proces, de *completion of lexical access*, zorgt voor een beweging van de aandacht naar het volgende woord. Woordfrequentie, woordlengte en voorspelbaarheid zijn belangrijke factoren die de moeilijkheid van de beide processen bepalen. Dit model is gebruikt om leespatronen van oudere

(Rayner, Reichle, Stroud, Williams, & Pollatsek, 2006) en jongere (Rayner et al., 2006; Reichle et al., 2013) lezers te simuleren, maar is nog nooit aangewend om verschillen in oogbewegingen voor eentalige en tweetalige personen te modelleren.

HET ONDERZOEK

Dit proefschrift heeft twee complementaire doelen. Het eerste is van methodologische aard en wordt behandeld in HOOFDSTUK 2. Het betreft de ontwikkeling van een tweetalig eye-tracking corpus. Het tweede is empirisch en wordt in HOOFDSTUKKEN 3 tot en met 6 behandeld. In deze hoofdstukken trachten we om dit eye-tracking corpus te gebruiken om tweetalige lexicale toegang in de eerste taal (L1) en de tweede taal (L2) tijdens zinsverwerking in een natuurlijke leescontext te onderzoeken. Dit zal op een zinsniveau (HOOFDSTUK 3) en een woordniveau (HOOFDSTUK 4-6) worden gedaan.

HOOFDSTUK 2: METHODOLOGISCH KADER

We ontwikkelden een tweetalig eye-tracking corpus: de Ghent Eye tracking Corpus (GECO). Dit is een grote dataset van leesgedrag van Nederlands-Engelse tweetaligen en Engelse eentaligen, die een volledige roman lazen. De roman was ‘De Zaak Styles’ van Agatha Christie. Deze roman bevat meer dan 5 000 zinnen en meer dan 50 000 woorden. Het is duidelijk dat een roman een erg narratieve context is en dat deze een veelvoud van zinnen inhoudt die semantische restricties opleggen aan de woorden.

Alle proefpersonen voerden een brede batterij van taalvaardigheidstesten uit, onder andere een spelling test (Gletschr voor Nederlands, De Pessemier & Andries, 2009; WRAT4 voor Engels, Wilkinson & Robertson, 2006), een vocabularium test (LexTALE, Lemhöfer & Broersma, 2012) en een lexicale decisietaak. Voor de tweetaligen gebeurde dit in de eerste (L1) en tweede taal (L2).

Wij maken dit corpus beschikbaar voor andere onderzoekers om vrij te gebruiken. [20]

Het ontwikkelen van een grote data set, met erg veel proefpersonen en/of stimulusmateriaal, heeft meerdere voordelen. Zo heeft zo'n dataset een grotere statistische power, die ervoor zorgt dat erg kleine effecten kunnen worden gedetecteerd. Ook wordt stimulusmateriaal vaak minder streng geselecteerd, waardoor er een groter bereik is van bepaalde stimuluskenmerken zodat er een meer representatieve steekproef van gedrag wordt bekomen. Ook kunnen onderzoekers experimentele hypothesen toetsen zonder hiervoor opnieuw data te verzamelen, wat natuurlijk erg tijd efficiënt is.

Zoals hierboven aangehaald ontbreekt een eye-tracking corpus voor tweetalig lezen. De eentalige corpora (bv., Frank, Fernandez Monsalve, Thompson, & Vigliocco, 2013; Kennedy & Pynte, 2005; Kliegl & Engbert, 2005) zijn veelvuldig gebruikt in het ontwikkelen van taalmodellen (Demberg & Keller, 2008; Engbert, Nuthmann, Richter, & Kliegl, 2005), theoretisering omtrent methode-effecten (bv., Kuperman, Drieghe, Keuleers, & Brysbaert, 2013) en specifieke hypothesetoetsing (bv., Kennedy & Pynte, 2008)

Zoals voorgaande corpora in het eentalig vakgebied, kan dit corpus eveneens worden gebruikt op een exploratieve wijze, waardoor interessante onderzoeksvragen en inzichten in tweetalige lezen kunnen worden ontdekt. Vanwege de grote hoeveelheid tekstmateriaal in het boek, kan dit corpus ook gebruikt worden voor toetsing van specifieke hypotheses. Ook hopen we dat deze dataset de generaliseerbaarheid van de leesmodellen kan valideren en dat deze data kan worden gebruikt voor de verdere ontwikkeling van taaltheorieën.

