The Scope of Bilingual Sentence Planning in L1 and L2

Effects of Complexity and Bilingual Profile

RENATE GJETNES

SUPERVISORS
Professor Linda Ruth Wheeldon
Professor Allison Louise Wetterlin

University of Agder, 2021
Faculty of Humanities and Education
Department of Foreign Languages and Translation

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to sincerely thank my supervisors, professors Linda Wheeldon and Allison Wetterlin, for their invaluable guidance, advice and feedback throughout my months of working on this thesis. I really could not have asked for better supervisors, and I am so grateful for everything they have done for me. I would also like to express my gratitude to Jan Zandhuis for his efforts in setting up and running the experiment in the experimental linguistics lab at the University of Agder. Thanks are also due to Eunice Fernandes for her work with the analyses, and to Mikael André Albrecht for the helpful course on calculating word frequencies. Additionally, I would like to thank my fellow MA students who worked in the lab alongside me, Karethe Nilsen and Ellinor Skjerli, for their help with running the experiment and for their support throughout the weeks of testing. My thanks also go out to the participants of this study, without whom there would have been no experiment. Lastly, I am thankful to my fiancée, family and friends for their encouragement throughout the process of working on this thesis.

ABSTRACT

Investigating the scope of bilingual sentence planning, this two-component study compared first and second language sentence production at different complexity levels in relation to aspects of bilingual profile. Norwegian-English bilinguals filled out a questionnaire detailing their language background and proficiency in both languages, after which they performed a sentence production task in each language, producing prepositional-structure and coordinate-structure sentences at varying levels of complexity. The results of the study revealed that first and second language planning operates with different planning scopes. Prior to speech onset, the scope of first language sentence-level encoding encompasses only the first functional phrase in the sentence, while in the second language the entire subject phrase is grammatically encoded. Additionally, the results showed that planning scope in the first language may be broken up by adding an adjective to the initial noun phrase of a coordinate structure, in which case only the first noun phrase is grammatically encoded prior to speech onset. Furthermore, a relationship was found between high English proficiency and shorter production latencies in English, as well as indications of a link between high English proficiency and adjective retrieval difficulties in Norwegian. Altogether, this study provides evidence for the two languages of a bilingual employing different planning scopes and strategies.

1 Introduction: Theoretical Background and the Purpose of this Study

Speech production is a complex process involving several steps, from the generation of a conceptual message and through the stages of linguistic, grammatical and phonetic encoding, after which the utterance can be physically articulated. The fluent articulation of a sentence requires a production plan. It is generally agreed that the sentence planning process is incremental rather than holistic, meaning that only part of the utterance is planned before articulation can take place. The scope of planning refers to how extensive this pre-planned part must be before speaking can begin. Planning scope in monolingual speakers has been extensively studied; however, much remains unknown about bilingual sentence planning. This thesis concerns the scope of bilingual sentence planning as it compares to monolingual sentence planning.

After reviewing the theoretical landscape surrounding monolingual speech production and sentence planning, there follows an examination of previous research on bilingual language processing and sentence planning, as well as the relationship between the two languages compared in this study, Norwegian and English. Finally, an experiment will be described that was designed to test how bilingual planning scope differs between first and second language production, and in what ways language production is affected by linguistic background and second language proficiency.

1.1 The Sentence Planning Processes

The production of a spoken utterance is a generative process involving several steps. Simply explained, the speaker generates what she wants to say, retrieves the required items from her mind’s library of words, assembles those items in the correct grammatical structure, clothes the intended utterance in speech sounds, and then out the utterance comes. However, anyone who has produced a full sentence will know that speaking can be complicated, even in one’s native language – disfluency and speech errors abound in day-to-day speech. Exploring the sentence production process in depth will allow for a better understanding of the complexity of speech production and the load it takes on cognitive processing.

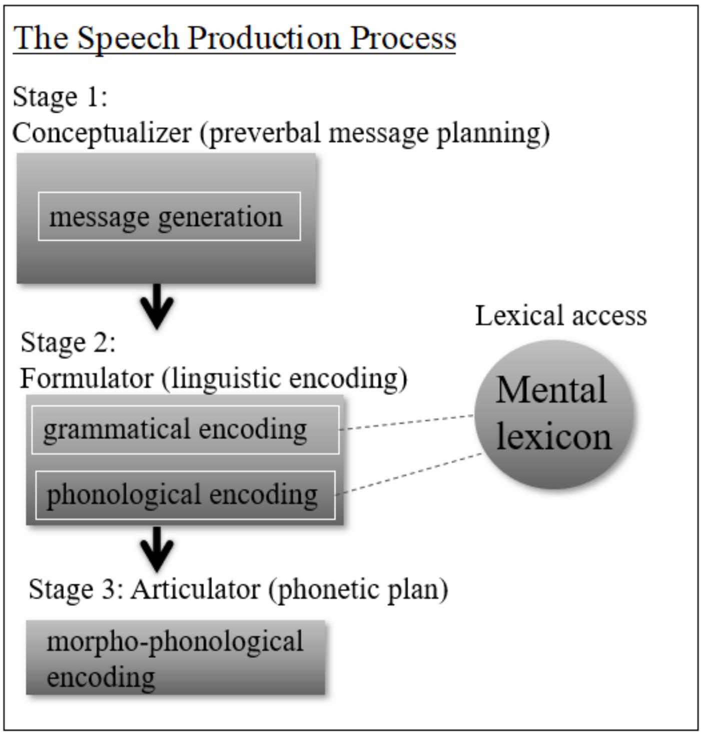


Figure 1

The three processing stages of speech production, based on Levelt's (1989) model

1.1.1 Message Generation

A model that can help explain the stages involved in sentence production is Levelt's (1989) three-stage model of speech production. The first stage in this model is message-level encoding, the generation of the preverbal message, which Levelt called the 'Conceptualizer' (see Figure 1). A message is a non-linguistic, conceptual representation of the speaker's intended utterance. According to Levelt (1989), the message is nonlinear and propositional in nature, and indicates the thematic roles of and relations between event participants, as well as the temporal structure and mood of the intended utterance. The contents of this preverbal message are also shaped by referential context, preceding exchanges between speaker and addressee, and what the speaker is attending to at the moment of speaking. In other words, at this stage, the knowledge of who is doing what and to whom is put together. For example, the concepts of ‘cat’ – agent, ‘mouse’ – patient, ‘chase’ – action and ‘yesterday’ – time, would be generated to form a preverbal message. Before it can be produced as a fully formed utterance, “the cat chased a mouse yesterday,” the preverbal message must be given linguistic form by going through a series of sentence planning processes.

1.1.2 Linguistic Encoding – Lexical Access

In order to give the conceptual message a verbal form, it needs to go through linguistic encoding, which takes place in the ‘Formulator’ stage (Levelt, 1989; see Figure 1). During this sentence-level stage of linguistic encoding, lexical items that correspond to the nonverbal concepts are retrieved from the mental lexicon, the storehouse of words in the mind, and assembled into a grammatically and syntactically correct structure. Linguistic encoding involves several processes, the first of which is lexical access and lexical retrieval, in which the required lexical items are accessed and collected from the mental lexicon.

Within the mental lexicon, the lexical items, also called lemmas (Kempen & Huijbers, 1983), contain the syntactic, semantic, morphological, and phonological information associated with the conceptual forms of words (e.g., La Heij, 1988; Levelt, 1989; Levelt, Roelofs & Meyer, 1999). During lexical access, the conceptual content of the preverbal message activates several related candidates for production, one of which must be retrieved for further processing. The activation of one lexical item results in a spread of activation to the syntactic structures in which the item can occur (e.g., Melinger & Dobel, 2005), as well as the activation of semantically, phonologically and thematically related lexical items (e.g., Levelt, 1989). The activation of the word ‘cat,’ for example, will also activate its semantic relative ‘tiger,’ its phonological relative ‘bat’ and its thematic relative ‘dog.’

The activation of related items is evident in studies showing that word production in picture naming experiments, in which participants are asked to describe pictures while their reaction times are measured, will be inhibited by the presence of related words. For example, La Heij (1988) demonstrated that the presence of a semantically related word within the same experimental trial will interfere with picture naming, an effect known as semantic interference (see also La Heij, 2005). This interference occurs because the semantically related word also receives activation during lexical access, causing the target word and the related word to compete for selection, slowing down retrieval and production of the target word. The lexical item that is selected is that which has the highest level of activation, which usually most corresponds to the meaning intended by the message (e.g., Roelofs, 1992).

1.1.3 Linguistic Encoding – Grammatical Encoding

After the appropriate lexical items have been retrieved from the mental lexicon, they must be placed into a syntactic structure via grammatical encoding. As mentioned above, lexical items contain syntactic, structural and grammatical information. Such information includes, for example, grammatical class such as noun, verb, etc., a noun's gender and number, a verb's tense, aspect and transitivity, or whether there are restrictions on the contexts in which a word can be used (Levelt, 1989). This information aids the selection process and also influences the choice of surrounding words. As this information is contained within the lemma, which is stored in the mental lexicon and accessed and selected during lexical retrieval, it follows that there must be a link between lexical retrieval and structure building, although models of grammatical encoding disagree on the degree to which these two production stages interact.

Traditional models, also called lexically driven models, (e.g., Bock & Levelt, 1994; Cleland & Pickering, 2003, 2006; Levelt, 1989; Levelt et al., 1999; Pickering & Branigan,

1998) assume some interplay between lexical retrieval and syntactic structure building, in which lexical selection precedes and drives the generation of syntactic structure. On the other hand, computational models such as Chang, Dell and Bock's (2006) Dual Path approach (see also Chang, 2002; Chang, Dell, Bock & Griffin, 2000) claim that lexical retrieval and grammatical encoding are separate, independent processes and that syntax is derived from the conceptual structure. Although no approach can account for all of the available data, Wheeldon (2011) evaluated the evidence in favour of both types of speech production models and concluded that lexical selection and syntactic structure building clearly interact during production, because associations between a word and the syntactic structure within which it has previously appeared can affect the processing of subsequent sentences. This phenomenon is known as repetition priming.

One type of repetition priming is structural priming. Studies of natural dialogue have shown that people tend to repeat sentence structures they have previously heard (Bock, 1986). Such structural priming effects have been shown to persist even when prime sentence and target sentence are separated by several unrelated sentences (e.g., Bock & Griffin, 2000; Bock, Dell, Chang & Onishi, 2007). Structural priming can facilitate sentence production – i.e. the production will be faster and more fluent – by reducing speech onset latency, the time needed to begin articulation, when the same syntax is produced consecutively (Smith & Wheeldon, 2001; Wheeldon & Smith, 2003). Structural priming has been shown to be independent of lexical content (e.g., Bock, 1989), meaning that the same sentence structure is likely to be repeated even if the words themselves are completely novel compared to the prime sentence. For example, the repetition of a passive sentence such as "the mouse is chased by the cat" makes a person more likely to continue using passive constructions when describing novel events, e.g., "the frisbee is caught by the girl" instead of the active structure in "the girl catches the frisbee." The independence of structural priming from lexical content indicates that lexical content and syntax are represented independently to some extent, though they are not necessarily fully independent (see Pickering & Ferreira, 1999, for a review).

Another kind of repetition priming is lexical priming. The repetition of lexical items has been proven to facilitate spoken word production over long time lags. For example, in picture naming experiments, the prior production of the target word will facilitate production of that same picture name even if a hundred trials intervene between prime and target trials (Wheeldon & Monsell, 1992). This effect occurs because using a word will boost its activation levels, speeding up the selection process in subsequent occurrences. An indication that lexical access and structure building are interactive processes comes from the finding that it is possible to combine structural and lexical priming. Structural priming effects increase upon the repetition of the content words, an effect known as ‘lexical boost’ (e.g., Cleland & Pickering, 2003; 2006; Corley & Scheepers, 2002; Mahowald, James, Futrell & Gibson, 2016). However, in some cases, repetition of the same word will interfere with rather than facilitate word production. If the repeated content word occurs in a different position in the target sentence than in the prime sentence, the competition between the activation of prime structure and target structure will cause a short-lived interference effect (e.g. Wheeldon, Smith & Apperly; 2011).

The evidence of a syntactic constraint to lexical retrieval further suggests an interaction between lexical retrieval and syntactic structure building. Speech error data reveal that only words from the same grammatical class, such as *mouse* and *house*, will compete for selection during lexical access (e.g., Fromkin, 1971; see also Griffin & Ferreira, 2006). There is also evidence that, regardless of a language’s word order typology (such as subject-verb-object, subject-object-verb etc.), lexical access occurs in the order of mention (e.g., Allum & Wheeldon, 2007, 2009). Accordingly, syntactic ordering processes appear to restrict lexical activation to only suitable candidates (e.g., Dell, Oppenheim & Kittredge, 2008). Thus, although evidence from structural priming studies suggest that lexical content and syntax are independently represented, the processes of lexical retrieval and syntactic structure building seem to be closely linked, with syntax having a strong influence on lexical access.

1.1.4 Phonological Encoding

Once the complex processes of lexical retrieval and syntactic structure building have been completed, the sentence elements must be given morphological, phonological and phonetic structure. The utterance has arrived at the third ‘Articulator’ stage, at which phonological encoding takes place (Levelt, 1989; see Figure 1). There is general consensus that phonological encoding is a sequential process of slotting morpho-phonological units into a phonological framework (e.g., Dell, 1986; Levelt, 1989; Levelt, et al., 1999; Roelofs, 2015). During phonological encoding, the lexemes associated with their corresponding lemmas are accessed. As maintained by lexically driven models of speech production, there is a division between linguistic units containing syntactic and semantic information, lemmas, and linguistic units containing morphological and phonological information, lexemes (e.g., Bock and Levelt, 1994; Cleland & Pickering, 2003; 2006; Levelt, 1989; Levelt et al., 1999; Pickering & Branigan, 1998). In other words, lexemes contain the composition of a word’s internal structure of morphemes and its organisation of speech sounds.

Research that supports this division between lemmas and lexemes has shown that syntactic and semantic information is accessed before phonological encoding takes place (e.g., Schriefers, Meyer & Levelt, 1990). Thus, lemmas operate during the lexical retrieval and grammatical encoding processes, and lexemes are accessed during the phonological encoding process. Additionally, evidence from speech error studies and morphological and phonological priming experiments suggests that morphemes are represented at the level of phonological encoding, indicating the existence of a morpho-phonological level containing both morphological and phonological details (e.g. Dell, 1986, Levelt et al., 1999). In other words, the generation of lexemes involves the construction of a morpho-phonological representation of the utterance, after which more detailed phonetic representations are constructed (see also Wheeldon & Konopka, 2018, for a review). The phonological content is then generated sequentially in the order of articulation (e.g., Damian, Bowers, Stadthagen-Gonzales & Spalek, 2010; Levelt & Wheeldon, 1994; Meyer & Schriefers, 1991).

After this journey through the speech production processes, the utterance is ready to be produced as a full-fledged utterance. Despite the complexity of this process, speaking seems to happen instantly and effortlessly, which begs the question: must all of the different encoding processes be completed before speaking can begin?

1.2 Sentence Planning: Incrementality and Planning Scope

It is important to distinguish between the scope of higher-level, conceptual message planning, and the scope of the lower-level planning processes that include lexical access, grammatical encoding and phonological encoding. A range of studies show that holistic sentence planning, i.e. that the whole of a sentence is generated and processed sequentially at each level prior to speech onset, is feasible in the case of very simple utterances (e.g., Bock, Irwin, Davidson & Levelt, 2003; Griffin & Bock, 2000; Konopka & Meyer, 2014; Kuchinsky, Bock & Irwin, 2011). However, it is unlikely that holistically planned sentences are typical in speech production, given the speed with which people can start producing complex utterances (e.g., Levelt, 1989). Instead, message encoding, linguistic encoding and phonetic encoding can operate incrementally, in parallel.

In incremental sentence planning, only a small portion of the intended utterance is constructed before speaking can begin, and segments of conceptual and linguistic information are continuously generated in increments after the onset of speech production (e.g., Kempen

& Hoenkamp, 1987; Levelt, 1989). Simply put, speakers make up an utterance in chunks as they go along, articulating early elements while the upcoming elements are still being planned. For example, in “the cat chased a mouse yesterday,” the message element concerning the agent, “the cat” is generated first, then its lexical information retrieved, and while this information is being grammatically encoded, the next message element concerning the action “chased” is being generated. Then, while “the cat” is being phonetically encoded and subsequently articulated, the grammatical encoding of the action will take place while the time information, “yesterday,” is being conceptualised, and so on. In this way, each element of the sentence is consecutively added to the utterance in small increments, and speaking can begin as soon as the processing of the first element has been completed, allowing for speed and fluency.

1.2.1 The Scope of Message Planning

A number of proposals exist in the literature based on the idea that message planning and lower-level planning processes can operate incrementally and with different scopes. Firstly, an important question is whether the whole message must be conceptualised before speaking can begin, or whether message planning can keep unfolding incrementally after speech onset. According to the Initial Preparation view, messages are generated holistically as a single element before being passed down to sentence-level processing, which may then proceed smoothly and independently without further interference from the message level. Fluent speech depends on the message being prepared in its entirety before speech onset, as adding to the message after speaking has begun may cause disfluency or ungrammaticality (Brown-Schmidt & Konopka, 2014).

A number of speech production models adhere to the principle of holistic message generation, at least at a clausal level. For example, in Ward’s (1992) ‘Flexible Incremental

Generator' (FIG) model, the first step of speech production is the holistic generation of the conceptual structure of a clause, followed by a second step of word-for-word grammatical encoding. Prior to speech onset, conceptual planning must be completed for the first clause of an utterance, but grammatical encoding needs to be completed for only the first word. In contrast to this model with its two clearly defined stages is that of Garrett (1975, 1982; as cited in Smith & Wheeldon, 1999), in which the different planning stages are simultaneously active, each working on different parts of the sentence. At the 'functional' planning stage, the lemmas within a clause are simultaneously accessed and assigned thematic roles, while the 'positional' stage specifies syntactic, morphological and phonological aspects of a clause in successive phrasal chunks. Prior to speech onset, functional level processing must be completed for at least the first clause of a sentence, and positional stage processes must be completed for the first phrase.