[20] Voor data op woordniveau: http://expsy.ugent.be/downloads/uschi/
Voor data op zinsniveau: http://expsy.ugent.be/downloads/uschi2/
Login: uschi
Password:pp02

Een goed voorbeeld van verder gebruik van de GECO is het onderzoeken van de organisatie van het tweetalige lexicon. Zo kunnen cross-linguale homograaf effecten worden onderzocht om een beter beeld te krijgen van de manier van representatie van woorden die overlap vertonen tussen talen. Ook kunnen effecten van de leeftijd waarop een woord geleerd werd (AoA) worden onderzocht om te kijken hoe een tweetalig vocabularium zich ontwikkelt. Een ander voorbeeld is het onderzoek naar zinscontext effecten. Het is tot nu toe helemaal nog niet duidelijk hoe de voorspelbaarheid van een bepaald woord de oogbewegingen van een persoon die leest in L2 zou beïnvloeden. Ook is het onbekend of lezers in een tweede taal ook gestructureerde principes gebruiken bij het verwerken van de syntactische structuur van een zin. De GECO kan aan al deze vragen en meer, op een significante manier bijdragen.

HOOFDSTUK 3: OOGBEWEGINSPATRONEN OP ZINSNIVEAU

Tweetaligen hebben meestal meer woorden en dus meer lexicale representaties in hun mentale lexicon dan eentaligen. Het is mogelijk dat deze groei van hun vocabularium een nadeel veroorzaakt voor tweetaligen tijdens het activeren en selecteren van een bepaald woord in vergelijking met eentaligen. In het veld van taalproductie zijn er een aantal studies die aantoonden dat meertaligen een productienadeel hebben ten opzicht van eentaligen: Tweetaligen zijn langzamer in het benoemen van prenten dan eentaligen (Gollan, Montoya, Fennema-Notestine, & Morris, 2005; Ivanova & Costa, 2008) en ze verkeren vaker in een *tip-of-the-tongue* state (Gollan & Acenas, 2004; Gollan & Silverberg, 2001). De *weaker links* account (Gollan & Acenas, 2004) beweert dat tweetaligen sowieso hun taalblootstelling of *exposure* verdelen over meerdere talen. Dit zou leiden tot zwakkere schakels, de zogenaamde *weaker links* in het taalsysteem van een tweetalige (Gollan, Montoya, Cera, & Sandoval, 2008). De auteurs van deze hypothese verwachten dat tweetaligen ook nadelen zullen ondervinden bij lexicale toegang tijdens lezen (Gollan et al., 2011). Ransdell en Fischler (1987) en Lehtonen et al. (2012) vonden inderdaad dat de reactietijden van tweetaligen langzamer waren dan die van eentaligen in een lexicale decisie taak in L1.

Om te onderzoeken of tweetaligen inderdaad een nadeel vertonen tijdens lexicale toegang tijdens het lezen van een boek voerden we twee belangrijke vergelijkingen van de oogbewegingen in de GECO op zinsniveau uit. De eerste was de vergelijking van het oogbewegingspatroon van ongebalanceerde Nederlands-Engels tweetaligen in hun L1 met het oogbewegingspatroon in hun L2. De lagere exposure aan L2 voor deze proefpersonen voorspelt tragere lexicale toegang in L2 (bv., Diependaele, Lemhöfer, & Brysbaert, 2013). We vonden inderdaad dat tweetaligen langzamer waren bij het lezen van L2 zinnen dan L1 zinnen. Ze maakten meer fixaties, die elk langer duurden, en sloegen minder woorden over bij het lezen in hun L2 dan bij het lezen in hun L1. De duur van de fixaties waren ook gevoeliger voor woordlengte in L2 dan in L1. Deze veranderingen van L2 lezen naar L1 lezen gelijken grotendeels op de veranderingen van kinderlijk naar volwassen lezen gerapporteerd in enkele eye-tracking studies (bv., Blythe, Liversedge, Joseph, White, & Rayner, 2009; Rayner, 1986). Wat niet gelijkend was, was de verandering in het aantal regressies die werden gemaakt. Regressies zijn oogbewingen die teruggaan in de tekst in plaats van vooruit. Het wordt meestal gevonden dat kinderen vaker regressies maken dan volwassenen (bv., Blythe, Häikiö, Bertam, Liversedge, & Hyönä, 2011; Buswell, 1922). Dit werd niet gevonden voor de tweetaligen die lazen in L2 in de GECO. Dit verschil kan worden verklaard door het feit dat kinderen een slechter werkgeheugen hebben (bv., Cowan, 1998; Cowan, Morey, AuBuchon, Zwilling, & Gilchrist, 2010) en/of een minder uitgebreid semantisch netwerk hebben dan volwassenen. Dit zou kunnen leiden tot meer tekstintegratieproblemen bij kinderen dan bij tweetalige volwassenen. Deze globale veranderingen voor L2 lezen zouden kunnen worden geïnterpreteerd als een vertraging van het proces van lexicale toegang. Elke verandering kan worden verklaard door een vertraging van de eerste fase van lexicale toegang, de *familiarity check* volgens het E-Z reader model (Reichle et al., 1998).

De tweede belangrijke vergelijking was deze tussen L1 lezen voor tweetaligen versus eentaligen. Volgens de weaker links theorie zouden de tweetaligen de frequentie van het taalgebruik tussen hun twee bekende talen moeten verdelen (bv., Gollan et al., 2011, 2008). Dit zou leiden tot zwakkere *links* tussen de lexicale en semantische representaties voor beide talen van de tweetalige persoon in vergelijking met de connecties in het taalsysteem van de eentalige. In contrast met de voorspellingen van deze theorie, vertoonden tweetaligen een grotendeels vergelijkbaar (en een even snel) oogbewegingspatroon als eentaligen. Dit betekent dat de zwakkere connecties tussen lexicale en semantische voorstellingen voor tweetaligen niet zo universeel zijn als gesuggereerd door sommige onderzoekers (Gollan et al., 2008).

Wij veronderstellen verder dat de validiteit van de *weaker links* hypothese afhankelijk is van welke tweetalige groepen men onderzoekt, omdat niet voor alle tweetalige groepen de L1 taalblootstelling noodzakelijk lager ligt dan voor eentalige personen. Ook kan het dat bij late tweetaligen (zoals de tweetaligen hier getest) een lagere taalblootstelling aan L1 na een bepaalde leeftijd niet meer zorgt voor een zwakkere connecties.

HOOFDSTUK 4: FREQUENTIE-EFFECTEN

Woordfrequentie is de meest invloedrijke lexicale variabele in woordherkenningsonderzoek en ze verklaart een groot deel van de variantie in reactietijden bij visuele woordherkenning (Brysbaert et al., 2011). Herhaalde blootstelling aan een bepaald woord verhoogt de basisactivatie van de relevante representatie, zodat lexicale selectie van het betreffende woord sneller wordt (bv., Monsell, 1991). De maximale snelheid van lexicale toegang is van nature beperkt, dus als een woord een zekere graad van blootstelling heeft bereikt, zal dit woord geen voordeel meer halen uit een extra blootstelling (Morton, 1970). Dit betekent dat een algemeen lager niveau van taalblootstelling zou leiden tot een groter verschil tussen de leestijden van laag- en hoogfrequente woorden. Inderdaad Gollan et al., (2011) voorspellen dat tweetalige frequentie-effecten groter zouden zijn dan eentalige frequentie-effecten als gevolg van een lagere taal blootstelling aan beide talen van de tweetalige persoon.

Woordfrequenties worden meestal gemeten door het tellen van het voorkomen van een bepaald woord in een grote hoeveelheid teksten (bv., Keuleers, Brysbaert, & New, 2010; van Heuven, Mandera, Keuleers, & Brysbaert, 2014) Het is aangetoond dat voor eentaligen met een klein vocabularium deze corpus woordfrequenties onnauwkeurig zijn, vooral voor laagfrequente woorden (Kuperman & Van Dyke, 2011). Als we dit idee veralgemenen naar tweetaligen, die een kleinere woordenschat in L2 dan in L1 hebben, kunnen we ook voor deze groep voorspellen dat deze corpus woordfrequenties niet erg nauwkeurig zijn.

We onderzochten voor het eerst de effecten van L1 en L2 vaardigheid op de grootte van de frequentie-effecten tijdens tekstlezen in L1 en L2. Hiervoor keken we naar het effect van woordfrequentie op de fixatietijden voor de inhoudswoorden die slechts eenmaal werden gefixeerd. Onze analyse van de GECO toonde aan dat onze groep van tweetaligen inderdaad een groter frequentie-effect hadden in L2 dan in L1. Omdat we een log-transformatie op de corpus woordfrequenties hadden toegepast voor de analyses, kan dit effect niet te wijten zijn aan een absoluut lagere blootstelling aan alle woorden, maar eerder aan een overschatting van de relatieve frequentie voor de corpus woordfrequenties van de laagfrequente woorden. Eentalige en tweetalige frequentie-effecten waren even groot tijdens L1 lezen. Dit impliceert een vergelijkbare grootte van het L1 vocabularium voor eentaligen en tweetaligen. Dit werd eveneens bevestigd door de resultaten van de LexTALE-test. De combinatie van de vergelijkbare grootte van de frequentie-effecten en de woordenschat voor tweetaligen en eentaligen in L1 laten zien dat de *weaker links* theorie (Gollan et al., 2008) niet geldt voor deze groep van tweetaligen, hetzij omdat de aanname van lagere blootstelling niet juist is of omdat deze lagere blootstelling niet noodzakelijkerwijs leidt tot zwakkere connecties in het tweetalig taalsysteem.

Voor tweetalig (in L1 en L2) en voor eentalig lezen was de grootte van het frequentie-effect beïnvloed door de grootte van het L1 vocabularium: een groter vocabularium maakte het frequentie-effect kleiner. Dit effect was even groot in de verschillende condities. Dit betekent dat de relatie tussen vocabularium grootte en lexicale toegang hetzelfde is voor beide groepen proefpersonen (eentaligen en tweetaligen) en over talen (L1 en L2). Deze resultaten tonen ook aan dat de lagere blootstelling aan laag frequente woorden voor mensen met een kleine woordenschat, zorgt voor tragere verwerking van deze laagfrequente woorden.