In contrast to Initial Preparation, Continuous Incrementality posits that sentence-level processing can begin while message-level information is still being generated, allowing for new information to be added to the message incrementally while speaking is ongoing. Throughout the planning process, sentence-level processes continue to interact with message-level processes, and new message-level information is rapidly incorporated as soon as it is available without causing disruptions (Brown-Schmidt & Konopka, 2014). A speech production model that supports incremental message planning is the IPG, the incremental procedural grammar, developed by Kempen and Hoenkamp (1987). In this model, the first, conceptual planning level continues to process fragments after the release of message output to the second level of grammatical encoding. The two levels thus process in parallel, and multiple fragments are also processed simultaneously within the same level, allowing for successive releases of fragments before the grammatical encoding of prior conceptual fragments is complete. Fragment size at either the conceptual or the grammatical level varies according to processing speed, and may thus correspond to lemmas or whole clauses. Hence, prior to speech onset, there is no fixed amount of planning that needs to be completed, and planning scope may be less than a clause. The IPG thus allows for less than a clausal scope of conceptual planning before speech onset.

On the one hand, there is some evidence of holistic message planning in line with Initial Preparation. Studies have shown that in cases where the message planning process is perceptually driven, e.g. stimulated by a visual scene, it is possible that a complete message could be encoded before speech onset, and that the rest of the planning process would subsequently be conducted incrementally (e.g. Bock et al., 2003; Griffin & Bock, 2000; Konopka & Meyer, 2014; Kuchinsky et al., 2011). For example, Griffin and Bock (2000) asked participants to describe line drawings of simple events while tracking their eye movements. Eye-tracking is useful because speakers are known to fixate the referents in a target sentence in the intended order of mention, allowing for insight into their planning process and the development of a message plan (e.g., Konopka & Brown-Schmidt, 2014). Griffin and Bock found that speakers' apprehension of the events guided their selection of event participants and thus that conceptualisation preceded formulation, supporting a holistic message planning scope.

On the other hand, Continuous Incrementality has received considerable support from empirical evidence (e.g., Brown-Schmidt & Konopka, 2008, 2014; Brown-Schmidt & Tanenhaus, 2006). For example, in two eye-tracked picture description experiments, Brown-Schmidt and Konopka (2014) tested the prediction that new message elements could be incorporated incrementally into an utterance after speech onset. In an interactive task, participant and experimenter sat in the same room at separate computers and took turns instructing each other to click on objects by describing their movements, as neither could see the other's screen. The participants' eye movements were tracked and their speech was recorded. The visual display was based on a picture description procedure developed by Brown-Schmidt consisting of a grid displaying a number of pictures, where the picture(s) to be described are marked by flashing or shifting in one direction. In this paradigm, several pictures can be marked for description simultaneously, allowing for the testing of the production of complex structures such as conjoined noun phrases (e.g., "the balloon and the fairy"). The inclusion of contrasting pictures, such as a size contrast, in the picture grid makes it possible to test the speed with which sentence plans can be updated with a modifier when the contrast is discovered after speech onset (e.g., "the balloon and the small fairy").

Brown-Schmidt and Konopka found that participants were able to incorporate new message elements, presented as size contrasts, into developing utterances even after speech onset without disrupting fluency, as long as the contrast was fixated before the onset of the phrase in which the modification occurs. Thus, this study suggests that messages can be incrementally updated after speech onset, in line with Continuous Incrementality. In a natural speaking environment, the content of the preverbal message is shaped by referential context and audience demands, and thus the message may have to be adjusted after speaking has begun in response to contextual demands. Brown-Schmidt and Konopka (2014) propose that the continuous incremental planning of small message elements would allow speakers to make these revisions as necessary without disrupting the flow of conversation.

1.2.2 Variation in Planning Scope

When it comes to sentence planning as a whole, speakers have been known to use different planning strategies when producing utterances under challenging conditions such as time pressure or increased cognitive load (Ferreira & Swets, 2002; Slevc, 2011; Wagner, Jescheniak & Schriefers, 2010). This implies a certain flexibility in sentence planning. Two accounts attempt to explain the use of different sentence planning strategies. According to

Linear Incrementality, each message element is encoded independently and in a linear succession at each level, allowing for the next element to be planned at the message level and encoded at sentence level while articulation is in progress according to the order of production (e.g., Gleitman, January, Nappa & Trueswell, 2007; Konopka & Meyer, 2014; Konopka, Meyer & Forest, 2018). Linear Incrementality has a lot in common with Continuous Incrementality in that both allow for the continuous planning of message elements after lower-level processing and articulation have begun. In contrast, Hierarchical Incrementality posits that message-level and sentence-level planning are separate processes, by which the larger conceptual representation is encoded first, and linguistic encoding follows after conceptual planning has finished specifying the relational content and the order of message elements (e.g. Bock et al., 2003; Griffin & Bock, 2000; Konopka & Meyer, 2014; Konopka et al., 2018). Accordingly, Hierarchical Incrementality largely corresponds to Initial Preparation, with the holistic generation of the message before sentence-level processing can begin. Which of these planning strategies to use has been proven to vary both from speaker to speaker, and within the same speaker depending on context (see Konopka et al., 2018, for a review). That these planning strategies may coexist within the same speaker to use when appropriate is important to note in a bilingual context, as will be discussed further in section 1.3.2.

One variable in determining the choice of planning strategy may be cognitive processing load. The preparation of complex utterances and the coordination between message-level and sentence-level processing increases processing requirements, which affects the accuracy and speed of processing and may have consequences for the degree of incrementality possible during speech production (e.g., Ferreira & Pashler, 2002; Pashler, 1994). Processing capacity limitations may constrain planning in favour of generating smaller and more economic increments, although individuals with a higher working memory capacity may opt for planning in larger increments (e.g., Swets, Jacovina & Gerrig, 2014; see also Wheeldon & Konopka, 2018). Word length is another factor that may have an effect on planning scope. Longer words, which are presumably more complex, should require more preparation than shorter words, yet studies of word length effects have yielded inconsistent results regarding word length effects on naming latencies (e.g., Damian et al., 2010; see also Ferrand & New, 2003, for a review). However, there are enough studies reporting an effect of syllable number on the onset of picture naming that there may be a connection between syllable number and planning scope (e.g., Meyer, Belke, Häcker & Mortensen, 2007; Meyer, Roelofs & Levelt, 2003; Santiago, MacKay, Palma & Rho, 2000). Thus, many factors contribute to considerable variation in sentence planning at the different processing levels, such that the sentence planning process is not uniform for all speakers in all contexts, but flexible and subject to individual and situational differences.

1.2.3 Finding the Minimal Planning Unit

Taking all this variation into consideration, the question remains as to what is the minimal unit of planning that needs to be processed at each level before speaking can begin. As each processing level operates with a different scope, the minimum planning scope for message generation is not necessarily the same as that for subsequent planning levels. One factor in determining the size of message increments may be whether the message planning process is perceptually or conceptually driven. As previously mentioned, messages driven by the uptake of perceptual information may operate with a holistic message planning scope (e.g., Griffin & Bock, 2000). However, visually driven utterances may also consist of very small increments which need not exceed a single message element at a time, as each object may be named in the order of mention without planning ahead to the next message elements (e.g., Griffin, 2001; Meyer, Sleiderink & Levelt, 1998; Meyer, Wheeldon, van der Maulen & Konopka, 2012). Thus, even when speakers are describing a visual scene, the size of message increments may vary considerably. Moreover, most natural language does not consist of simply describing things that the speaker can see. Utterances that are planned conceptually rather than perceptually, and thus have structures that are less predictable and more complex, require more information at the message level, such as the identities and roles of the protagonists and how they are related (e.g., Konopka & Brown-Schmidt, 2014).

The past couple of decades has seen much development in research on planning scope and the search for what constitutes the minimal conceptual and grammatical planning units required for speech onset. A notable example is a study by Smith and Wheeldon (1999), who sought to investigate the minimum scope of processing needed before speaking can begin. Levelt and Maassen (1981) had previously demonstrated that speech onset latencies are longer for sentences with conjoined noun phrases, e.g., “The square and the diamond are rising,” than for coordinates sentences, such as “The square is rising and the diamond is rising.” This finding suggests that the whole utterance cannot have been planned before speech onset, as coordinate sentences are more complex than conjoined noun phrases and should therefore have taken longer to articulate if the entire utterance had been planned before speaking. To narrow it down further, Smith and Wheeldon’s (1999) study tested whether the minimum processing scope before speech onset consists of a word, a phrase or a clause, by measuring the speech onset latency for the production of different sentence types.

Specifically, they tested whether the first phrase of an utterance received more attention from conceptual and grammatical encoding before speech onset than the remainder of the utterance. Sitting in a sound-attenuated booth, participants were shown three pictures in a line on a computer monitor, e.g., a dog, a foot and a kite. Each picture would move up or down, and participants were asked to describe the scenario from left to right using a single clause sentence. Depending on the number of pictures moving in tandem, the display would prompt the production of one of the following two sentence types: a complex-simple sentence which began with a conjoined noun phrase, such as “the dog and the foot move above the kite,” or a simple-complex sentence beginning with a simple noun phrase, such as “the dog moves above the foot and the kite.” Reaction times were automatically recorded. Smith and Wheeldon found that latencies were indeed greater to complex-simple sentences than to simple-complex sentences, indicating that grammatical processing of the first phrase is prioritised. The difference in latency is due to the higher complexity of the initial phrase of complex-simple sentences compared to the simple-complex sentences. If grammatical encoding had been completed for either the whole clause or only the first word, there would have been no significant difference in latency between the sentence types.

This finding is consistent with a phrasal processing scope for grammatical encoding rather than a clausal scope, meaning that grammatical encoding does not encompass the entire sentence before speech onset, nor is lexical access completed for the entire sentence before speech onset. Rather, only the first phrase is grammatically encoded prior to speech onset. Additionally, Experiment 2 and 3 revealed that latencies to double clause sentences were longer than to single clause sentences, indicating that some processing of the second clause is initiated but not completed before speech onset and thus that conceptual encoding reaches beyond the first clause. Smith and Wheeldon’s findings thus indicate the existence of a phrasal scope for lexical access and a clausal scope for higher-level, conceptual processing prior to lexical access.

One limitation of the Smith and Wheeldon (1999) study is that it does not reveal the processing level at which the phrasal scope operates, and another is that there is still a possibility that the processing scope is lexical rather than phrasal, in which case speakers prefer to retrieve the two initial content words of an utterance prior to speech onset rather than the first phrase. In the case of the simple noun phrases, the second content word was always the verb ‘moves,’ whereas in the conjoined noun phrases it was a different picture name, which might explain the increased speech onset latency for conjoined noun phrases. The critical planning unit, i.e. the scope of grammatical planning before speech onset, could be the subject phrase or the agent, or even a smaller grammatical unit such as the initial noun phrase. In order to examine which of the above constitutes the critical planning unit, Allum and Wheeldon (2007) took advantage of the head-final properties of Japanese, in which the agent or head of the subject phrase is the second grammatical unit in a sentence, to test the hypothesis that the critical planning unit comprises the initial verb argument phrase, i.e. the initial phrase that represents a major sentence unit such as subject, object, etc. with its associated elements.

Allum and Wheeldon designed four picture description experiments in which items, pictured as line drawings, were grouped together by colour to form phrases, thereby creating Japanese sentences that increased in complexity for each experiment. Examples are sentences that translate into the likes of “the dog is above the clock and the flower” and “the dog and the clock are above the flower” etc (Experiment 1), “the flower above the dog is red” and “the dog and the clock are red” (Experiment 2), “the flower and apple and trousers above the dog are red” and “the trousers and apple above the dog and flower are red” (Experiment 3), etc. Inside a sound-attenuated booth, the pictures, grouped together in different ways according to the type of sentence to be elicited, were displayed on a computer screen in front of the participant, who had beforehand been instructed in the procedure and the required sentence patterns per display. Reaction times were automatically recorded, and errors were recorded by an experimenter seated outside the booth.

What Allum and Wheeldon found was, first, that sentence production latencies increased depending on the length of the sentence-initial verb argument phrase, suggesting that this phrase was the minimal unit of planning before speech onset (Experiment 1).

Second, reaction times also increased when the grouping of elements within the verb argument phrase increased the size of the sentence-initial functional phrase, even when the length of the verb argument phrase was the same overall. Accordingly, the initial functional phrase, even if it was a subordinate phrase, was processed more thoroughly than the verb argument phrase as a whole (Experiment 2-3). A functional phrase is any unit that serves a single function, such as a modifier, theme or agent, and cannot be broken down into lesser functions. For example, the subject phrase “the flower above the dog” can be split into two functional phrases, ‘the flower’ and ‘above the dog,’ whereas “the flower and the dog” consists of one functional phrase. In Japanese, the literal translation of the English sentence “The flower above the dog is red” would be *“Dog above flower red is,” which would make the initial functional phrase ‘dog above’ a subordinate phrase and leave the head of the phrase, ‘flower,’ in second position. Thus, the scope of grammatical encoding does not necessarily include the head of the phrase, and may encompass less than the initial verb argument phrase as a whole. When Experiment 2 was replicated in English with native English speakers, results showed a similar effect of sentence production latencies being dependent on the size of the first functional phrase rather than the subject phrase as a whole, indicating that the results from the previous experiments were not caused by language differences (Experiment 4). Furthermore, Experiments 2 and 4 showed that sentence-initial conjoined noun phrases (henceforth referred to as coordinate structures) had longer speech onsets than structures in which noun phrases were modified by prepositional phrases (henceforth referred to as prepositional structures), suggesting that coordinate structures have longer processing scopes than prepositional structures.

These findings expand on those by Smith and Wheeldon (1999) by confirming that the first verb argument phrase is more thoroughly processed than subsequent phrases, and adding that the initial functional phrase, a smaller unit than the verb argument phrase, is processed more thoroughly than the whole verb argument phrase, even if it does not represent a major element in the sentence. Thus, it appears that speech can be initiated once the smallest functional unit has been lexicalised. Since the initial functional unit is larger in coordinate structures, in that both nouns need to be lexicalised, it follows that planning scope for coordinate structures is larger than that of prepositional structures.

Building on the 2007 study, Allum and Wheeldon (2009) conducted a second study in both Japanese and English in order to investigate at which processing level the functional phrase scope operates. Specifically, how does the prelinguistic and presumably nonlinear message ensure that lexical access transpires in the correct order, regardless of language? By including a picture preview in advance of the picture description experiment in order to activate lexical information associated with one of the trial picture names, they were able to examine how planning is affected by the prior access of lexical information. The results showed that, for both English and Japanese, a preview of the first noun in a coordinate structure substantially reduced onset latency. There was also a lesser effect previewing the second noun in a coordinate structure. However, there was no effect of preview on a second noun that was not a part of the initial verb argument phrase, such as with prepositional structures which consist of a noun phrase modified by a prepositional phrase containing a second noun. In Japanese, the modifying prepositional precedes the head noun phrase, but preview still had no effect on the second noun even though this was the head of the verb argument phrase. These findings were taken as evidence that only the lexical items in the initial phrase need to be retrieved before speech production can begin. Additionally, the results suggest that the order of lemma activation is affected by word order, such that syntax mediates between message generation and lexical access. This is consistent with Chang et al.'s (2006) Dual Path model in which conceptual elements interact with grammatical/syntactic processes before lexical access occurs. Additionally, when comparing sentence production in Japanese and English, Allum and Wheeldon found that preverbal message planning seems to be universal across languages, which suggests that any L1-L2 differences in bilingual sentence production are likely not due to constraints on conceptual encoding processes. The final experiment also supported the findings of Allum and Wheeldon (2007) that coordinate structures take longer to plan than prepositional structures.

In sum, these studies by Smith and Wheeldon (1999) and Allum and Wheeldon (2007, 2009) demonstrated that linguistic encoding is completed for the first functional phrase of an utterance prior to speech onset, but that higher-level conceptual planning can extend much farther ahead, up to and including the second clause of the utterance before articulation occurs. Additionally, initial coordinate structures have longer speech onsets than initial prepositional structures. This means that coordinates take longer to plan, indicating that both nouns in the coordinate structure are planned before speech can begin, while in a prepositional structure, speaking can begin as soon as the initial noun phrase is planned, and the prepositional phrase is planned incrementally after speech onset. These are important findings around which the present study is heavily based (see section 1.4).

As regards the scope of phonological planning, there is evidence in favour of a variety of planning scopes. For example, Meyer (1996) found that the scope of phonological planning prior to speech onset extends only as far as to the first noun, even in cases where the lexical items for both nouns in a coordinate structure have been selected at speech onset, demonstrating that the lexical/grammatical planning unit is larger than that of the phonological level. Other studies have shown that the scope of phonological encoding does not exceed a single phonological word (Levelt, 1989; Levelt et al., 1999; Wheeldon & Lahiri, 1997; 2002). A phonological word is a unit consisting of a single stressed syllable and any unstressed syllables associated with it, e.g., “the um<u>brel</u>la in the...,” and may thus encompass more than a lexical word. Speech error data suggest that phonological encoding operates with a phrasal planning scope (e.g., Garrett, 1975, 1980, as cited in Wheeldon & Konopka, 2018).

Phonological priming and interference effects suggest a fluid unit of phonological encoding, as phonological similarity between words (e.g., ‘mouse’ – ‘mouth’) has also been shown to affect speech production. For example, Schnur, Costa and Caramazza (2006) found that a phonologically similar distractor word would facilitate speech production when presented alongside a picture depicting an action which the participant must describe while ignoring the distractor word. For example, the presence of the distractor word ‘jug’ facilitated the onset of the sentence “the girl jumps” compared to the unrelated distractor word ‘sneer.’ However, the presence of phonologically similar words has also been found to interfere with planning. For example, Oppermann, Jescheniak and Schriefers (2010) found that a distractor word, e.g., ‘*mauk*,’ related to the subject or object, e.g., ‘*Maus*’ (German: ‘mouse’) interfered with the target utterance if the subject or object appeared in a non-initial phrase. Both of these findings suggest that phonological planning can surpass phrase boundaries. The great variation in the results of phonological planning studies indicates that the unit of phonological encoding is not fixed, but that phonological planning scope seems to be flexible and to some extent independent of the scopes of conceptual and grammatical encoding.