HOOFDSTUK 5 EN 6: CROSS-LINGUALE EFFECTEN

Zoals hierboven beschreven, veronderstelt het enige model van tweetalig lezen, het BIA+ model (Dijkstra & van Heuven, 2002), dat lexicale toegang taalonafhankelijk is. Dit betekent dat bijvoorbeeld tijdens lezen in de L1, de lexicale representaties van L2 eveneens geactiveerd worden. Er zijn veel studies die bewijs vonden voor dit idee (bv., Dijkstra, Timmermans, & Schriefers, 2000; Lemhöfer et al., 2008; van Heuven & Dijkstra, 1998). De meeste studies hebben cognaten gebruikt om deze kwestie te onderzoeken (bv., Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010; Duyck et al., 2007; Van Assche, Drieghe, Duyck, Welvaert, & Hartsuiker, 2011) en vinden duidelijke cognaatfacilitatie effecten in L2. Onderzoek dat kijkt naar lezen in een L1 context is echter zeldzaam (Titone, Libben, Mercier, Whitford, & Pivneva, 2011; Van Assche, Duyck, Hartsuiker, & Diependaele, 2009) en ontbreekt volledig in een L1 paragraaf context. GECO is dus zeer gepast om deze hypothese te onderzoeken, omdat de proefpersonen het boek compleet in een eentalige paragraaf context lazen. Ook vereist het stillezen weinig tussenkomst van de proefleider, dus biedt dit opzet een goede blik op taalverwerking hoe deze dagelijks gebeurd. We onderzochten de herkenning van identieke en niet-identieke cognaten (alleen de zelfstandige naamwoorden) in het GECO corpus in zowel vroege en late oogbewegingen in zowel L1 en L2 context.

In onze analyses van L2 lezen vonden we kortere eerste fixatietijden voor woorden met meer cross-linguale orthografische overlap met hun vertaling in L1. Ook hadden zelfstandige naamwoorden meer kans om overgeslagen te worden met stijgende cross-linguale orthografische overlap. Daarnaast werden identieke cognaten langer gefixeerd en werd er vaker teruggekeerd naar deze woorden dan naar andere woorden. In L1 lezen vonden we kortere eerste fixatietijden voor woorden met meer cross-linguale orthografische overlap, maar alleen wanneer deze woorden meer dan 9 letters lang waren. Ook was de totale leestijd van identieke cognaten korter, maar alleen wanneer ze hoogfrequent waren. Dit is de eerste demonstratie van cognaatfacilitatie in het lezen van doorlopende tekst in L1, waarbij de zinnen niet geconstrueerd waren zodat de cognaten weinig semantische restricties opgelegd kregen.

Een andere manier om cross-linguale activering van lexicale representaties te onderzoeken is te kijken naar de leessnelheid van woorden met een variabel aantal cross-linguale buurwoorden. Volgens de meest prominente modellen van visuele woordherkenning, zoals het IA model (McClelland & Rumelhart, 1981) en het Multiple Read Out model (MROM) (Grainger & Jacobs, 1996), moet de lezer bij het lezen de correcte lexicale representatie kiezen uit een set van orthografisch gelijkende representaties. Deze gelijkaardige woorden worden aangeduid als buren (Coltheart, Davelaar, Jonasson, & Besner, 1977). Het aantal buren van een doelwoord en de frequentie van deze buurwoorden hebben een invloed op de leestijd van het doelwoord (bv., Perea & Pollatsek, 1998; Pollatsek, Perea, & Binder, 1999). Het IA model verklaart deze effecten door te veronderstellen dat er laterale inhibitie is tussen deze lexicale kandidaten (zie figuur 1), waardoor het groter aantal buren de lexicale toegang vertragen. De hypothese van tweetalige parallelle lexicale toegang gaat ervan uit dat voor tweetaligen deze pool van lexicale kandidaten zich uitbreidt tot de niet-doeltaal (Dijkstra & van Heuven, 2002; van Heuven & Dijkstra, 1998).

Tot nu toe is er maar 1 studie die deze cross-linguale buureffecten onderzocht (van Heuven & Dijkstra, 1998). Ze deden dit met taken waarbij woorden in isolatie werden gepresenteerd, zoals de lexicale decisie en de woordidentificatie taak. We onderzochten het effect van het aantal cross-linguale buurwoorden in een gegeneraliseerde lexicale decisietaak en in het GECO. Voor L2 woordherkenning vonden we langzamere reactietijden in de veralgemeende lexicale decisie taak voor woorden met meer cross-linguale buren. In tegenstelling vonden we in de oogbewegingsdata dat L2 naamwoorden sneller werden gelezen en vaker overgeslagen werden als ze meer cross-linguale buren hadden. Dit toont aan dat bij de verwerking van L2 woorden, in of uit een zin context, lexicale kandidaten van L1 actief zijn en ofwel de lexicale toegang naar het doelwoord vertragen of versnellen. Voor L1 woordherkenning vonden we een hoger foutenpercentage voor hoogfrequente en een lager foutenpercentage voor laagfrequente woorden met meer cross-linguale buren in de veralgemeende lexicale decisie taak. In het lezen van L1 tekst vonden we dat hoogfrequente naamwoorden met een hogere frequente cross-linguale buur vaker werden overgeslagen, maar laagfrequente naamwoorden werden minder vaak overgeslagen. Ook waren er minder regressies naar woorden met een hogere frequente cross-linguale buur. Samengevat toonde het GECO aanzienlijke effecten van het aantal cross-linguale buren en hun frequentie. Deze effecten waren grotendeels faciliterend, terwijl de resultaten van de veralgemeende lexicale decisie taak wezen op negatieve effecten van cross-linguale buren. Deze bevindingen suggereren sterk dat lexicale kandidaten worden geselecteerd op basis van orthografische gelijkenis, niet op basis van taal. Ook suggereren ze dat tijdens het lezen van uitgebreide tekst in de meeste gevallen deze lexicale kandidaten uit de niet-doeltaal daadwerkelijk de lexicale toegang van het te lezen woord versnellen.

Hoewel deze resultaten het idee van een tweetalige taalonafhankelijk lexicale toegang, zoals door het BIA+ model (Dijkstra & van Heuven, 2002) voorgesteld, ondersteunen, zijn onze resultaten niet compatibel met de aanname van absolute lettercodering van het model. In het BIA+ model wordt voor elke letter van een woord een absolute positie gecodeerd. Deze veronderstelling over locatie-specifieke letterverwerking werd gekopieerd van het IA-model (McClelland & Rumelhart, 1981). Omdat de codering van letterposities bepaalt welke woorden orthografisch op elkaar lijken, bepaalt deze dus ook de set van lexicale kandidaten of welke woorden worden beschouwd als buurwoorden. In een lettercoderingssysteem zoals door het BIA + model wordt voorgesteld, zouden slechts *substitution* buren de lexicale toegang tot de doelgroep woord beïnvloeden. Dit zijn buren waarbij 1 enkele letter is vervangen door een andere letter. Uit onze analyses bleek echter dat *addition* en *deletion* buren, woorden gevormd door het toevoegen of verwijderen van 1 letter aan het woord, ook leestijden van woorden beïnvloeden. Dit betekent dat het lettercoderingsmechanisme van het BIA + model zou moeten worden versoepeld. Een voorbeeld van zo'n soepelere lettercodering is het mechanisme beschreven in het overlap model van Gomez, Ratcliff, and Perea (2008).

CONCLUSIES

Het onderzoek in dit proefschrift was gemotiveerd door het proberen te begrijpen van het immens complexe leesproces, dat meertaligen elke dag uitvoeren. We onderzochten de beter omlijnde vraag: Hoe krijgen late tweetaligen toegang tot de correcte lexicale representatie? Dit proefschrift draagt bij tot het tweetalig onderzoeksveld op 3 belangrijke manieren. Ten eerste, de GECO (Ghent Eye Tracking Corpus), wat verzameld werd voor dit proefschrift, is toegankelijk voor andere onderzoekers. Dit maakt hypothese toetsing en model evaluatie meer kostenefficiënt en makkelijker. Ten tweede, in de eerste twee empirische hoofdstukken van dit proefschrift tonen we aan dat taalvaardigheid, en meer bepaald taal *exposure*, een centraal concept is in het begrijpen van leesverschillen tussen L1 en L2 tweetalig lezen en tussen L1 tweetalig en eentalig lezen. Ongebalanceerde tweetaligen vertonen immers een nadeel in L2 lezen in vergelijking met L1 lezen, terwijl dezelfde tweetaligen geen nadeel vertonen in L1 in vergelijking met een andere groep eentaligen. Voor de grootte van frequentie-effecten vonden we gelijkaardige resultaten. Ten derde, onze twee laatste empirische hoofdstukken ondersteunen in sterke mate de hypothese van taalonafhankelijke lexicale toegang, die werd voorgesteld in het BIA+ model (Dijkstra & van Heuven, 2002), en dit in een natuurlijke L1 en L2 leescontext. Dit is de eerste keer dat cognaatfacilitatie wordt gedetecteerd in L1 lezen van doorlopende tekst, die ook semantische restricties bevatte. Ook is het de eerste keer dat effecten van het aantal cross-linguale buren worden gevonden in het lezen van tekst in L1 of L2.

REFERENTIES

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition*, *24*(4), 477–492.