The above sections have established that planning scope may be subject to variation depending on such variables as processing level, changes in linguistic and visual context, cognitive processing load, priming effects, word length, and sentence structure. However, if planning scope depends on so many different circumstantial elements, then what happens when the language changes? The question is whether bilingual speakers employ the same planning strategies as monolinguals, i.e. whether the sentence planning process of a bilingual person remains the same regardless of which language that person is speaking.

1.3 Bilingualism and Sentence Planning

1.3.1 Language-Specific vs. Language-Nonspecific Processing

Much remains to be explained regarding how bilingual speakers coordinate message-level and sentence-level processes in their two languages, and in what ways second-language (L2) planning differs from the first language (L1). To understand bilingual production entails understanding how information is passed between processing levels, how the system knows how to restrict selection processes to apply only to the target language, and how non-target lexical items affect the selection processes. There is evidence to suggest that the intention to speak in either the L1 or L2 is generated as a part of the preverbal message (e.g., Costa, Miozzo & Caramazza, 1999; De Bot, 1992; Green, 1998). Furthermore, research has shown that the semantic systems of a bilingual activate in parallel for both of the bilingual's mental lexicons, and that both lexicons are active in parallel (e.g., Colomé, 2001; Hermans, Bongaerts, De Bot & Schreuder, 1998; van Hell & Dijkstra, 2002). The question is how the bilingual processing system navigates between the languages such that only the intended language is selected for production.

Two views attempt to explain bilingual language processing. According to the language-specific view, during lexical access, the meaning of each word contains a feature specifying the language to which it belongs (e.g., Green, 1998; Levelt, 1989). For example, despite being translations, the English word 'table' and the Norwegian word '*bord*' do not express the same meaning or contain the same conceptual content; the meaning of 'table' includes its specification as an English word, and the meaning of '*bord*' includes a Norwegian feature. When selecting a lexical item, a language-sensitive selection mechanism ensures that the lexical item chosen matches the speaker's intended meaning and intended language (Roelofs, 1998). Thus, by deciding to speak in either language, the planning process will be language-specific from the start of conceptualisation until the articulation of the utterance (e.g. Costa et al., 1999). The existence of the other language will not interfere with the lexical selection process.

In contrast, the language-nonspecific view is that the contents of the preverbal message will activate all relevant lexical items during lexical access regardless of language, and thus that words from the non-target language will also act as competitors during the lexical selection process. The lexical item that is most highly activated will end up being selected, due to an inhibition mechanism which suppresses activated words from the non-target language (Green, 1998). Because of this competition between the two languages' lexicons, lexical items from the non-target language will interfere during lexical access, which may cause disruptions in speech (e.g., Hermans et al., 1998).

One important difference between the two views lies in the role of the non-target language during lexical selection. Language-specific models postulate that the non-target language is irrelevant during lexical selection and thus that there is no competition, as the lexical selection mechanism only considers lexical items from the target language (e.g., Costa et al., 1999; Roelofs, 1998). The language-nonspecific view, however, maintains that lexical items from both lexicons of a bilingual will compete for selection. Accordingly, if the presence of words from one language can be proven to interfere with the selection of semantically related lexical items in the other language, this suggests that the items compete for selection and thus that lexical selection is language-nonspecific. On the other hand, if the presence of words from one language can be shown to facilitate the selection of semantically related items from the other language, this suggests that the word from the one language spreads activation to semantically related words in the other's lexicon, due to the lexicons sharing semantic systems. The lexical items will therefore receive higher levels of activation, facilitating production (Costa et al., 1999).

As it is, neither view can wholly explain bilingual language processing. Both language-specific models and language-nonspecific models yet leave important questions unanswered, such as how the language-specific models' selection mechanism knows how to restrict selection to only the lexical items of the target language, and how the language-nonspecific models can satisfactorily explain the mechanism that prevents the selection of words from the non-target language (Costa, 2005). In his review on the subject of language-specific versus language-nonspecific speech processing, Costa (2005) hypothesises that proficiency in the L2 may cause a shift from language-nonspecific toward language-specific processing in bilingual speakers. As such, it may be that there is not one correct answer to the question of whether language processing is language-specific or language-nonspecific, and that the language-specificity of bilingual language processing is subject to individual variation, much like planning strategies.

1.3.2 Bilingual Planning Strategies

The heightened processing load of speaking in a second language has consequences for message preparation as well as on the incrementality of sentence planning. It is assumed that linguistic encoding is slower and more difficult in the L2 than in the L1, which in turn slows down L2 production. The coordination of message- and sentence-level processes appears to be different between L1 and L2 sentence production, as L2 message-level encoding appears to be more extensive compared to the L1 (e.g., Konopka et al., 2018). To compensate for the heightened difficulty of producing sentences in the L2, bilinguals may employ different planning strategies than they would if they were speaking their L1. Furthermore, bilinguals' planning strategies may be shaped by their experience with each language, which may influence message generation and predict the difficulty of linguistic encoding.

Empirical evidence suggests that under cognitive load, such as when speaking in a less-practiced language, speakers will use smaller planning windows during sentence production in order to spread the cognitive effort across the utterance (e.g., Ferreira & Swets, 2002; Wagner et al., 2010). Therefore, bilinguals may opt for a linearly incremental planning strategy, breaking the utterance up into smaller increments. However, although this planning strategy would reduce processing load per increment, the speaker would also need to continuously plan new increments after speech onset, which may lead to disfluencies and repairs. Another option would be to go for a hierarchical planning strategy, planning a larger message before speech onset in the L2 and beginning linguistic encoding after the message has been generated. Speakers have been known to shift from linear to hierarchical planning if the linguistic encoding processes become more difficult to perform (e.g., Konopka & Meyer, 2014). Using this strategy would presumably lead to less disfluencies and need for repairs, but it would also entail a delay in speech onset.

Konopka, Meyer and Forest (2018) set out to examine to what extent bilingual planning is influenced by linguistic experience, and specifically whether bilinguals would opt for a linearly incremental planning strategy or a hierarchically incremental planning strategy depending on their experience with the language. They conducted a series of experiments in which Dutch-English bilinguals generated event descriptions in both the L1 and L2 prompted by a picture grid (as described above in section 1.2.1), while their gaze patterns were tracked and their speech onsets measured. After the experiment, participants completed a questionnaire about their linguistic background and L2 experience, as well as a LexTale test in both languages. A LexTale test measures vocabulary knowledge and general proficiency in a language. Results revealed that speakers tended to select ‘easier’ sentence structures, such as actives over passives, in their less-practiced language, English. However, more proficient English speakers used the more complex passive structures more frequently in English than did low-proficient speakers. Additionally, initiation of sentences in the L2 was slower than in the L1. Furthermore, eye tracking revealed that the agent of the sentence was fixated earlier in L1 sentences than in L2 sentences, indicating that speakers delayed linguistic encoding in L2 because of a separation of message-level encoding and sentence-level encoding. This suggests that L2 speakers devoted more attention to extensive message planning, consistent with Hierarchical Incrementality. In the L1, on the other hand, production was more consistent with a linearly incremental planning strategy with greater syntactic flexibility. This finding suggests that L2 planning is generally more hierarchically incremental than linear, and that earlier and more expansive message generation is preferred in a less-experienced language.

Konopka and colleagues also found that, when including a preview of half of the target pictures before each block (Experiment 2), in L2 sentences the picture preview removed the effect of delaying linguistic encoding and facilitated conceptual processing. When including a noun preview before beginning the experiment, in which half of the nouns were names of agents in the target events (Experiment 3), this priming significantly facilitated linguistic encoding of target agents in both languages, but linguistic encoding in the L2 was facilitated to a greater extent than in the L1. Furthermore, when including a preview of the most frequently used verbs for describing the target pictures, this verb preview had a strong faciliatory effect on speech onsets. The way the previewed verbs functioned as a prime, increasing the efficiency of conceptual encoding, suggests that early sentence planning includes the sentence verb. Importantly, the preview manipulation had a larger benefit in the L2 than the L1, suggesting a broader scope of planning in L2 than in L1.

Another finding by Konopka et al. (2018) was that the more experience the participants had with the task, the more linearly incremental was their planning process, showing that participants are able to utilise the strategy of Linear Incrementality in the L2 if they feel comfortable with the task. Additionally, the more proficient L2 speakers tended to shift towards linearly incremental planning, showing greater syntactic flexibility. This suggests that the choice of planning strategy is experience-driven, and that the higher the proficiency, the more L1-like the bilinguals' planning process will be. For this reason, factors related to L2 experience and proficiency are of great interest in relation to the bilingual sentence planning process, and of particular importance for the purposes of the present study.

1.3.3 The L1-L2 Relationship

As this study focuses on Norwegian-English bilinguals, it is important to consider the relationship between these languages and how they may affect each other in the bilingual mind. Second language learning is to a large degree shaped by transfer from the L1, therefore the structure of the L1 has great importance when it comes to learning the L2. The transfer will be more successful if the two languages are closely related conceptually, materially and linguistically (e.g. MacWhinney, 2005). Both being Germanic languages, Norwegian and English belong to the same language family and thus follow the same general structure when it comes to syntax; both as a general rule follow the SVO (declarative) and VS (interrogative) word order, both prefer active sentence structures over passives, and both languages rely on word order to mark the syntactic function of a phrase, as neither language marks grammatical case (e.g., nominative, accusative) like many other languages do. Thus, both languages are subject to similar constraints on word order during grammatical encoding.

However, some language-specific constraints do exist, whose backward dependencies may affect processing scope (e.g., Wheeldon, 2012). For example, in Norwegian the indefinite article takes a different form according to the gender of the noun, as in '*et hus*' (neuter) – 'a house,' but '*en sykkel*' (masculine) – 'a bicycle.' The speaker would need to first retrieve the noun in order to know which indefinite article to select. This would not make any difference in languages such as English where nouns do not have gender. Relevant to this study, Norwegian adjectives are inflected according to the plurality, gender and definiteness of the noun they qualify. In the case of attributive adjectives, which are placed before the noun, the speaker is required to first retrieve the correct form of the noun in order to select the correct form of the adjective. For example, ‘a green apple’ would be ‘*et grønt eple*,’ (neuter noun), but ‘a green book’ would be ‘*ei grønn bok*’ (feminine noun), and when put in the definite form, ‘the green book’ would be ‘*den grønne boka*.’ Such language-specific features may create differences between the processing scope of one language compared to the other that should be taken into account when studying bilingual sentence planning.

As Norwegian and English are related, it is also important to consider that there will be a large number of cognates. Within the field of psycholinguistics, cognates are translation words that are similar in orthographic and/or phonological form in two languages, e.g. the Norwegian ‘*papir*’ and English ‘paper’. Cognates are known to affect the speed of sentence production. As mentioned above, during lexical selection, competition increases the more semantically similar the distractor word is to the target word due to the semantic interference effect. Therefore, a cognate may be assumed to slow down lexical retrieval by virtue of its similarity in meaning to the target word.

However, rather than interfering, cognates have been found to facilitate the retrieval of the target word and lead to faster naming latencies in bilingual production. Costa, Caramazza and Sebastian-Galles (2000) examined the effect of cognates on picture naming latencies, and found that bilinguals named pictures with cognate names faster compared to noncognate names, while monolinguals named both cognate and noncognate pictures equally fast. This effect was greater when bilinguals named pictures in their less dominant language. These findings suggest a strong cognate facilitation effect on bilingual lexical retrieval, and also that phonological segments appear to be activated by similar lexical items in a language non-specific way (Costa et al., 2000). The cognate facilitation effect has also been observed in a study of the tip-of-the-tongue (TOT) phenomenon in bilingual speakers, where bilinguals experienced fewer TOT-states with cognate words than with noncognates (Gollan & Acenas, 2000, as cited in Costa, 2005). The fact that cognates have been known to facilitate picture naming in bilinguals is important to keep in mind for the present experiment, as its success depended on avoiding priming and facilitation effects from one language to the other. Therefore, only non-cognates and translation words that differed phonetically and orthographically were used as stimuli in the present experiment.

1.4 The Purpose of the Present Study

The purpose of this study was to test the scope of bilingual sentence planning in relation to language background and proficiency. As described in section 1.2.3, studies by Smith and Wheeldon (1999) and Allum and Wheeldon (2007, 2009) have shown that linguistic encoding is completed for the first phrase of an utterance, but that there is also some conceptual planning for much later elements in the utterance. These studies also showed that planning scope depends on the type of sentence being produced, as sentence-initial coordinate structures take longer to process and produce than simple noun phrases or prepositional structures. Furthermore, there is evidence from Konopka et al. (2018) that when speaking in a second language, people are more likely to employ a larger message planning scope and to delay speech onset than when they speak in their first language. Nonetheless, Konopka and colleagues found that bilinguals with a higher second language proficiency behaved more like first language speakers when speaking in their second language (see section 1.3.2).

As previously mentioned, there is much that is yet unknown regarding sentence planning in a second language. Inspired by the abovementioned studies, the present study endeavours to shed light on aspects of bilingual production around which little research has been done compared to the research on bilingual control mechanisms and individual word production (e.g., Konopka et al., 2018). Building on the findings of Konopka and colleagues (2018) the present study aimed to further examine how planning scope differs between L1 and L2 sentence production, and in what ways bilingual speech production in the L2 is affected by their linguistic background and experience. Furthermore, this thesis aimed to investigate whether the fact that planning scope varies depending on sentence type, as shown by Smith and Wheeldon (1999) and Allum and Wheeldon (2007, 2009), is also true for bilingual L2 language production. A novel addition in this respect is the aim to discover whether adding complexity in the form of adjective modifiers to subject position noun phrases affects planning scope in either language, and if so, whether complexity effects differ between the languages.

Due to the high saturation of English both in informal channels, such as TV and the general media, and in formal educational settings, Norwegian-English bilinguals can be expected to have a generally high level of English proficiency. Indeed, in 2020, Norway ranked in fifth place in the EF English Proficiency Index ranking of English proficiency by country (EF Education First, 2020). As such, Norwegian bilinguals are aptly suited for participation in a study of bilingual sentence production. Hence, this study was designed to test the sentence production process of Norwegian-English bilinguals based on the following experimental questions:

1. How does bilingual profile and proficiency level affect performance in L2 sentence production?
2. How does planning scope differ between L1 (Norwegian) and L2 (English)?

3. Does adding sentence complexity in the form of adjectives to the noun phrases affect scope of planning in the L1 and/or the L2, and are there between-language differences in complexity effects?

There were two main components to the study: a questionnaire and a laboratory experiment. In order to gain understanding of each participant's language background and proficiency level, the participants were asked to fill out a questionnaire detailing their language experience and second language acquisition, which was used to determine the participant's bilingual profile (see section 2.1.2). Questionnaire data would be subjected to a factor analysis in order to extract common factors concerning linguistic background and proficiency level that could be related to performance in the experimental tasks. The factor analysis would thus allow for the assessment of how different aspects of participants' profile may affect processing scope in the L2 as compared to the L1. Of particular interest was what aspects of profile might feed into rapid and accurate structure building in the L2 as compared to the L1. The questionnaire used in this experiment was an adapted version of the Language Experience and Proficiency Questionnaire, or LEAP-Q, developed by Marian, Blumenfeld and Kaushanskaya (2007). The LEAP-Q was created in response to the need for a reliable method of determining bilingual language status in which there was a predictable consistency between self-reported proficiency and first and second language performance. The reliability and efficiency of the LEAP-Q has been established on the basis of two quantitative studies of multilinguals and bilinguals (Marian et al., 2007). The LEAP-Q is thus a solid foundation on which to build a bilingual language profile for use in this experiment.

The experiment, conducted in both Norwegian and English, was a picture description task inspired by the picture grid paradigm developed by Brown-Schmidt (detailed in section 1.2.1) in which complexity effects could be added in the form of contrastive pictures. Two changes were made from the original paradigm to the picture grid used for the present study. Firstly, while the picture grids used in Brown-Schmidt and colleagues' studies (e.g. Brown-Schmidt & Konopka, 2014) had up to 20 images within one grid, the grid used in the present experiment was simplified to include only four images. Second, in the original picture grid paradigm, the marked images would flash or shift; however, the grid used in the present study had no movement, and the images to be described were instead framed in red squares (see section 2.1.3). These changes were made for the sake of clarity and in order to restrict the levels of complexity to only what was most relevant, and to avoid unnecessary disfluencies and longer speech onsets due to participants not being able to locate the contrasting image.

Based on the study's aims and experimental questions, the general predictions for this study were, firstly, that the higher the L2 proficiency level, the more L1-like a person's performance in the L2 experimental tasks will be. Thus, such aspects of profile as age of acquisition, usage rates and vocabulary were expected to play an important role in a person's level of proficiency in the L2. Second, although speech onset for L2 production can be expected to be longer compared to L1 production due to increased cognitive processing load and a larger planning scope, L2 planning scope will be similar to L1 planning scope in that sentence-initial coordinate structures will take longer to produce than prepositional structures. The third general prediction was that adding modifiers to either noun phrase in a coordinate structure will either a) break down planning scope into smaller units, in which case we will see faster speech onsets in coordinate structures than in prepositional structures, or b) make no difference to scope, in which case we should see a consistent increase in speech onset latency according to the rise in sentence complexity level.