Blythe, H. I., Häikiö, T., Bertam, R., Liversedge, S. P., & Hyönä, J. (2011). Reading disappearing text: Why do children refixate words? *Vision Research*, *51*(1), 84–92. http://doi.org/10.1016/j.visres.2010.10.003

Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2009). Visual information capture during fixations in reading for children and adults. *Vision Research*, *49*(12), 1583–1591. http://doi.org/10.1016/j.visres.2009.03.015

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: a review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology*, *58*(5), 412–424. http://doi.org/10.1027/1618-3169/a000123

Bultena, S., Dijkstra, T., & van Hell, J. G. (2014). Cognate effects in sentence context depend on word class, L2 proficiency, and task. *The Quarterly Journal of Experimental Psychology*, *67*(6), 1214–1241. http://doi.org/10.1080/17470218.2013.853090

Buswell, G. T. (1922). *Fundamental reading habits: a study of their development*. University of California Libraries.

Caramazza, A., & Brones, I. (1979). Lexical Access in Bilinguals. *Bulletin of the Psychonomic Society*, *13*(4), 212–214.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and Performance VI* (pp. 535–555). London: Academic Press.

Cowan, N. (1998). *The Development of Memory in Childhood*. Psychology Press.

Cowan, N., Morey, C. C., AuBuchon, A. M., Zwilling, C. E., & Gilchrist, A. L. (2010). Seven-year-olds Allocate Attention Like Adults Unless Working Memory is Overloaded. *Developmental Science*, *13*(1), 120. http://doi.org/10.1111/j.1467-7687.2009.00864.x

Deacon, T. W. (1998). *The Symbolic Species: The Co-evolution of Language and the Brain*. W. W. Norton & Company.

De Groot, A. M., Delmaar, P., & Lupker, S. J. (2000). The processing of interlexical homographs in translation recognition and lexical decision: support for non-selective access to bilingual memory. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *53*(2), 397–428. http://doi.org/10.1080/713755891

Dehaene, S., Le Clec'H, G., Poline, J.-B., Le Bihan, D., & Cohen, L. (2002). The visual word form area: a prelexical representation of visual words in the fusiform gyrus. *Neuroreport*, *13*(3), 321–325.

Demberg, V., & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, *109*(2), 193–210. http://doi.org/10.1016/j.cognition.2008.07.008

De Pessemier, P., & Andries, C. (2009). *GL&SCHR*. Leuven/Apeldoorn: Garant.

Diependaele, K., Lemhöfer, K., & Brysbaert, M. (2013). The word frequency effect in first-and second-language word recognition: a lexical entrenchment account. *The Quarterly Journal of Experimental Psychology*, *66*(5), 843–863.

Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of Cognates and Interlingual Homographs: The Neglected Role of Phonology. *Journal of Memory and Language*, *41*(4), 496–518.

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language*, *62*(3), 284–301. http://doi.org/10.1016/j.jml.2009.12.003

Dijkstra, T., Timmermans, M., & Schriefers, H. (2000). Cross-language effects on bilingual homograph recognition. *Journal of Memory and Language*, *42*, 445–464.

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(03). http://doi.org/10.1017/S1366728902003012

Dussias, P. E., & Cramer Scaltz, T. R. (2008). Spanish–English L2 speakers' use of subcategorization bias information in the resolution of temporary ambiguity during second language reading. *Acta*

Psychologica, *128*(3), 501–513. http://doi.org/10.1016/j.actpsy.2007.09.004

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 663–679. http://doi.org/10.1037/0278-7393.33.4.663

Ellis, E. (2006). Monolingualism: The unmarked case. *Estudios de Sociolinguistica*, *7*(2), 173–196.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A Dynamical Model of Saccade Generation During Reading. *Psychological Review*, *112*(4), 777–813. http://doi.org/10.1037/0033-295X.112.4.777

Eurobarometer, S. (2006). Europeans and their Languages. *European Commission*. Retrieved from http://pol.gu.se/digitalAssets/759/759844_Europeans_and_their_Languages_-_EC_2006.pdf

Frank, S. L., Fernandez Monsalve, I., Thompson, R. L., & Vigliocco, G. (2013). Reading time data for evaluating broad-coverage models of English sentence processing. *Behavior Research Methods*, *45*(4), 1182–1190. http://doi.org/10.3758/s13428-012-0313-y

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), 246–269. http://doi.org/10.1037/0278-7393.30.1.246

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew–English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(5), 1122.

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*(3), 787–814. http://doi.org/10.1016/j.jml.2007.07.001

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). http://doi.org/10.1017/S136672890100013X

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General*, *140*(2), 186–209. http://doi.org/10.1037/a0022256

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review*, *115*(3), 577–600. http://doi.org/10.1037/a0012667

Grainger, J., & Jacobs, A. M. (1996). Orthographic Processing in Visual Word Recognition: A Multiple Read-out Model. *Psychological Review*, *103*(3), 518–565. http://doi.org/10.1037/0033-295X.103.3.518

Grosjean, F. (1982). Preface. In *Life with two languages: An introduction to Bilingualism.* (p. viii). Cambridge, MA: Harvard University Press.

Grosjean, F. (1992). Another view of bilingualism. In *Cognitive processing in bilinguals* (pp. 51–62). Oxford, England: North-Holland.

Grosjean, F. (2010). *Bilingual: Life and Reality*. Harvard University Press.

Hoffmann, C. (2014). *Introduction to Bilingualism*. Routledge.

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, *127*(2), 277–288. http://doi.org/10.1016/j.actpsy.2007.06.003

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review*, *87*(4), 329.

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research*, *45*(2), 153–168. http://doi.org/10.1016/j.visres.2004.07.037

Kennedy, A., & Pynte, J. (2008). The consequences of violations to reading order: An eye movement analysis. *Vision Research*, *48*(21), 2309–2320. http://doi.org/10.1016/j.visres.2008.07.007

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior*

Research Methods, *42*(3), 643–650. http://doi.org/10.3758/BRM.42.3.643

Kliegl, R., & Engbert, R. (2005). Fixation durations before word skipping in reading. *Psychonomic Bulletin & Review*, *12*(1), 132–138.

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the Mind During Reading: The Influence of Past, Present, and Future Words on Fixation Durations. *Journal of Experimental Psychology: General*, *135*(1), 12–35. http://doi.org/10.1037/0096-3445.135.1.12

Kuperman, V., Drieghe, D., Keuleers, E., & Brysbaert, M. (2013). How strongly do word reading times and lexical decision times correlate? Combining data from eye movement corpora and megastudies. *The Quarterly Journal of Experimental Psychology*, *66*(3), 563–580. http://doi.org/10.1080/17470218.2012.658820

Kuperman, V., & Van Dyke, J. A. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language*, *65*(1), 42–73. http://doi.org/10.1016/j.jml.2011.03.002

Lehtonen, M., Hultén, A., Rodríguez-Fornells, A., Cunillera, T., Tuomainen, J., & Laine, M. (2012). Differences in word recognition between early bilinguals and monolinguals: behavioral and ERP evidence. *Neuropsychologia*, *50*(7), 1362–1371. http://doi.org/10.1016/j.neuropsychologia.2012.02.021

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325–343. http://doi.org/10.3758/s13428-011-0146-0

Lemhöfer, K., Dijkstra, T., & Michel, M. C. (2004). Three Languages, One ECHO: Cognate Effects in Trilingual Word Recognition. *Language and Cognitive Processes*, *19*(5), 585–611.

Lemhöfer, K., Dijkstra, T., Schriefers, H., Baayen, R. H., Grainger, J., & Zwitserlood, P. (2008). Native language influences on word recognition in a second language: A megastudy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 12–31. http://doi.org/10.1037/0278-7393.34.1.12

McClelland, J., & Rumelhart, D. (1981). An interactive activation model of context effects in letter perception. An account of basic findings. *Psychological Review*, *88*, 375–407.

Monsell, S. (1991). The Nature and Locus of Word Frequency Effects in Reading. In D. Besner & G. W. Humphreys (Eds.), *Basic Processes in Reading: Visual Word Recognition* (pp. 148–197). Hillsdale, NJ, England: Erlbaum.

Morton, J. (1970). A Functional Model for Memory. In D. A. Norman (Ed.), *Models of Human Memory* (pp. 203–254). Elsevier.

Neville, H. J., & Bavelier. (2000). Specificity and Plasticity in Neurocognitive Development in Humans. In M. S. Gazzaniga (Ed.), *The New Cognitive Neurosciences*. MIT Press.

Perea, M., & Pollatsek, A. (1998). The effects of neighborhood frequency in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 767.

Pinker, S. (1997). Foreword. In *Why our children can't read.* New York: Free Press.

Pollatsek, A., Perea, M., & Binder, K. S. (1999). The Effects of "Neighborhood Size" in Reading and Lexical Decision. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(4), 1142–1158.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology*, *52*(1), 1–56. http://doi.org/10.1016/j.cogpsych.2005.06.001

Ransdell, S. E., & Fischler, I. (1987). Memory in a monolingual mode: When are bilinguals at a disadvantage? *Journal of Memory and Language*, *26*(4), 392–405.

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology*, *41*(2), 211–236.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506. http://doi.org/10.1080/17470210902816461

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The Effects of Frequency and Predictability on Eye Fixations in Reading: Implications for the E-Z Reader Model. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(4), 720–732. http://doi.org/10.1037/0096-1523.30.4.720

Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and Aging*, *21*(3), 448–465. http://doi.org/10.1037/0882-7974.21.3.448

Reichle, E. D., Liversedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2013). Using E-Z Reader to examine the concurrent development of eye-movement control and reading skill. *Developmental Review*, *33*(2), 110–149. http://doi.org/10.1016/j.dr.2013.03.001

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*(1), 125.