2 EXPERIMENT AND QUESTIONNAIRE

A picture description experiment was devised aiming to elicit the articulation of different sentence types, which could be modified by adding adjectives so as to manipulate whole sentence complexity or first phrase complexity, thus making it possible to investigate variations in processing scope by measuring speech onset latency. In this way it would be possible to examine differences in speech onset time as an indicator of sentence complexity for each sentence type. The longer the speech onset, the more complex the sentence. Three pairs of sentence types were generated based on complexity effects, as shown in Figure 2:

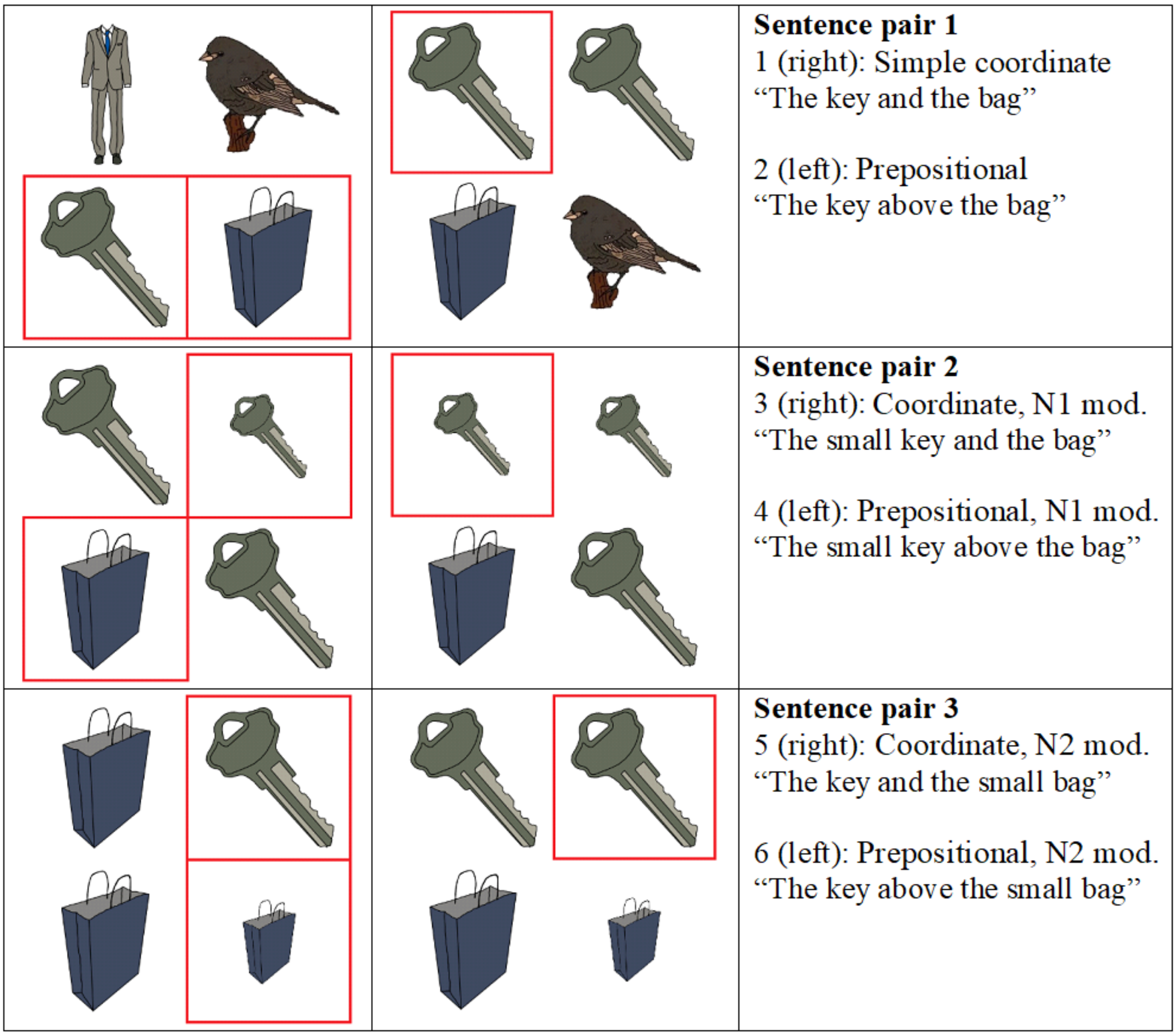


Figure 2

Examples of experimental stimuli and sentence types

Sentence pair 1:

1. Simple coordinate E.g.: The key and the bag
2. Prepositional: E.g.: The key above the bag

These first two sentences (Figure 2, top row) are control sentences used for comparison of scope manipulation, as both are unmodified by adjectives. The first sentence type is a simple coordinate structure. The second sentence type consists of a prepositional structure, which could occur with either the preposition ‘above’ or ‘below.’ Based on previous work by Smith and Wheeldon (1999) and Allum and Wheeldon (2007, 2009), the prediction was that sentence type 2 would have a faster speech onset than sentence type 1 because a coordinate structure will require longer processing time than a prepositional structure, as the lemma of both heads need to be accessed before speech onset.

Sentence pair 2:

3. Coordinate, N1 modified: E.g.: The small key and the bag
4. Prepositional, N1 modified: E.g.: The small key above the bag

In the second sentence pair (Figure 2, middle row) an adjective has been added to qualify the first noun phrase (N1) in both sentences. The aim was to investigate whether and in what ways this added complexity affected processing scope in these two sentence types as compared to each other and to the two control sentences. If adding a modifier to the first head noun affects planning scope, such that only the first noun phrase is prioritised for grammatical encoding, there should be a smaller complexity difference between sentence type 3 and 4, and sentence type 3 should have a shorter speech onset latency (SOL) than sentence type 1. If adding a modifier to the first head noun makes no difference to scope, there should be a larger complexity difference between sentence type 3 and 4, and sentence type 3 will have a longer SOL than sentence type 1.

Sentence pair 3:

5. Coordinate, N2 modified: E.g.: The key and the small bag
6. Prepositional, N2 modified: E.g.: The key above the small bag

An adjective was added to the second noun phrase (N2) of each sentence in the third sentence pair (Figure 2, bottom row). Adding complexity to the N2 made it possible to examine if and how processing scope was affected as compared to the second sentence pair, in which the N1 was modified. If adding a modifier to the N2 breaks up planning scope to include only the first noun phrase, there should be a smaller complexity difference between sentence type 5 and 6, and sentence type 5 will have a shorter SOL than sentence type 1. If adding a modifier to the second noun has no effect on scope, there should be a larger complexity difference between sentence type 5 and 6, and sentence type 5 will have a longer SOL than sentence type 1 and 3.

The sentences used for this experiment were simple descriptive utterances without a verb, so as to keep any complexity effects confined to the noun phrases. Verbs introduce the possibility of a language-specific complexity effect in that English has subject-verb agreement, which Norwegian does not, which would lead to a variation in verb form from coordinate structures to prepositional structures in English, but not in Norwegian. For example, a full English prepositional sentence would be “the key above the small bag is framed,” but a coordinate sentence would be “the key and the small bag are framed.” Additionally, having to inflect the verb in concord with the subject may lead to subject-verb agreement errors in the L2. The sentences were thus designed to answer the prompt "Uniquely identify the framed pictures" in the least complex manner.

2.1 Method

2.1.1 Participants

27 participants between the ages of 18 and 34, of whom 20 were female, took part in the production experiment. All of the participants were bilinguals who spoke Norwegian as their first language (L1) and English as their second language (L2), with no other home languages. The participants received a 300 NOK gift card at the university bookstore as compensation for participating.

2.1.2 Bilingual Profiling Questionnaire

Before taking part in the experiment, participants were asked to fill out a Bilingual Profiling Questionnaire (henceforth referred to as the BPQ) detailing their language experience and the acquisition of their L2. The BPQ was used to determine each participant's bilingual profile, including their age of acquisition, vocabulary, usage rates and general proficiency.

As described in section 1.4, the questionnaire used in this experiment was an adapted version of the LEAP-Q by Marian et al. (2007). However, a number of changes were made from the LEAP-Q to the BPQ used in the present study. Importantly, the BPQ was specially adapted for Norwegian L1 speakers, with questions ascertaining that Norwegian is their first language and questions specifically pertaining to their use of Norwegian as their L1. The order of the various parts of the BPQ also differed from that of the LEAP-Q. The LEAP-Q starts with language background and then proceeds to questions about US immigration, which is irrelevant to the present experiment, followed by level of education and subsequently more language questions. In the BPQ, all of the formalities such as personal information and criteria for participating were contained in Part 1, ‘Screening questions’, while language background was detailed in Part 2 (see Appendix A). There was also an added Part 3 that specifically addressed Norwegian and English proficiency. The question of whether the participants considered themselves a “reasonably good speaker of English” was also included at the very beginning of the questionnaire (Appendix A, Part 1, question 5). The order presented in the BPQ made it easy to ascertain that all criteria were met at the very start, before proceeding to the language questions. Thus, the changes in structure made for a more time-efficient questionnaire, tailored to the specific needs of the present experiment.

The BPQ was also more detailed than the LEAP-Q, with added questions about accidental and intentional language mixing both from English into Norwegian and vice versa (Appendix A, Part 3, Q7 & Q8), a question about which language was used for different mental and linguistic tasks (Appendix A, Part 2, Q9), and a question about whether there had been a decline in fluency in either language (Appendix A, Part 2, Q8). The questions about proficiency level were also much more comprehensive. While the LEAP-Q only details speaking, spoken language comprehension and reading comprehension (Marian et al., 2007), the BPQ additionally included pronunciation, writing, grammar, vocabulary and spelling (Appendix A, Part 3, Q4). A question about percentage of time spent speaking each language in total was also added (Appendix A, Part 2, Q4), whereas the LEAP-Q only includes reading and speaking with other bilinguals (Marian et al., 2007). Another example of the BPQ’s higher level of detail than the original LEAP-Q is the added questions about vision, hearing or language impairments, as well as right- or left-handedness (Appendix A, Part 1, questions 6-9). There was also a question about whether the participant had taken part in experiments at the linguistics lab at the University of Agder before (Appendix A, Part 1, question 14), which if so, should facilitate preparing the participant for testing. In contrast to the LEAP-Q, the BPQ was made to be anonymous, including only a participant number and no name or date of birth that could be used to identify the participant. The BPQ also included questions about country of birth and country of residence (Appendix A, Part 1, questions 10-11), which are important aspects of a person's bilingual profile. Last, but not least, the BPQ added a 'non-binary' option under "Gender", for which the LEAP-Q only had the two options of 'male' and 'female' (Appendix A, Part 1, question 2; Marian et al., 2007). This addition made the BPQ more inclusive in terms of gender identity than the original LEAP-Q. Altogether, the amended version of the LEAP-Q used in this experiment aimed to be more specialised, efficient and comprehensive than the original, while still retaining the reliability and validity of the original LEAP-Q.

2.1.3 Production Experiment Stimuli

The design used for the trial stimuli was a picture grid displaying four drawn pictures, in which one or two pictures were marked by a red frame, as illustrated in Figure 2 (see also Appendix B). The stimuli consisted of a total of 80 pictures and their corresponding names, paired to make 40 items, each item occurring in six conditions according to sentence type as shown in Figure 2 (see also section 2.1.4). The pictures used for the experiment were selected from MultiPic, a standardised set of 750 coloured line drawings constructed by Duñabeitia et al. (2018) for use in cognitive experiments. The pictures in this databank are all drawn by the same artist in the same style, are matched for visual complexity, and represent common concrete and easily nameable concepts across several different European languages. The chosen pictures corresponded to nouns with high name agreement in both languages. All trial picture names were non-cognates that did not begin with the same initial sound in Norwegian and English.

Some of the pictures were modified for use in the present experiment for reasons of saliency. For example, the colours used in certain pictures were very vivid, making the picture stand out more than the others within the same grid. These pictures were therefore modified to a duller hue. In other cases, the picture had too little colour compared to the other pictures, in which case colour was filled in to even out the overall visual complexity of the picture grid. The modified version was used consistently throughout the experiment whenever such pictures occurred in any picture grid, regardless of whether or not the picture in question was an experimental item.

Furthermore, the experimental item pictures each had a version that was modified such that it would correspond to one of the following five adjectives used to add complexity effects to the noun phrases in the experimental items: ‘big’, ‘small’, ‘sparkly’, ‘yellow’ and ‘pink’ (see Figure 3). Like the picture names, all adjectives are non-cognates that do not start with the same initial sound in English and Norwegian. For the ‘big’ and ‘small’ effects, the pictures were made significantly larger or smaller than the other pictures of the same type within that picture grid. For the ‘sparkly’ effect, yellow sparkles in a line art drawing style

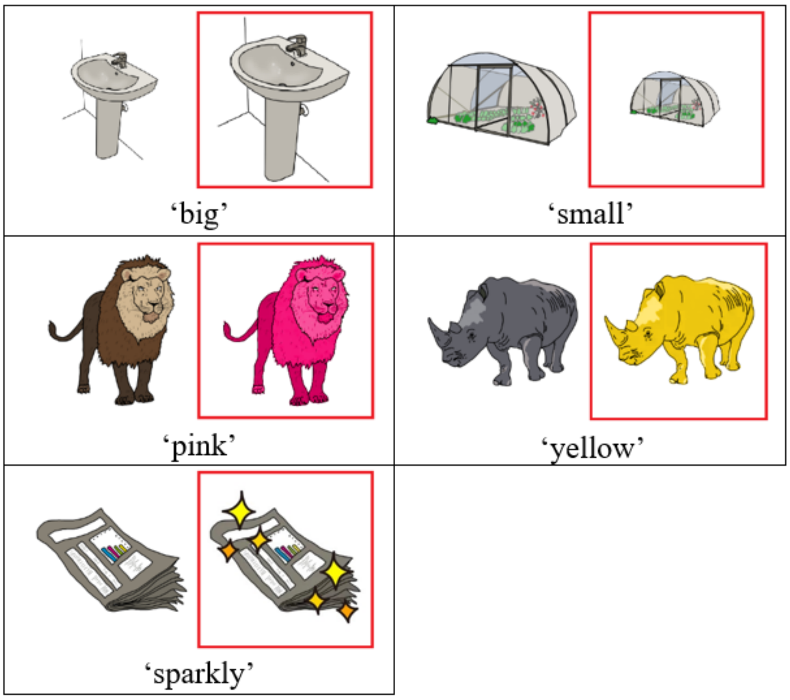


Figure 3

Examples of contrasts

were added to the picture using the LINE Camera picture editing app for smart phones. For the ‘yellow’ and ‘pink’ effects, the hue of the picture was changed to yellow or pink in stark contrast to the ‘normal-coloured’ pictures of the same type within the picture grid. The adjective used for each noun was carefully selected based on the corresponding picture to make sure the modification was clearly and contrastively prominent in relation to the other pictures of the same type. Each adjective modification occurred in four of the 20 items in each experimental list, such that all adjectives occurred an equal number of times throughout each list.

2.1.4 Production Experiment Designs

There were six conditions corresponding to sentence types, in which scope was manipulated using phrase type with two levels (coordinate/prepositional) and/or complexity (no modification/N1 modification/N2 modification), as shown in Figure 2. All conditions were within subject and within item. In other to elicit the required sentence types, the participant was instructed to uniquely identify the framed picture(s) of the trial stimuli. The layout of the pictures within the grids guided the participant to produce the various sentence types (see Figure 2; see also Appendix B). The participants were instructed to read the picture names in the order top-to-bottom, left-to-right, as this is the order in which people are taught to read in both English-speaking and Norwegian cultures. Complexity effects were added to the highlighted pictures by the inclusion of similar pictures within the grid, such that adjectives and/or preposition phrases were necessary in order to identify the target picture (see Figure 2; Appendix B). The layout of the highlighted pictures (see Appendix B) varied both within and between trial items, as well as within block, and each layout occurred approximately an equal number of times in total to ensure that participants would not detect a pattern and start developing strategies based on image layout.

Table 1

Matched numbers for length and frequency across languages (sample)

Word EN	Word NO	Syl. length EN	Syl. length NO	Phon. length EN	Phon. length NO	EN frequency	NO frequency
car	bil	1	1	3.2	3	5.44	4.97
city	by	2	1	4	2	5.4	5.35
road	vei	1	1	3	2	5.28	5.24
dog	hund	1	1	3	3	5.17	4.87
bed	seng	1	1	3	3	5.12	4.01
key	nøkkel	1	2	2	5	5.07	4.14
butterfly	sommerfugl	3	3	8	8	4.03	3.97
corner	hjørne	2	2	4.6	4.5	4.96	4.02
blanket	teppe	2	2	7	4	3.93	3.59
bag	pose	1	2	3	4	4.89	4.08
leg	bein	1	1	3	3	4.88	4.48
cabin	hytte	2	2	5	4	3.9	4.23
brain	hjerne	1	2	4	4.5	4.84	3.74
farm	gård	1	1	3.4	3	4.81	4.33
cheese	ost	1	1	3	3	4.81	3.64
bird	fugl	1	1	3.4	3	4.85	4.03

Two experimental lists were constructed that each had two different versions, English and Norwegian, each consisting of 20 items (see Appendix C). Within item, each of the six sentence types occurred once. Each of the 20 items consisted of a pair of picture names exclusive to that item. There was no phonological or semantic relationship between the picture names within item, so as to avoid potential priming effects. To control for potential word length and frequency effects, noun pairs were matched for length and frequency, as far as possible, within and across languages (see Table 1). Word frequency norms for British English were collected from SUBTLEX-UK, developed by van Heuven, Mandera, Keuleers and Brysbaert (2006). Norwegian word frequency norms were found in the 'Norwegian Web as Corpus' database (NoWaC; Guevara, 2010). Both of the official Norwegian written standards, *Bokmål* and *Nynorsk*, were collected and combined to find the total frequency of usage regardless of written standard. For ease of comparison, both word frequency lists were converted into Zipf-values, in which 1-3 represents low frequency words, and 4-7 represents high frequency words (Heuven et al., 2006). See Appendix D for a complete list of picture names used in experimental trials.

Within item, picture names were matched for word length within language and ideally also across languages, i.e. 'sheep' and 'map' in English and '*sau*' and '*kart*' in Norwegian (see Table 2 below). However, it was unavoidable that, as the nouns produced would always be definite, and since Norwegian marks definiteness by a suffix, the Norwegian word was in many cases one syllable longer than the English word, e.g. 'the umbrella' = '*paraplyen*.' Each item was assigned one of five adjectives, described above in section 2.1.3, each of which occurred four times across items within a list (see Table 2). The 20 items within an experimental list, each occurring in six conditions, totalled 120 experimental trials. When 40 filler trials were added (see section 2.1.5), each experiment had a total of 160 trials.

The two experimental lists were matched for word animacy, word length and word frequency, and were completed in each language, Norwegian and English, by sentence type (six conditions, as described in section 2) within participants and within items (see Appendix C). Furthermore, the lists were organised so that the English List 1 was identical to the Norwegian List 2 in terms of stimuli, fillers and practice block items, and the English List 2 was likewise the same as the Norwegian List 1, thereby ensuring that a participant could be assigned either List 1 or List 2 across languages and be exposed to novel stimuli in both lists (see section 2.1.7).

Each experimental list was divided into six blocks of 26 trials, including fillers. Within block, each item occurred once, and each condition occurred three to four times (see Appendix C). Within block, trials were pseudo-randomised such that the same sentence type never occurred on consecutive trials, and the initial phrase type, such as a simple noun phrase, did not occur more than twice in a row. Prepositional sentence types occurred equally often with the preposition 'above' as with 'below.' The same adjective did not appear in the

Table 2

Example of list construction (English List 1)

	English noun pair	Norwegian noun pair	EN adj.	NO adj.	EN phrase	NO phrase
1	city - corner	by (m)* - hjørne (n)	pink	rosa	the pink city / the pink corner	den rosa byen / det rosa hjørnet
2	key - bag	nøkkel (m) - pose (m)	small	liten	the small key / the small bag	den lille nøkkelen / den lille posen
3	pharmacy - wheelbarrow	apotek (n) - trillebår (m/f)	big	stor	the big pharmacy / the big wheelbarrow	det store apoteket / den store trillebåren
4	hippo - apron	flodhest (m) - forkle (n)	yellow	gul	the yellow hippo / the yellow apron	den gule flodhesten / det gule forkleet
5	frame - wave	ramme (f) - bølge (m)	sparkly	glitrende	the sparkly frame / the sparkly wave	den glitrende ramma / den glitrende bølgen
6	bottle - pocket	flaske (f) - lomme (f)	big	stor	the big bottle / the big pocket	den store flaska / den store lomma
7	cone - shrimp	kjegle (m) - reke (f)	sparkly	glitrende	the sparkly cone / the sparkly shrimp	den glitrende kjeglen / den glitrende reka
8	grape - chin	drue (m/f) - hake (f)	small	liten	the small grape / the small chin	den lille druen / den lille haka
9	lemon - trophy	sitron (m) - pokal (m)	big	stor	the big lemon / the big trophy	den store sitronen / den store pokalen
10	turkey - carrot	kalkun (m) - gulrot (f)	pink	rosa	the pink turkey / the pink carrot	den rosa kalkunen / den rosa gulrota
11	desert - hunter	ørken (m) - jeger (m)	small	liten	the small desert / the small hunter	den lille ørkenen / den lille jegeren
12	rabbit - barrel	kanin (m) - tønne (f)	yellow	gul	the yellow rabbit / the yellow barrel	den gule kaninen / den gule tønna
13	arrow - river	pil (f) - elv (f)	pink	rosa	the pink arrow / the pink river /	den rosa pila / den rosa elva
14	suit - farm	dress (m) - gård (m)	sparkly	glitrende	the sparkly suit / the sparkly farm	den glitrende dressen / den glitrende gården
15	cloud - fox	sky (m) - rev (m)	yellow	gul	the yellow cloud / the yellow fox	den gule skyen / den gule reven
16	tyre - beak	dekk (n) - nebb (n)	big	stor	the big tyre / the big beak	det store dekket / det store nebbet
17	branch - steak	grein (m) - biff (m)	pink	rosa	the pink branch / the pink steak	den rosa greinen / den rosa biffen
18	tray - shark	brett (n) - hai (m)	small	liten	the small tray / the small shark	det lille brettet / den lille haien
19	spoon - pond	skje (m/f) - dam (m)	sparkly	glitrende	the sparkly spoon / the sparkly pond	den glitrende skjea/ den glitrende dammen
20	duck - smoke	and (f) - røyk (m)	yellow	yellow	the yellow duck / the yellow smoke	den gule anda / den gule røyken

*Letters in parentheses denote gender: (m) - masculine, (f) - feminine, (n) - neuter same position, i.e. qualifying first or second noun, on consecutive trials. Additionally, the visual layouts varied as much as possible within block, and the same visual layout never occurred on consecutive trials within block. These constraints aimed to ascertain syntactic and visual variation, while simultaneously making the order of items feel random for the participant, so as to prevent pattern recognition and strategy development. The order of the blocks within the list was counterbalanced and rotated between participants.

2.1.5 Fillers

In order to break up the syntax of the spoken sentences and thus avoid repetition priming, 40 filler trials were distributed pseudo-randomly across each experimental list (see Appendix C). Each of the six blocks contained two of each filler type, and there were three types of filler sentences:

1. There are no pictures

   - A grid of four red frames that are all empty of pictures.

2. All the pictures are the same

   - A grid of four identical pictures

3. There are three 'picture name'; e.g. "there are three candles"

   - A grid in which three identical pictures are all displayed within the same angled red box frame

As a general rule, the filler sentences were never more complex than the coordinate sentences. To avoid priming effects, the filler sentences did not contain nouns that were used as experimental items in that list. See Appendix B for visual stimuli examples.

2.1.6 Procedure

Half of the participants were assigned to List 1 and half to List 2. Each participant came to the experimental linguistics lab at the university for testing on two separate days, each day testing one language only. Thus, participants completed one experimental list in Norwegian on one day and another with a different set of stimuli in English on another day, making a total of 40 items for each participant throughout the two days of testing. The reason why all participants started with a Norwegian list was to avoid the dominant language inhibition effect. This is the phenomenon by which the first language, which is usually the dominant language, is suppressed or inhibited by the use of the second language, resulting in slower L1 processing (Meuter & Allport, 1999). To ensure natural language use, participants were free to speak in their own dialect during the Norwegian trials, and for the English trials, no constraints were placed on accent. The assignment of lists to language was counterbalanced across participants.

Before coming in for the experiment, participants were sent a document in the language to be tested, containing clear instructions and examples of sentence types they would be required to produce (see Appendix E). Participants were asked to read these instructions carefully before arriving for testing. As this experiment did not seek to investigate lexical retrieval, it was important that participants were familiar with all of the picture names and adjectives to be produced to ensure that reaction times reflected syntactic processing and not lexical retrieval difficulties. Therefore, upon arriving for testing, participants were instructed to thoroughly examine a laminated booklet containing lists of picture names and adjectives to be used during testing, as well as a repetition of the examples from the instruction document, so that they would be comfortable with the experimental tasks and picture names, avoiding unnecessary speech onset delays due to confusion over the experimental task.

The testing was conducted individually for each participant inside a sound-attenuated booth within the experimental linguistics lab at the university with a window through which the experimenter and the participant could see each other. The experimenter was seated in front of two monitors, one of which showed a real-time display of the participant's monitor, and the other on which experiment progress and information about block, trial, reaction times and error codes were displayed. The Presentation software was used to run the experiment. Communication between experimenter and participant was made possible through an intercom. The experimenter was equipped with a Sennheiser GSP 350 headset. Inside the booth, Creative SBS270 speakers conveyed the experimenter's voice, and a Røde VideoMic NTG microphone captured the participants' responses, which were recorded to allow for further examination and analyses if needed. All interactions between experimenter and participant remained in the language of testing, so as not to inhibit the target language by use of the other language.

Inside the testing booth, the participant sat in front of a computer monitor and a keyboard. At the start of the experiment, the computer monitor displayed the most important instructions again to ensure that the participant knew the procedure. A ready check followed, requiring a key press from the experimenter when the participant had confirmed she or he was ready to begin.

Each trial began by the appearance of a fixation cross in the centre of the screen for 500 ms, accompanied by a low beep, followed by a 500 ms blank screen before the stimulus appeared on the screen and the voice key timer started. This timer automatically registered reaction time errors and deemed any reaction times of more than 3000 ms invalid. Upon speech onset, a voice key was triggered and started recording until the participant had finished speaking. The stimulus remained on-screen until the experimenter pressed a key code for correct trial or a key code specifying an error. Upon the key press from the experimenter, the next trial would begin.

The experimental blocks were preceded by two blocks of 18 practice trials to allow the participants to familiarise themselves with the procedure. The practice trials were identical in design to the experimental trials, but were made up of non-experimental items and fillers. Breaks were inserted between practice blocks, between the practice blocks and the experiment proper, and between experimental blocks. These breaks gave the participant the chance for a short rest and made possible communication between experimenter and participant, and required a key press from the experimenter before the next block could begin. Including breaks, experiment duration varied from approximately 20 to 30 minutes.

3 RESULTS

3.1 Results from Questionnaire Data

3.1.1 Participants

The participants for the questionnaire study were all bilinguals aged 18 to 34. Fifty-one participants (thirty-six females) filled out the BPQ; however, only thirty-seven of them took part in the production experiment for this study. The education level of the participants varied from upper secondary (one participant) to completed Master's degree (M = 16.79 years of education, range = 12.5 to 23 years). No participants reported any hearing, learning or language disability, and all had normal or corrected-to-normal vision. All but three of the participants were right-handed. Participants completed the questionnaire independently prior to the experiment.

3.1.2 Language Background Usage

All participants spoke Norwegian as their L1 except for three, whose L1 was English with Norwegian as their L2. One participant was born in the USA, the rest were born in Norway. All participants were Norwegian residents. Of the participants, eighteen spoke a third language, and one spoke a fourth language. All but four reported Norwegian as their most dominant language. Three reported English to be the most dominant, and one did not report a dominant language. All participants reported having acquired Norwegian first, followed by English, except for one participant who acquired English first and Norwegian subsequently. All participants identified mainly with Norwegian culture except two, who identified mainly with British and American cultures, respectively. When asked if they had become less fluent in one of their languages, fifteen participants reported becoming less fluent in English, and six reported a loss of fluency in Norwegian. When asked which language they used for different tasks, two participants reported using English for simple maths, seven used English when dreaming, fourteen used English when expressing anger or affection, and sixteen used English when talking to themselves. Forty-four participants reported accidentally mixing words or sentences from Norwegian and English, and all but five of the participants reported intentionally mixing words or sentences from Norwegian and English (see Table 3 below for variable means and ranges).

Table 3

Participants' background and current language use

Background and language use					
	M	Range		M	Range
Participants' age (years)	25.18	18-34	Language choice: Norwegian (%)	82.22	10-100
			Language choice: English (%)	17.37	0-90
Years of education	16.79	12.5-23			
			Accidental L2 mixing into L1[1]	3.67	0-8
Exposure to L1 (%)	59.64	25-95	Accidental L1 mixing into L2[1]	1.73	0-7
Exposure to L2 (%)	39.18	5-75			
			Intentional L2 mixing into L1[1]	4.41	0-10
Speaking L1 (%)	77.8	20-100	Intentional L1 mixing into L2[1]	2.08	0-9
Speaking L2 (%)	21.18	0-80			
Reading L1 (%)	52.59	10-99			
Reading L2 (%)	46.75	1-90			

[1]Range: 0 = never, 5 = half of the time, 10 = all of the time

3.1.3 Language Exposure and Learning

As shown in Table 4, exposure to English was mostly in the context of TV or streaming, music and other media, and reading, while exposure to Norwegian was mostly in the context of family, friends and colleagues, and reading. The highest contributors to the learning of English were education and reading, followed by TV or streaming, music and media and interactions with friends or colleagues (see Table 4 below for variable means and ranges).

Table 4

Language exposure and learning

Language: Norwegian	M	Range	Language: English	M	Range
Immersion duration: (years:months)			Immersion duration: (years:months)		
In a country	24:6	17:9-32:7	In a country	1:6	0:0-17:3
In a family	25:3	18:8-34:11	In a family	1:11	0:0-34:11
In a school (all of the time)	14:4	0:0-27:0	In a school (all of the time)	0:9	0:0-13:0
In a school (some of the time)	5:3	0:0-27:0	In a school (some of the time)	6:9	0:0-18:1
In a workplace (all of the time)	5:1	0:0-14:0	In a workplace (all of the time)	0:1	0:0-3:0
In a workplace (some of the time)	1:11	0:0-12:11	In a workplace (some of the time)	1:5	0:0-6:0
Contribution to language learning[1]			Contribution to language learning[1]		
Friends/colleagues	7.75	0-10	Friends/colleagues	6.08	0-10
Family	9.33	5-10	Family	2.49	0-10
Reading	6.84	0-10	Reading	7.20	2-10
Education	7.53	2-10	Education	7.57	0-10
Self-instruction	1.27	0-10	Self-instruction	2.71	0-10
TV/streaming	3.92	0-10	TV/streaming	7.55	2-10
Music/media	3.25	0-10	Music/media	6.37	0-10
Current exposure to language[2]			Current exposure to language[2]		
Friends	9.14	4-10	Friends	3.76	0-10
Family	9.35	0-10	Family	1.10	0-9
Reading	5.31	0-10	Reading	6.94	1-10
Self-instruction	1.12	0-10	Self-instruction	1.69	0-10
TV/streaming	3.41	0-10	TV/streaming	8.14	4-10
Music/media	3.33	0-10	Music/media	8.00	4-10
Age milestones (years)			Age milestones (years)		
Started hearing	0.12	0-3	Started hearing	7.06	0-14
Attained speaking fluency	4.49	2-16	Attained speaking fluency	13.9	6-23
Started reading	5.18	3-8	Started reading	7.84	5-13
Attained reading fluency	8.14	5-19	Attained reading fluency	13.24	7-22

[1]Range: 0 = not a contributor, 5 = moderate contributor, 10 = most important contributor

[2]Range: 0 = never, 5 = half of the time, 10 = almost always

3.1.4 Proficiency Measures

Participants' self-reported proficiency ratings in English were generally high. As Table 5 (below) shows, on a scale from 0-10, the lowest average self-reported measure for all sub-categories was 6.92. Participants reported the highest proficiency ratings within the categories of understanding and reading, while grammar and spelling had the lowest ratings. In Norwegian, self-reported proficiency level was high in all sub-categories, the lowest ratings falling within the categories of grammar, vocabulary and spelling (see Table 5).

Table 5

Self-reported proficiency measures

Language: Norwegian	M	Range	Language: English	M	Range
Self-reported proficiency level[1]			Self-reported proficiency level[1]		
Speaking	9.53	5-10	Speaking	7.78	4-10
Pronunciation	9.47	6-10	Pronunciation	7.02	2-10
Understanding	9.76	6-10	Understanding	8.49	6-10
Reading	9.55	3-10	Reading	8.31	3-10
Writing	9.12	5-10	Writing	7.76	4-10
Grammar	8.43	4-10	Grammar	6.92	3-10
Vocabulary	8.67	6-10	Vocabulary	7.18	4-10
Spelling	8.69	4-10	Spelling	6.98	3-10

[1]Range: 0 = none, 5 = adequate, 10 = perfect

3.1.5 Factor Analysis

The participants' responses were analysed by factor analysis so as to group variables together according to co-variance patterns, which indicate that they measure the same underlying factor. Factor analysis thus allows for the comparison of statistical clustering of questions. The data from the written answer variables, described in the above paragraph, were removed from the analysis input data, as well as data from questions with little to no variation in the answers. For example, many of the answers to questions about Norwegian were highly similar or identical due to the uniform nature of the participant set. Of the original 128 variables, 46 remained. A correlation matrix of these variables was produced. Only variables with a correlation of 0.3 or more with another variable were included in the final data set, which meant that one variable was excluded from the analysis set due to insufficient co-variation. There were also number of variables with too high a correlation at 0.9 or above, in which case the variable deemed least relevant was removed. For example, where the correlation was between L1 and L2 versions of a variable, such as Question 3 regarding time exposed to each language (Appendix A, part 2), L1 variables were removed as these were deemed less relevant than the L2 measures for the present study. For a list of variables removed due to too high or insufficient correlation, see Appendix F.

The remaining 38 variables formed the input to the factor analysis. This analysis yielded four factors, based on patterns of variables and commonalities underlying variable clusters. These factors account for close to 50 % of the variance in the data. Within these factors, positive loadings signify criteria contributing to the factor's underlying construct, while negative loadings signify criteria in opposition to the underlying construct. Factor names were composed on the basis of the general nature and commonalities of the cluster of variables within. The output of the factor analysis can be seen in Table 6 (below).

The first factor, which accounted for most of the variance, was named 'English Proficiency,' as the highest-loading variables are associated with proficiency in English, such as skills in English grammar, writing, reading, vocabulary, listening, speaking and pronunciation, as well as exposure to reading, interactions with friends or colleagues, music and media, and TV or streaming in English. This factor also includes variables concerning the contribution of reading and interacting with family to the learning of English. Interestingly, a higher proficiency in English is accompanied by a higher rate of accidental mixing of English into Norwegian (Q7 Accidental mixing: EN to NO, Table 6). This factor also includes some variables associated with age of Norwegian acquisition, such as fluency in speaking and reading, indicating that a higher age of fluency in Norwegian is related to higher English proficiency.

The second factor was interpreted as relating to spoken English, as the factor is strongly influenced by the variable 'Q4 Time spent speaking L2' (Table 6), and it also

Table 6

Factor analysis output

English Proficiency		Spoken English Proficiency		Informal Learning of English		Age of English Acquisition	
Q4 EN proficiency: grammar	0.84	Q4 Time spent speaking L2	0.68	Q2 EN learning: TV/streaming	0.74	Q5 EN reading fluency age	0.69
Q4 EN proficiency: writing	0.82	Q5 NO speaking fluency age	0.6	Q2 EN learning: music/media	0.65	Q5 EN started hearing age	0.65
Q4 EN proficiency: reading	0.82	Q5 NO reading fluency age	0.59	Q3 EN exposure: TV/streaming	0.61	Q5 EN speaking fluency age	0.62
Q4 EN proficiency: vocabulary	0.78	Q6 Choosing to speak L2	0.59	Q3 EN exposure: music/media	0.56	Q5 EN started reading age	0.61
Q4 EN proficiency: listening	0.74	Q3 Time exposed to L2	0.57	Q2 EN learning: reading	0.49	Q2 EN learning: friends	0.51
Q4 EN proficiency: speaking	0.71	Q2 EN learning: family	0.44	Q3 Time exposed to L2	0.46	Q2 NO learning: reading	0.48
Q4 EN proficiency: pronouncing	0.69	Q3 EN exposure: friends	0.41	Q6 Switching proficiency	0.44	Q2 NO learning: TV/streaming	0.47
Q3 EN exposure: reading	0.65	Q7 Accidental mixing: EN to NO	0.37	Q3 EN exposure: friends	0.35	Q7 Accidental mixing: NO to EN	-0.37
Q6 Switching proficiency	0.6	Q4 EN proficiency: vocabulary	0.33	Q2 EN learning: self-instruction	0.35	Q8 Intentional mixing: EN to NO	-0.58
Q2 EN learning: reading	0.59	Q4 EN proficiency: pronouncing	0.33	Q5 Time spent reading L2	0.31		
Q2 NO learning: school	0.43	Q5 Time spent reading L2	0.33	Q5 NO speaking fluency age	-0.3		
Q7 Accidental mixing: EN to NO	0.41	Q2 NO learning: reading	-0.37	Q5 EN reading fluency age	-0.4		
Q2 NO learning: reading	0.41	Q2 NO learning: school	-0.4				
Q5 Time spent reading L2	0.38	Q3 NO exposure: reading	-0.44				
Q5 NO speaking fluency age	0.36	Q3 NO exposure: TV/streaming	-0.71				
Q5 NO reading fluency age	0.35						
Q3 EN exposure: friends	0.35						
Q3 EN exposure: music/media	0.35						
Q2 EN learning: school	0.33						
Q6 Choosing to speak L2	0.31						
Q2 EN learning: family	0.31						
Q3 EN exposure: TV/streaming	0.31						
Proportion Var.	**0.19**	**Proportion Var.**	**0.11**	**Proportion Var.**	**0.09**	**Proportion Var.**	**0.09**
Cumulative Var.	**0.19**	**Cumulative Var.**	**0.3**	**Cumulative Var.**	**0.39**	**Cumulative Var.**	**0.47**

Note : The questions are collected from separate parts of the questionnaire, thus the same question number may be used more than once to refer to different questions. For example, ‘Q5 Time spent reading L2’ and ‘Q5 NO speaking fluency age’ are different questions from separate parts of the questionnaire.

includes variables related to vocabulary and pronunciation in English, the amount of time the L2 is chosen when speaking, and English accidentally intruding when speaking Norwegian. This factor has therefore been named ‘Spoken English Proficiency.’ Another strong variable in this factor is ‘Q5 NO speaking fluency age,’ indicating that a higher age of speaking fluency in Norwegian is associated with higher English proficiency. This factor also includes variables with negative loadings related to Norwegian education, indicating that higher educational proficiency in Norwegian is associated with poorer performance in speaking English. Altogether, this factor seems to suggest that speaking more English relates to delay in Norwegian proficiency.

In the third factor, the variables with the highest positive loadings were all related to the learning of and exposure to English via informal channels such as TV/streaming and music and media, which led to this factor being named ‘Informal Learning of English.’ Variables that also played a part in this factor were related to reading in English, time exposed to English and proficiency in switching between languages, as well as English interactions with friends and self-instruction of English. The ages of achieving speaking fluency in Norwegian and reading fluency in English had weak negative loadings on this factor, indicating that a lower age of fluency in Norwegian is associated with a higher contribution of informal channels to the learning of English, and a higher degree of exposure to informal English.

As the fourth factor included several age-related variables with large positive loadings connected to the learning of English, this factor was named ‘Age of English Acquisition.’ Other variables with sizeable positive loadings are related to the learning of English through interacting with friends, as well as variables related to the learning of Norwegian through reading and watching TV or streaming. This covariance indicates that participants who acquired English later had more contribution from friends to their learning of English, and that the later participants acquired English, the more they would read and watch TV in Norwegian. Negatively loaded variables in this factor are related to language mixing. The accidental intrusion of Norwegian into English has a weak negative loading on this factor, and the intentional mixing of English into Norwegian has a relatively large negative loading, implying that there is a relationship between age of English acquisition and language mixing. This factor suggests that higher age of fluency in reading and speaking English is associated with less accidental and intentional mixing of the languages.

3.2 Results from Experimental Data

The experimental data from twenty-seven participants provided 6480 observations. One subject with a 62 % error rate was removed from analyses. Timeout trials with zero reaction time (RT), which occurred if participants did not start speaking before the timer ran out, were removed from analyses. These removals make up 424 of the 6480 observations, or 6.53 %. Trials were coded with an operator response indicating whether the participant's response was correct or erroneous, in which case the key code specified the type of error. Most errors were disfluency errors, as these made up more than 35 % of the errors and almost 10 % of the data. Finally, seven trials coded as a stimulus error were discarded, and two additional trials coded as experimenter errors were also removed. The RT analysis was restricted to correct trials only, which amounted to 73.9% of the data. Altogether, there was minimal removal of data. See Appendix G for data loss and mean error rate by participant. The conditions for the analyses were Scope (coordinate or prepositional structures), Language (English/Norwegian) and levels of Complexity: zero complexity (Zero), first noun phrase modification (NP1) and second noun phrase modification (NP2).

3.2.1 Error Analysis

For the error analysis, a linear mixed effects model was run with a maximal structure that included random intercepts for subject (27) and item (20), as well as random slopes to account for variability between subjects in the effect of language and condition, and variability between items in the effect of condition (see Table 7 below for model syntax). Forward difference coding was conducted such that each level of the complexity condition was compared to the next. In this manner, two contrasts were coded: NP1.zero and NP1.NP2. The first contrast, NP1.zero, is the mean of NP1 minus the mean of Zero, i.e., the contrast between the likelihood of making an error on NP1 and the likelihood of making an error on Zero. The second contrast, NP2.NP1, represents the errors on NP2 relative to errors on NP1; i.e., the mean likelihood of making an error on NP2 minus the mean likelihood of making an error on NP1. To avoid collinearity, the levels of Scope and Language were centred.

Table 7 shows the output from the linear effects model. Significant variables are shown in bold text. The analysis yielded significant effects of Scope, interactions between

Table 7

Analysis on errors

	Errors			
	Estimate	Std. Error	Z value	P value
(Intercept)	-1.21	0.2	-6.05	**<0.01**
scope [coord, -0.55; prep., 0.5]	**0.26**	**0.1**	**2.55**	**0.01**
language [English, -0.5; Norwegian, 0.5]	0.1	0.2	0.51	0.61
NP2.NP1	-0.09	0.08	-1.19	0.24
NP1.zero	**0.61**	**0.08**	**7.49**	**<0.01**
scope:language	**-0.78**	**0.13**	**-5.96**	**<0.01**
scope:NP2.NP1	**-0.41**	**0.16**	**-2.63**	**0.01**
scope:NP1.zero	0.05	0.16	0.33	0.74
language:NP2.NP1	0.01	0.15	0.06	0.96
language:NP1.zero	-0.17	0.16	-1.03	0.3
scope:language:NP2.NP1	-0.44	0.31	-1.4	0.16
scope:language:NP1.zero	0.55	0.33	1.7	0.09

Scope and Language, as well as effects of Complexity in the contrasts between the adjective positions, NP1 and NP2. Figure 4 illustrates the mean error rate for each condition in each language. As this graph shows, although the pattern of errors differs from Norwegian to English, both languages were on average equally error-prone.

In Figure 4, mean error rates and standard deviation are shown for each phrase type and complexity condition in both English and Norwegian. There was a main effect of Scope; averaged across languages, participants made more errors in prepositional structures than coordinate structures. The interaction of Scope with Language is due to the difference between the phrase types across languages. As can be seen, in English the pattern for the coordinate structures versus the prepositional structures is very regular across the different complexity levels, in that participants consistently make more errors in prepositional structures than in coordinate structures. In Norwegian, the opposite is true except for in NP1 sentences. In other words, prepositional structures in Norwegian are less error-prone than coordinate structures.

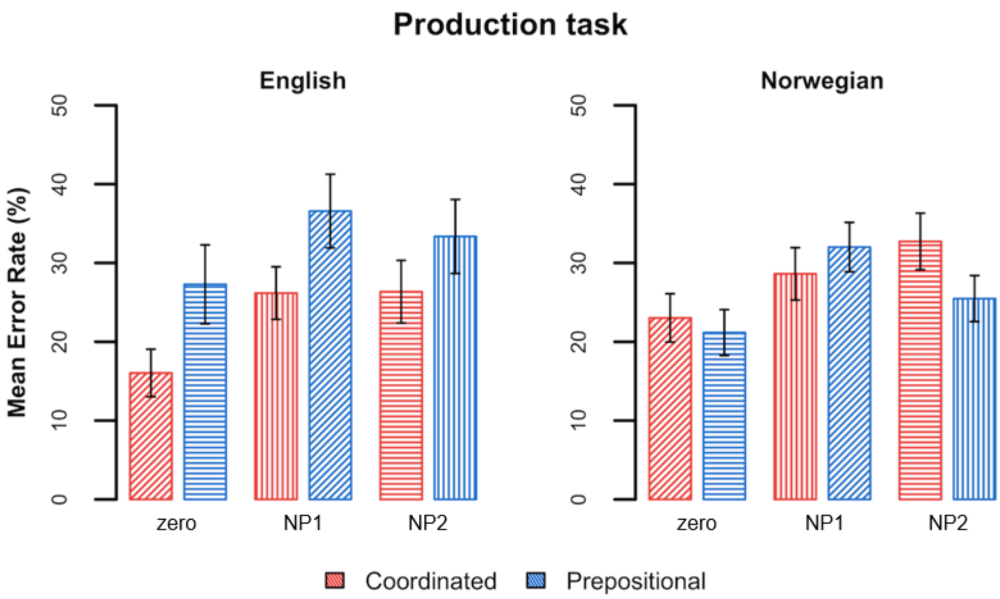


Figure 4
Error analysis

The main effect of the complexity difference between zero adjectives and first phrase adjectives, NP1.zero (Table 7), indicates that, across languages, participants made more errors in NP1 trials compared to Zero trials. The contrast between NP2 and NP1 complexity levels interacted with Scope such that coordinate structures had more errors in the NP2 condition than in the NP1 condition, while prepositional structures had more errors in the NP1 condition.

3.2.2 RT Analysis

The pattern of errors is mirrored in the speech onset latency (RT) data, and as such there is no evidence of a speed/accuracy trade-off in the data. The RT data were submitted to a linear mixed affects model analysis. Table 8 shows the output from the RT analysis.

Table 8

Analysis on speech onset latency (RT data)

	RT			
	Estimate	Std. Error	Z value	P value
(Intercept)	1337.6	44.06	30.36	< 1e-04
scope [coord, -0.55; prep., 0.5]	5.24	11	0.48	0.63
language [English, -0.5; Norwegian, 0.5]	**106.05**	**32.28**	**3.29**	**0**
NP2.NP1	**56.11**	**11.46**	**4.89**	**< 1e-04**
NP1.zero	10.48	11.15	0.94	0.35
scope:language	**-42.43**	**18.29**	**-2.32**	**0.02**
scope:NP2.NP1	-24.38	22.93	-1.06	0.29
scope:NP1.zero	18.21	22.3	0.82	0.41
language:NP2.NP1	**88.61**	**22.94**	**3.86**	**0**
language:NP1.zero	**-85.05**	**22.29**	**-3.82**	**0**
scope:language:NP2.NP1	**-171.73**	**45.87**	**-3.74**	**0**
scope:language:NP1.zero	**98.42**	**44.61**	**2.21**	**0.03**

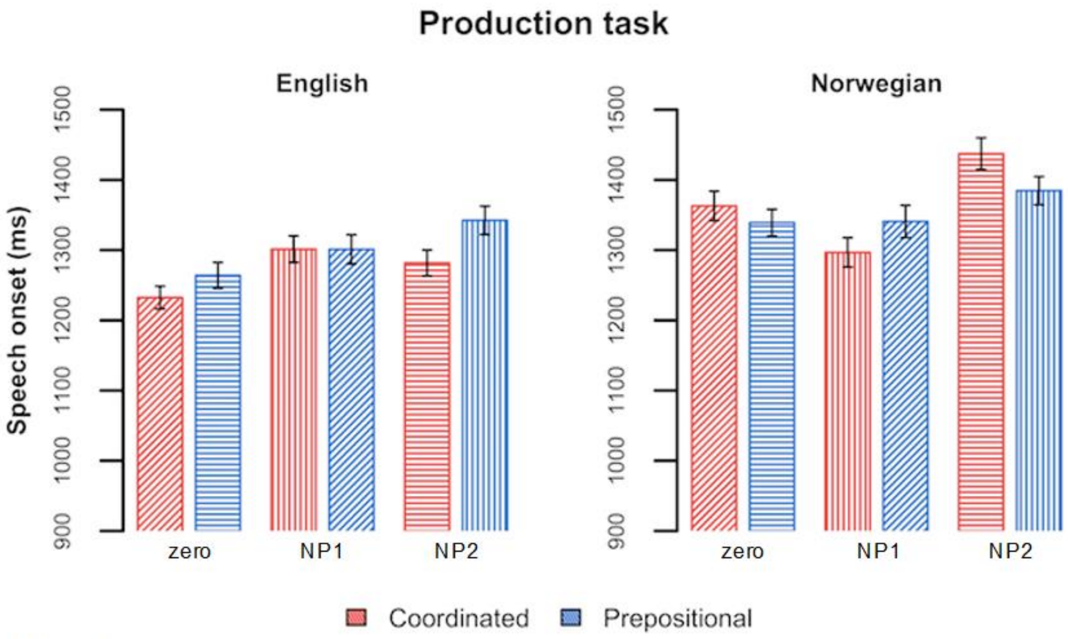


Figure 5

Speech onset latencies (RT data averages)

The RT analysis revealed several significant effects, shown in bold text in Table 8. Firstly, there is a main effect of Language. Figure 5 illustrates how RTs were considerably faster in English than in Norwegian. Furthermore, Scope interacts with Language (scope:language, Table 8), such that in English, the prepositional structures are either slower than or similar to coordinate structures, while in Norwegian coordinates take longer to initiate than prepositional structures. The exception is the NP1 condition, in which in Norwegian the prepositional structures are slower in the NP1 condition, whereas reaction times are approximately the same in English for both structures.

The difference between the two Complexity conditions, NP2.NP1, is also significant, as NP2 is slower overall than NP1 across languages. However, both NP2.NP1 and NP1.zero interact with Language and Scope. The three-way interaction of Scope, Language and NP2.NP1, is seen especially in the coordinate condition in Norwegian; i.e. the effect of increasing complexity in NP2 is seen most strongly in coordinate structures compared to prepositional structures in Norwegian, see Figure 5. The three-way interaction of Scope, Language and the contrast between Zero and NP1 (scope:language.NP1.zero, Table 8), can be observed in Norwegian, in which zero complexity sentences have longer RTs than NP1 in the coordinate condition, while the opposite is true for English.

A follow-up analysis was run in order to further examine the nature of the interactions, in which the data for English and Norwegian were analysed separately, see Table 9. This follow-up analysis showed that the interaction between Scope and the difference between NP1 and NP2 (scope:NP2.NP1) is significant in Norwegian, while in English this interaction is not significant. In Norwegian, the pattern between coordinate and prepositional structures is the opposite for NP1 and NP2 (see figure 5). Thus, averaged across languages, the difference in Complexity effects between NP1 and NP2 in the Coordinate condition will most likely either have no difference or only a slight increase due to the

Table 9

Follow-up analysis

RT - English

	Estimate	Std. Error	Z value	P value
(Intercept)	1285.25	42.72	30.09	**<0.01**
scope [coord, -0.55; prep., 0.5]	**25.68**	**12.94**	**1.99**	**0.05**
NP2.NP1	12.07	15.57	0.77	0.44
NP1.zero	**52.36**	**15.16**	**3.45**	**<0.01**
scope:NP1.zero	-29.84	30.37	-0.98	0.33
scope:NP2.NP1	**60.02**	**31.21**	**1.92**	**0.05**

RT - Norwegian

	Estimate	Std. Error	Z value	P value
(Intercept)	1393.9	51.29	27.18	**<0.01**
scope [coord, -0.55; prep., 0.5]	-16.2	21.21	-0.76	0.45
NP2.NP1	**100.59**	**16.79**	**5.99**	**<0.01**
NP1.zero	**-32.73**	**16.32**	**-2.01**	**0.04**
scope:NP1.zero	**67.71**	**32.65**	**2.07**	**0.04**
scope:NP2.NP1	**-111.03**	**33.59**	**-3.31**	**<0.01**

increase in Norwegian, as illustrated in Figure 5, whereas the prepositional condition would see a decrease in Complexity effects between NP1 and NP2 when averaged. In English, there is a main effect of Scope in which prepositional structures take longer to initiate than coordinate structures, overall, see Figure 5. In English, there was also a main effect of NP1.zero (see Table 9); i.e., NP1 sentences have longer onsets than zero complexity sentences regardless of scope.

The pattern of results for Norwegian is different. There was no main effect of Scope. However, there were significant effects of Complexity, as well as main effects of Scope interacting with Complexity. There was a main effect of NP2.NP1 in Norwegian (see Table 9), as NP2 sentences took longer to produce than NP1 sentences (see Figure 5). This contrast significantly interacted with Scope (scope:NP2.NP1, Table 9), as this effect could be seen especially in coordinate structures. The contrast NP1.zero, on the other hand, had a negative main effect, meaning that latencies were longer for Zero than for NP1 sentences. The interaction of this contrast with Scope is because the effect was due to the difference between coordinate structures rather than prepositional structures.

3.3 Effects of Individual Differences in Language Factors

In order to examine how the factors from the questionnaire analysis might relate to participants' performance in the experimental task, analyses including these factors were run on the experimental data for the two contrasts NP1.zero and NP2.NP1 in each language. Table 10 shows the output from the first of these analyses which includes the difference between NP1 and Zero in the English experiments. The pattern of effects for each factor is illustrated in the four graphs in Figure 6.

Table 10

Individual differences from factors – English: NP1.zero contrast

RT - English - NP1 - Zero				
	Estimate	Std. Error	Z value	P value
(Intercept)	1273.96	43.72	29.14	<0.01
scope [coord, -0,47; prep, 0.53]	14.07	15.19	0.93	0.35
complexity [AdjFirst, -0.53; Zer, 0.47]	**-50.22**	**15.19**	**-3.31**	**<0.01**
EngProficiency	**-84.48**	**45.49**	**-1.86**	**0.06**
SpokenEngProficiency	-35.64	37.84	-0.94	0.35
InfEngLearning	-4.3	35.77	-0.12	0.9
AoAEng	40.68	48.84	0.83	0.4
scope:complexity	29.24	30.41	0.96	0.34
scope:EngProficiency	11.23	16.21	0.69	0.49
scope:SpokenEngProficiency	-4.54	13.53	-0.34	0.74
scope:InfEngLearning	**26.96**	**13.03**	**2.07**	**0.04**
scope:AoAEng	34.75	19.77	1.76	0.08
complexity:EngProficiency	11.35	16.22	0.7	0.48
complexity:SpokenEngProficiency	-6.33	13.56	-0.47	0.64
complexity:InfEngLearning	-5.74	12.97	-0.44	0.66
complexity:AoAEng	32.56	19.51	1.67	0.1
scope:complexity:EngProficiency	45.02	32.44	1.39	0.17
scope:complexity:SpokenEngProficiency	8.04	27.14	0.3	0.77
scope:complexity:InfEngLearning	-14.05	26	-0.54	0.59
scope:complexity:AoAEng	11.88	39.28	0.3	0.76

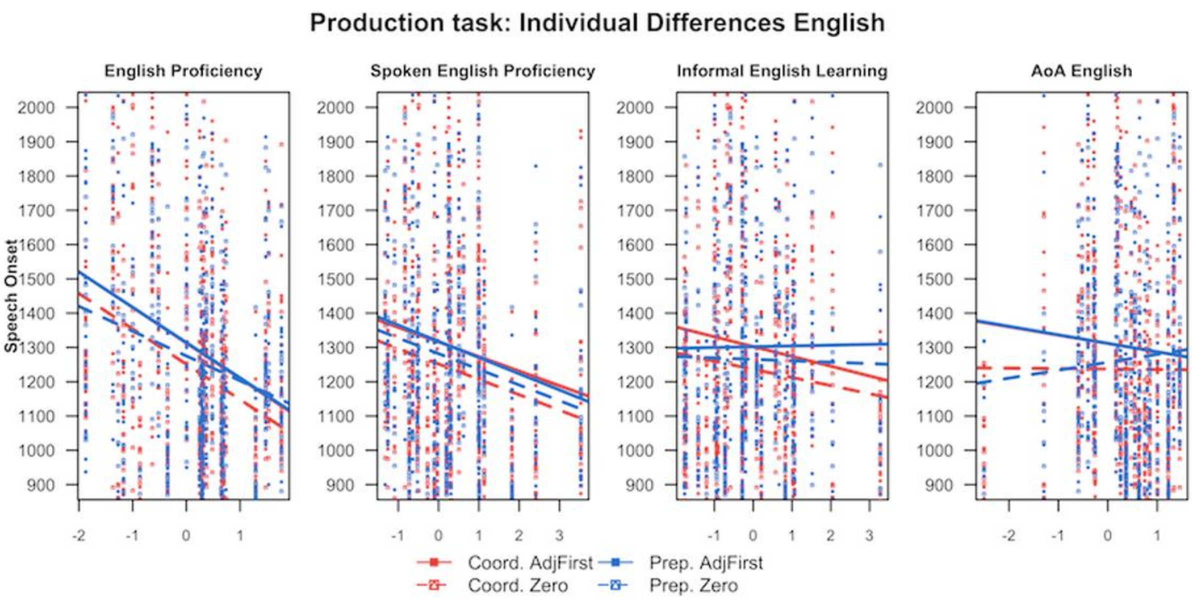


Figure 6

Individual differences from factors – English: NP1.zero contrast

As Table 10 shows, there is a main effect of the 'English Proficiency' factor that is borderline significant. As can be seen, as participants' English proficiency level increases, their RTs also get faster across conditions.[1] There is also another significant interaction of Scope with the 'Informal Learning of English' factor: as informal English learning increases, RTs to coordinate structures reduce, while RTs to prepositional structures increase slightly. Accordingly, coordinate structures are more strongly affected by the interaction between Scope and this factor.

Table 11

Individual differences from factors – English: NP2.NP1 contrast

RT - English - NP2 - NP1				
	Estimate	Std. Error	Z value	P value
(Intercept)	1307.42	45.48	28.75	<1e-04
scope [coord, -0,47; prep, 0.53]	24.77	16.17	1.53	0.13
complexity [AdjFirst, -0.5; AdjSecond 0.5]	15.85	16.09	0.99	0.33
EngProficiency	-75.95	48.35	-1.57	0.12
SpokenEngProficiency	-26.5	40.22	-0.66	0.51
InfEngLearning	-10.63	37.98	-0.28	0.78
AoAEng	34.95	51.94	0.67	0.5
scope:complexity	50.72	32.19	1.58	0.12
scope:EngProficiency	-7.1	17.4	-0.41	0.68
scope:SpokenEngProficiency	16.37	14.4	1.14	0.26
scope:InfEngLearning	11.42	13.62	0.84	0.4
scope:AoAEng	28.31	21.65	1.31	0.19
complexity:EngProficiency	**33.09**	**17.39**	**1.9**	**0.06**
complexity:SpokenEngProficiency	11.13	14.41	0.77	0.44
complexity:InfEngLearning	-18.97	13.54	-1.4	0.16
complexity:AoAEng	14.79	20.88	0.71	0.48
scope:complexity:EngProficiency	10.61	34.72	0.31	0.76
scope:complexity:SpokenEngProficiency	48.67	28.78	1.69	0.09
scope:complexity:InfEngLearning	-45.38	27.12	-1.67	0.09
scope:complexity:AoAEng	-12.15	42.18	-0.29	0.77

[1] In Figure 6, although only the blue line is visible, the blue line representing NP1 sentences with prepositional structures, and the red line representing NP1 sentences with coordinate structures, are overlapping.

Table 11 displays the output from the analysis of the difference between NP2 and NP1 in the English experiments. This analysis yielded no significant effects, although there was a borderline effect of the interaction between Complexity and the 'English Proficiency' factor. As can be seen' in Figure 7, RTs decrease more strongly for NP1 sentences than NP2 sentences across scope conditions.

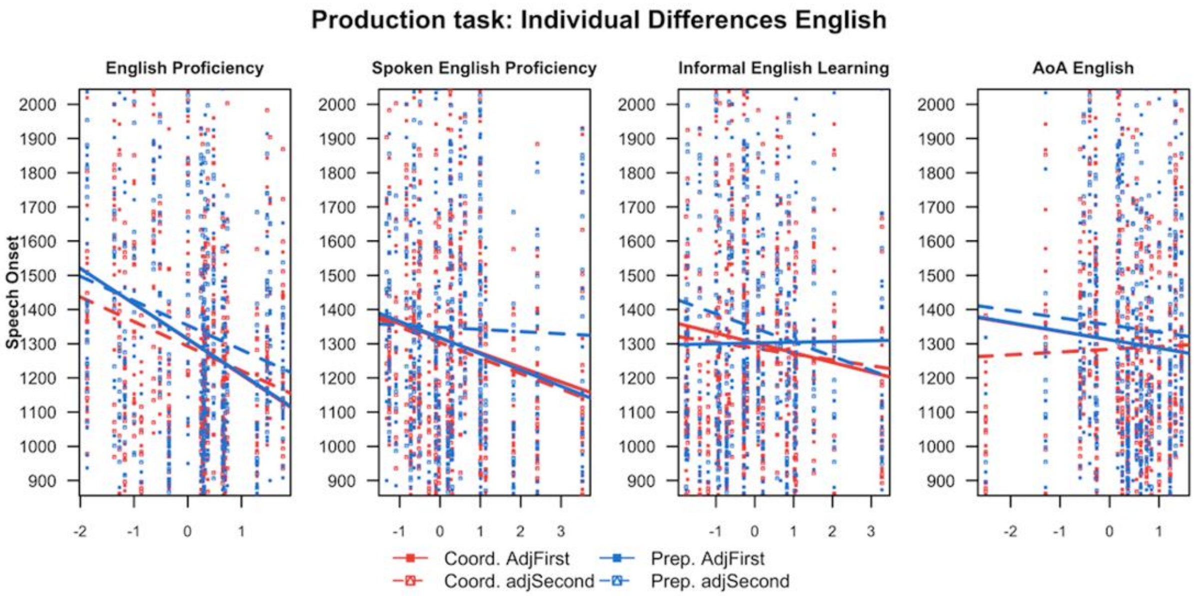


Figure 7

Individual differences from factors – English: NP2.NP1 contrast

The next analysis examined the interactions between the factors and the same two contrasts in the Norwegian data. Table 12 shows the output from the analysis including the NP1.zero contrast. The pattern of effects for each factor are shown in Figure 8.

There is a main effect of Complexity interacting with the 'Spoken English Proficiency' factor, as well as a main effect of Complexity interacting with the 'Informal Learning of English' factor. The more complex condition NP1 has slower RTs in Norwegian as the English proficiency and level of informal English learning increases.

Table 12

Individual differences from factors – Norwegian: NP1.

RT - Norwegian - NP1 - Zero				
	Estimate	Std. Error	Z value	P value
(Intercept)	1372.56	55.46	24.75	<1e-04
scope [coord, -0,47; prep, 0.53]	10.57	16.81	0.63	0.53
complexity [AdjFirst, -0.53; Zer, 0.4'	30.67	16.8	1.83	0.07
EngProficiency	-65.37	58.97	-1.11	0.27
SpokenEngProficiency	15.43	49.12	0.31	0.75
InfEngLearning	15.35	46.44	0.33	0.74
AoAEng	30.11	62.94	0.48	0.63
scope:complexity	**-81.49**	**33.66**	**-2.42**	**0.02**
scope:EngProficiency	-25.31	18.02	-1.4	0.16
scope:SpokenEngProficiency	13.79	15.32	0.9	0.37
scope:InfEngLearning	2.72	15.07	0.18	0.86
scope:AoAEng	-30.35	18.83	-1.61	0.11
complexity:EngProficiency	22.93	17.99	1.27	0.2
complexity:SpokenEngProficiency	**-47.03**	**15.37**	**-3.06**	**0**
complexity:InfEngLearning	**-59.9**	**15**	**-3.99**	**<1e-04**
complexity:AoAEng	7.21	18.82	0.38	0.7
scope:complexity:EngProficiency	-1.7	36.07	-0.05	0.96
scope:complexity:SpokenEngProficie	51.59	30.76	1.68	0.09
scope:complexity:InfEngLearning	-35.7	30.06	-1.19	0.24
scope:complexity:AoAEng	-30.83	37.71	-0.82	0.41

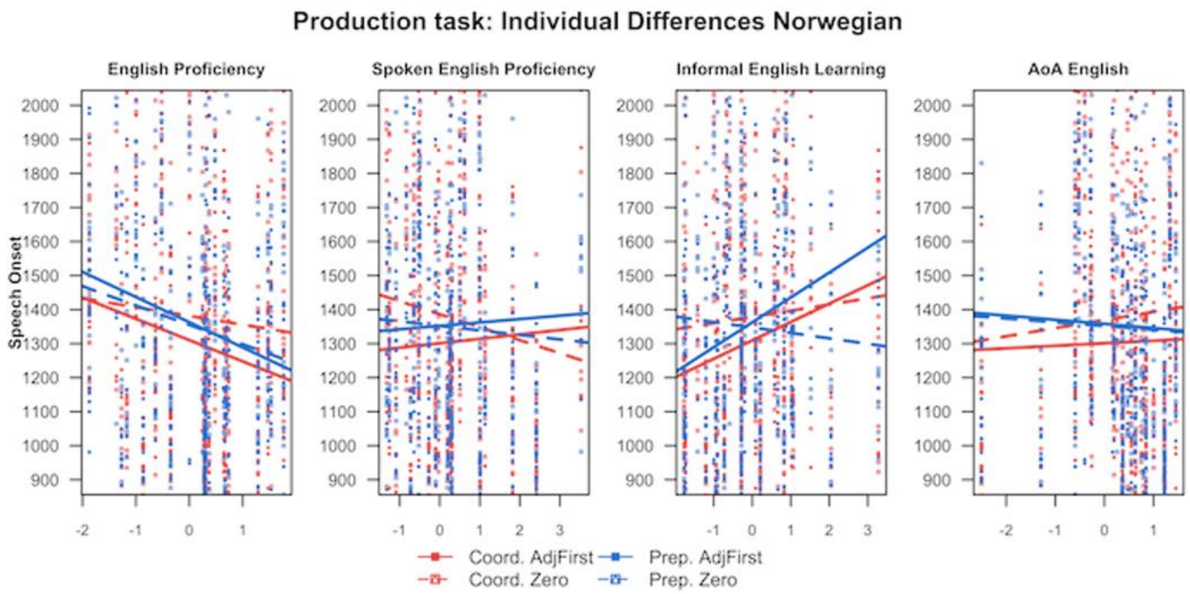


Figure 8

Individual differences from factors – Norwegian: NP1.zero contrast

The analysis of the NP2.NP1 contrast in relation to the factors (see Table 13) reveals a borderline significance for Scope interacting with the ‘Age of English Acquisition’ factor, such that a higher age of acquisition of English is associated with slower RTs in coordinate structures in Norwegian, but faster RTs in prepositional structures. More importantly, there is a significant main effect of Complexity interacting with the ‘Spoken English Proficiency’ factor such that as spoken proficiency in English increases, Norwegian NP1 sentences have longer speech onsets, while NP2 sentences have shorter speech onsets.

Table 13

Individual differences from factors – Norwegian: NP2.NP1

RT - Norwegian - NP2 - NP1				
	Estimate	Std. Error	Z value	P value
(Intercept)	1411.17	56.41	25.02	< 1e-04
scope [coord, -0,47; prep, 0.53]	-12.74	17.01	-0.75	0.45
complexity [AdjFirst, -0.53; Zer, 0.47]	**98.14**	**16.96**	**5.79**	**< 1e-04**
EngProficiency	-77.32	59.93	-1.29	0.2
SpokenEngProficiency	26.98	49.94	0.54	0.59
InfEngLearning	32.93	47.19	0.7	0.49
AoAEng	29.92	63.99	0.47	0.64
scope:complexity	**-129**	**33.9**	**-3.81**	**0**
scope:EngProficiency	-1.01	18.34	-0.06	0.96
scope:SpokenEngProficiency	9.95	15.67	0.63	0.53
scope:InfEngLearning	-3.76	15.27	-0.25	0.81
scope:AoAEng	**-35.6**	**19.43**	**-1.83**	**0.07**
complexity:EngProficiency	1.38	18.25	0.08	0.94
complexity:SpokenEngProficiency	**-33.29**	**15.59**	**-2.14**	**0.03**
complexity:InfEngLearning	-27.01	15.25	-1.77	0.08
complexity:AoAEng	8.66	19.39	0.45	0.66
scope:complexity:EngProficiency	47.17	36.46	1.29	0.2
scope:complexity:SpokenEngProficiency	41.45	31.2	1.33	0.18
scope:complexity:InfEngLearning	-38.35	30.44	-1.26	0.21
scope:complexity:AoAEng	-43.74	38.8	-1.13	0.26

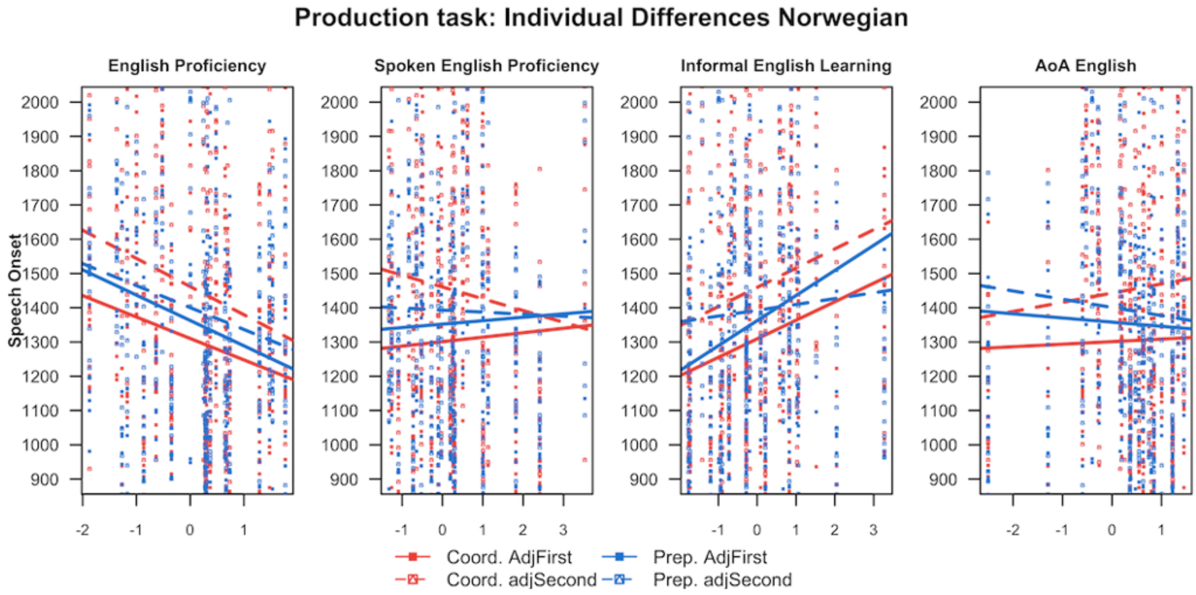


Figure 9

Individual differences from factors – Norwegian: NP2.NP1 contrast

4 Discussion

The aims of this study were to investigate how the scope of planning for bilingual sentence production differs between L1 and L2 sentence production in relation to linguistic background and experience, and to examine whether bilingual L2 planning scope interacts with sentence complexity the same way as does monolingual planning scope, as found by Smith and Wheeldon (1999) and Allum and Wheeldon (2007, 2009). To this end, the picture description experiment reported above was devised to test L1 and L2 production of coordinate and prepositional sentence types of different levels of complexity, accompanied by a bilingual profiling questionnaire. The following sections will discuss the results of this study.

4.1 Interpretation of Results

4.1.1 The Findings from the Experimental Analyses

First and foremost, the pattern of sentence planning scope was found to differ significantly between languages, as reflected in both the error data and the RT data. This difference is evident in the finding that in Norwegian, only the first functional phrase in a sentence is prioritised for grammatical encoding before speaking can begin, while in English, the entire subject phrase is grammatically encoded prior to speech onset. Evidence for this pattern can be seen in the difference between the production latencies of the different sentence structures between the languages. Results from both the error analysis and the RT analysis point towards coordinate structures being more demanding to produce than prepositional structures in Norwegian, while prepositional structures are more challenging than coordinate structures in English. The error analysis revealed that prepositional structures were less error-prone than coordinate structures in Norwegian, while in English, prepositional structures were more error-prone than coordinate structures. Similarly, the RT analysis revealed that, on the whole, RTs for prepositional structures are either slower than or similar to coordinate structures throughout in English. At first glance this finding seems sensible, as prepositional structures are arguably more complex than coordinate structures due to their hierarchical structure, which would reasonably cause longer speech onset latencies for these sentences. However, in Norwegian, prepositional structures had shorter RTs than coordinate structures despite being presumably more complex. The fact that prepositional structures take longer to produce than coordinate structures in English therefore suggests that in English, participants employ a more extensive planning scope, completing grammatical encoding of the entire subject phrase, i.e. the whole prepositional structure, before speech onset. These results are in line with the findings of Konopka et al. (2018) that L2 speakers utilise a larger planning scope than L1 speakers.

That the opposite is true for Norwegian, with longer speech onsets and increased error rates for coordinate structures, is consistent with prioritising the first functional phrase only for grammatical encoding, echoing the findings of Smith and Wheeldon (1999) and Allum and Wheeldon (2007, 2009). As previously mentioned, prepositional structures are more complex grammatically than coordinate structures. The only way coordinate structures would be more complex than prepositional structures is if speakers were not encoding the whole subject phrase grammatically, but instead prioritising the initial functional phrase. For example, if only the initial functional phrase is prioritised for prepositional structures, the scope of the first functional phrase in "[the dog] above the car" is smaller than the entire subject phrase, "[the dog above the car]." In contrast, for coordinate structures the first functional phrase would constitute the entire subject phrase and thus be considerably larger, e.g. "[the dog and the car]." Thus, a coordinate structure would require a larger planning scope than a prepositional structure. As such, the pattern for the L1, Norwegian, suggests a smaller scope in which the initial functional phrase only is prioritised for grammatical encoding prior to speech onset, while the pattern for the L2, English, indicates that the entire subject phrase is grammatically encoded prior to speech onset regardless of sentence type. This finding thus reveals a major difference between L1 and L2 planning scope.

The pattern of complexity effects provides further evidence for a different processing scope in the L2 than the L1. The RT analysis showed that when complexity is added to the first phrase of a coordinate structure in Norwegian, speech onset was reduced compared to zero complexity sentences. This onset reduction suggests that the added complexity breaks up the planning scope of the coordinate structure, and that only the first, modified noun phrase is grammatically encoded prior to speech onset, e.g. "[the yellow dog] and the car." In other words, speakers will prioritise the first noun phrase for encoding before moving on to process the second noun phrase incrementally after speech onset. In contrast, when the second noun phrase of a coordinate was modified by an adjective in Norwegian, speech onset was considerably slower than in the zero condition, and such sentences were also more error-prone than the zero condition. Speech onset to NP2 coordinate structures was also slower compared to the NP2 prepositional structures, consistent with the pattern seen in the zero condition in which coordinate structures had slower RTs than prepositional structures. Altogether, these complexity effects suggest that in the L1, Norwegian, scope is not broken up when adding complexity to the second noun phrase, but that the second noun phrase of the coordinate structure, including the adjective, is also processed for grammatical encoding before speech onset. Thus, it is only when complexity is added to the first noun phrase of a coordinate that a breaking up of scope can be found, while adding complexity to the second noun phrase will further extend the scope of grammatical planning prior to speech onset, leading to longer speech onset latencies.

Importantly, in the L2, English, there was no evidence for a breaking up of scope in any condition, as speech onsets increased with each complexity level. English NP1 sentences consistently had longer RTs than zero complexity sentences, regardless of scope, and prepositional structures had longer RTs overall than coordinate structures, except for NP1 sentences where the RTs for both structures were the same. In English, there were also more errors in prepositional structures than coordinate structures whether they constitute the first or the second noun phrase of the sentence. Thus, planning scope appears to remain regular in English across conditions, with speech onset latencies and error rates reflecting complexity levels without evidence of scope breaking up into smaller increments. This pattern of complexity effects offers further evidence of tangible scope differences between L1 and L2 planning.

The finding that in English, participants seem to complete grammatical encoding of the entire subject phrase before speech onset, regardless of whether it is a coordinate or a prepositional structure, is not consistent with linearly incremental planning in the L2, in which the speaker breaks up the utterance into smaller planning windows at risk of disfluency in order to initiate production faster. Rather, the slower speech onsets for increased complexity levels in the L2 are more in line with Hierarchical Incrementality in which larger message elements are encoded before speech onset, thereby delaying speech onset in favour of producing fluent utterances. The finding that the English experimental trials were not more error-prone than the Norwegian trials offers further support for a prioritisation of fluency at the expense of speed. L1 production, on the other hand, is more in line with Linear Incrementality, as suggested by the less extensive planning scope prior to speech onset and the breaking up of scope into smaller increments by increasing N1 complexity in a coordinate structure. Thus, these results mirror those of Konopka et al. (2018), who found that bilinguals prefer to use a hierarchically incremental planning strategy and employ a larger planning scope in their less practiced language, while employing a linearly incremental planning strategy with smaller planning windows in their native language.

As previously stated, the RT analysis and the error analysis are largely in agreement, with the pattern of errors mirroring that of the RT analysis. Thus, there is no evidence of a speed/accuracy trade-off in the data, meaning that participants were not slowing down in order to be more accurate. As such, slower reaction times can be judged to reliably reflect the level of production difficulty. However, at this point an unexpected finding revealed by the RT analysis should be addressed, namely that participants were faster in English than in Norwegian. This finding is contrary to the assumption that L2 production is slower and more difficult than L1 production due to the increased cognitive load of speaking in a less practised language, especially considering the finding described above, that speaking in the L2 involves a more extensive planning scope than in the L1. The faster speech onsets for English trials are also contrary to the findings of Konopka et al. (2018), who found a delay in L2 speech onset compared to the L1 tasks.

The faster initiation of English production may be a practice effect due to the participants having already completed the task in Norwegian. However, such a practice effect was not found in an unrelated bilingual language comprehension experiment that was run simultaneously with the production experiment. This experiment was also conducted on separate days in Norwegian first and English subsequently, with the same participant group. In this comprehension experiment, participants were slower to respond in English than in Norwegian. Thus, the effect of participants being faster in English seems to be particular to complex production tasks. To find out whether this difference in speed is a practice effect or whether it can be contributed to language, the experiment would have to be counterbalanced by running another 27 participants in the opposite order, with English first and Norwegian second. This could not be done due to the limitations of this study; however, it presents an interesting opportunity for future research.

4.1.2 The Findings from the Bilingual Profiling Questionnaire

The factor analysis generated four factors that were taken to represent ‘English Proficiency,’ ‘Spoken English Proficiency,’ ‘Informal Learning of English’ and ‘Age of English Acquisition.’ All of these factors were interpreted as representing English learning and usage. Interestingly, age-related variables played a large part in all of the findings. Firstly, a higher age of fluency in Norwegian played a part in the factor ‘Informal Learning of English,’ suggesting that a lower age of fluency in Norwegian is associated with informal learning in English. In other words, people who became fluent in Norwegian at a young age reported a higher contribution of TV, music and reading to the learning of English, and a higher degree of exposure to English TV, music and interactions with friends. This finding suggests that informal channels play a larger part in learning English in cases where Norwegian fluency is attained at a younger age. It may be that people who become fluent in their L1 at a younger age are better equipped to *acquire* the L2 implicitly through informal immersion rather than explicitly *learning* it via formal schooling, and thus that there is a link between early L1 acquisition and implicit L2 acquisition. The investigation of this link poses an intriguing opportunity for further research.

The factor specifically associated with age-related variables, 'Age of English Acquisition,' suggested that people who acquire English later in life have more contribution to their English learning from interacting with friends. This finding implies that informal learning through social interactions plays a larger part in L2 acquisition than formal schooling when the language is learnt later in life. Additionally, a later age of acquisition of English was associated with more reading and watching TV in Norwegian, suggesting that later L2 acquisition is associated with more L1 usage in daily life. This finding can be interpreted as late L2 learners using their L1 at the expense of L2 learning, and thus that the frequent use of Norwegian is detrimental to the learning of English. Such an attrition effect of the use of one language negatively affecting that of the other is to be expected. Furthermore, a later English age of acquisition made it less likely for participants to intentionally insert English words into Norwegian speech, known as code-switching. Thus, thus finding suggests that frequent code-switching can be associated with a lower age of L2 acquisition.

The 'English Proficiency' factor revealed an association between age of Norwegian fluency and proficiency in English, in that participants who acquired Norwegian fluency at a later age were more proficient in English. Likewise, the factor 'Spoken English Proficiency' suggested that an older age of fluency in Norwegian was associated with a higher level of English speaking proficiency, and thus that speaking more English is delaying Norwegian proficiency. It was not unexpected to find some attrition effects of English usage on

Norwegian. However, that the amount of English, and especially spoken English, is relating negatively to participants' self-rating of Norwegian age of fluency is unusual given that this participant group represented Norwegian-dominant bilinguals who have grown up immersed in their first language, and who would in all probability not have been exposed to English to a large extent until after they were fluent in Norwegian. All in all, there seems to be a clear relationship between people's L2 proficiency level and their ages of L1 and L2 acquisition, although the precise nature of this relationship is a question for further research.

4.1.3 Interpreting Individual Differences in Experimental Performance

When factoring in the effects of individual differences in language factors on participants' performance in the experimental task, the findings become less conclusive than those reported above. Firstly, in the Norwegian experiments, it was interesting to find an increase in RTs for the NP1 complexity level as a result of higher speaking proficiency in English. The slower speech onsets for the initial modified noun phrase could reflect a trend of having increased difficulty retrieving adjectives in Norwegian due to interference from English, especially since Norwegian adjectives are arguably more complex than English adjectives because their form has to agree with that of the noun. It is possible that this interference effect reflects a competition between the two languages' lexicons during lexical access, in line with the language-nonspecific view, although further study is needed in order to give conclusive interpretations on this matter.

Another interesting finding was that NP2 sentences in Norwegian have shorter speech onsets as spoken English proficiency increases. As there was otherwise no evidence for a breaking up of planning scope for sentences in which the second noun is modified in the L1 (see section 4.1.1), finding that a higher proficiency in English spoken usage may influence planning scope for Norwegian NP2 sentences warrants further investigation. Furthermore, it was unexpected to find that for the Norwegian trials, a higher age of acquisition of English led to slower speech onsets in coordinate structures, but faster RTs to prepositional structures. It is unclear in what way acquiring fluency in English at a later age would affect the production of coordinate and prepositional structures in Norwegian, especially as there is no evidence of coordinate structures being planned to a greater extent than prepositional structures in English. On the contrary, prepositional structures were more error-prone and had longer RTs in English than coordinate structures. Accordingly, this effect of English age of acquisition on Norwegian speech onset latencies to different sentence structures is another subject for future research.

As regards the English experimental trials, it was found that as English proficiency increases, RTs get faster in English across conditions, and complex sentences have the strongest reduction of speech onset. These findings were unremarkable, as it is to be expected that higher English proficiency equals faster reaction times. Surprising, however, was the finding that the factor ‘Informal Learning of English’ was associated with reduced RTs to coordinate structures and increased RTs to prepositional structures in English. In other words, participants who reported a higher degree of contributions from informal channels to their learning of English had longer speech onset latencies to prepositional structures and shorter latencies to coordinate structures. As prepositional structures seem to be more difficult to produce than coordinate structures in English, given that they have longer RTs and are more error-prone overall than coordinate structures, it would have been expected to see stronger effects of proficiency measures in English on prepositional structures, i.e. a reduction in RTs for prepositional structures. That the opposite was the case, and participants instead became faster in coordinate structures due to more informal learning of and exposure to English, is therefore surprising and warrants further exploration.

Lastly, it was interesting to find that only the 'Spoken English Proficiency' factor showed signs of attrition effects on Norwegian proficiency. The more general 'English Proficiency' factor did not appear to negatively affect Norwegian proficiency, which is the opposite of what was expected, and may suggest that this factor is not necessarily a true L2 factor. As many of the Norwegian variables had to be removed due to insufficient co-variation or being too similar to their English equivalents (see section 3.1.5) before running the factor analysis, some of the resulting factors may in truth reflect more general language proficiency effects on both L1 and L2 processing. The finding that the 'English Proficiency' factor shows a similar pattern of effects in both languages and that there are no clear effects of proficiency in one language being detrimental to the other, which is what most research on bilingualism would suggest, raises doubt as to the extent to which this factor is a true English factor. For example, this factor reflecting general language proficiency rather than English proficiency specifically could explain why higher degrees of this factor has only beneficial effects on the speech onset latencies of both languages. As it is unclear whether this factor is a true reflection of English proficiency, it is difficult to say anything conclusive about whether a higher English proficiency leads to native-like performance in English for Norwegian-English bilinguals, as predicted prior to the experiments. More research on this aspect of the study regarding proficiency effects on bilingual planning and performance is therefore needed.

To conclude this section, it should be noted that the participant set for the study turned out to be a very uniform group of bilinguals, with a high mean level of education of almost seventeen years. They had a generally high level of English proficiency, as on a scale from 0-10, the average self-reported number for all sub-categories was never below 6.92 (see Table 5). The mean age of attaining speaking and reading fluency in English was approximately thirteen, meaning that all participants, at eighteen years or older, had had at least five years of experience with fluent English usage. Hence, both the experiment and the questionnaire study would have possibly yielded quite different results had the participant set had a wider range of proficiency level. For future research, it would be interesting to see what effects of bilingual profile could be found on experimental performance with a larger and more heterogeneous group of participants.

4.2 Experimental Observations and Issues

Throughout the weeks of running the experiment, observations were made which made clear some limitations and issues pertaining to the design and execution of the experiments. The most notable of these will now be addressed briefly with suggestions for how improvements might be made. The foremost issue that presented itself early in the course of testing is that not all of the translations were as non-cognate as had been thought. English has synonyms for ‘above’ and ‘below,’ instead of which it is also possible to say ‘over’ and ‘under,’ which happen to be the Norwegian translations for these words. For this reason, many of the Norwegian participants tended to intuitively use ‘over’ and ‘under’ during the English experiment instead of ‘above’ and ‘below,’ despite being thoroughly instructed in the target sentence types and despite being repeatedly corrected during and after the practice blocks. This was also true to some extent for the adjective ‘sparkly,’ for which the Norwegian translation was ‘*glitrende*.’ However, although not as frequently used, an English alternative is the cognate ‘glittering,’ which led to some Norwegians intuitively preferring this adjective over ‘sparkly’ due to its resemblance to ‘*glitrende*.’ These translation issues observably led to high disfluency rates, especially in cases where the participant had completed the Norwegian experiment only the day before, which seemed to prime the problematic forms. The Norwegian experiment thus appeared to interfere with participants’ performance in the English experiment when not separated by enough time. This problem could have been avoided by spacing out the experiments so that all participants had a longer interval, for example a week, between the two experiments. Alternatively, and perhaps preferably, more thought could have been put into the cognate status of the synonyms of translation words.

It was also observed that, in both languages, participants often paused slightly before the second noun phrase in a sentence, whether it be a coordinate or a prepositional structure. There is a possibility that such hesitations reflect a restructuring in planning strategy while production was ongoing. Distinct hesitations mid-sentence were coded as fluency errors, which as previously mentioned represented the bulk of the experimental errors (see section 3.2). Due to the limitations of this thesis, conducting a breakdown of error type was not feasible due to constraints of time and space. However, if analysing these fluency errors in greater detail had been possible, it would have been interesting to investigate how bilingual speakers might utilise planning strategies while speaking is ongoing, and what aspects of L2 experience and proficiency might factor into the choice of planning strategy.

Lastly it is worth mentioning that, during the course of testing, it became apparent that the order of naming, top-to-bottom and left-to-right, was not intuitive to all participants. It seems that some people naturally prioritise a top-to-bottom order, while others prioritise left-to-right. Accordingly, in trials wherein the top right and bottom left pictures of the picture grid stimulus were framed, some participants, contrary to instructions, would read the bottom left corner image before the top right one, and thus the produced utterance would not match the target sentence. For this reason, it might be best to avoid this particular layout in future studies (see Appendix B, layout 4).

4.3 Summary and Conclusion

To sum up the results yielded by this study, firstly, it was found that Norwegian and English speakers employ significantly different planning scopes. Whereas only the first functional phrase is prioritised for grammatical encoding before speech onset in Norwegian, in English the entire subject phrase is grammatically encoded prior to speaking. Data from both the error analysis and the RT analysis revealed a larger processing scope for coordinate structures than prepositional structures in the L1, suggesting that native speakers employ smaller planning scope when producing prepositional structures than when producing coordinates. The fact that these scope differences between sentence types were not found in L2 trials suggests that in the L2, the entire subject phrase is grammatically encoded prior to speech onset, reflecting a major scope difference between L1 and L2 planning. Additionally, the finding that adding complexity to the first noun in a coordinate structure will cause L1 speakers to break up planning scope into smaller increments, but that L2 speakers will not break up scope regardless of noun phrase complexity, provides further evidence of tangible scope differences between L1 and L2 planning. Furthermore, the finding that L2 speakers complete grammatical encoding of the entire subject phrase before speech onset regardless of sentence type is consistent with employing a hierarchical planning strategy in the less practiced language.

The results from the factor analysis revealed links between a younger age of fluency in Norwegian and informal learning of English, as well as between a later age of English acquisition and informal English learning. Late L2 learners also appear less prone to code switching. Additionally, a high English speaking proficiency seems to be connected to a later age of Norwegian fluency. As for the relationship between bilingual profile and experimental performance, a high proficiency in English was predictably linked to faster speech onsets in English. Furthermore, English appeared to interfere with the retrieval of adjectives in Norwegian in highly proficient English speakers, which may reflect competition between the words from both languages during lexical retrieval. Some unexpected associations between age of fluency in both languages and proficiency in English were found, indicating an undeniable relationship between age of acquisition and L2 proficiency, the understanding of which requires further examination. Regrettably, several findings relating to proficiency and language background factors proved difficult to explain, not least due to uncertainty as to whether the ‘English Proficiency’ factor truly reflected English proficiency