Reichle, E. D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: Accounting for initial fixation locations and refixations within the EZ Reader model. *Vision Research*, *39*(26), 4403–4411.

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review*, *16*(1), 1–21. http://doi.org/10.3758/PBR.16.1.1

Titone, D., Libben, M., Mercier, J., Whitford, V., & Pivneva, I. (2011). Bilingual lexical access during L1 sentence reading: The effects of L2 knowledge, semantic constraint, and L1–L2 intermixing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(6), 1412–1431. http://doi.org/10.1037/a0024492

Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *Journal of Memory and Language*, *64*(1), 88–107. http://doi.org/10.1016/j.jml.2010.08.006

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. *Psychological Science*, *20*(8), 923–927.

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review*, *9*(4), 780–789.

Van Heuven, W. J. B., & Dijkstra, T. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language*, *39*, 458–483.

Van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190.

Wilkinson, G. S., & Robertson, G. . (2006). *Wide Range Achievement Test 4 professional manual.* Lutz, FL: Psychological Assessment Resources.

APPENDIX
DATA STORAGE FACT SHEET

% Data Storage Fact Sheet

% Name/identifier study: Bilingual and Monolingual Eye tracking corpus.

% Author: Uschi Cop, Denis Drieghe, Emmanuel Keuleers, Nicolas Dirix, Eva Van Assche, Wouter Duyck

% Date: 2015

1. Contact details

===

1a. Main researcher

- name: Uschi Cop

- address: Henri Dunantlaan 2, 9000 Gent

- e-mail: uschi.cop@ugent.be

1b. Responsible Staff Member (ZAP)

- name: Wouter Duyck

- address: Henri Dunantlaan 2, 9000 Gent

- e-mail: wouter.duyck@ugent.be

If a response is not received when using the above contact details, please send an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies

===

* Reference of the publication in which the dataset is reported:

Cop, U., Drieghe, D., & Duyck, W. Presenting GECO: An Eye-Tracking Corpus of Monolingual and Bilingual Sentence Reading. Manuscript submitted for publication.

Cop, U., Drieghe, D., & Duyck, W. Eye Movements in Natural Reading: a Comparison of Monolingual and Bilingual Novel Reading. Revised Manuscript submitted for publication.

Cop, U., Keuleers, E., Drieghe, D., & Duck, W. (2015). Frequency Effects in Monolingual and Bilingual Natural Reading. Psychonomic Bulletin & Review, in press. doi: 10.3758/s13423-015-0819-2

Cop, U.,Van Assche, E., Drieghe, D., & Duyck, W. Eye Movements in Natural Reading: a Comparison of Monolingual and Bilingual Novel Reading. Revised Manuscript submitted for publication.

Cop, U.,Dirix, N., Drieghe, D., & Duyck, W. Cross-Lingual Neighborhood Effects in Generalized Lexical Decision and Natural Reading Revised Manuscript submitted for publication.

* Which datasets in that publication does this sheet apply to?:

The sheet applies to all data reported in these studies

3. Information about the files that have been stored

===

3a. Raw data

----

* Have the raw data been stored by the main researcher? [x] YES / [ ] NO

If NO, please justify:

```
* On which platform are the raw data stored?

  - [x] researcher PC

  - [ ] research group file server

  - [ ] other (specify): ...


* Who has direct access to the raw data (i.e., without intervention
of another person)?

  - [x] main researcher

  - [ ] responsible ZAP

  - [ ] all members of the research group

  - [ ] all members of UGent

  - [ ] other (specify): ...



3b. Other files

-----------------------------------------------------------


* Which other files have been stored?

  - [x] file(s) describing the transition from raw data to reported
results. Specify: See methodology and results section in the article

  - [x] file(s) containing processed data. Specify: Excel files,
Pivot tables and figures are available on my personal computer

  - [x] file(s) containing analyses. Specify: See results section in
the article

  - [x] files(s) containing information about informed consent: a
blank copy along with a detailed information letter is saved on my
personal computer

  - [x] a file specifying legal and ethical provisions: Documents
submitted to the Ethical Committee as well as their letter of
approval are saved on my personal computer
```

  - [ ] file(s) that describe the content of the stored files and how this content should be interpreted. Specify: ...

  - [ ] other files. Specify: ...

* On which platform are these other files stored?

  - [x] individual PC

  - [x] research group file server

  - [ ] other: ...

* Who has direct access to these other files (i.e., without intervention of another person)?

  - [x] main researcher

  - [x] responsible ZAP

  - [x] all members of the research group

  - [ ] all members of UGent

  - [ ] other (specify): ...

4. Reproduction

===

* Have the results been reproduced independently?: [ ] YES / [x] NO

* If yes, by whom (add if multiple):

   - name:

   - address:

   - affiliation:

   - e-mail: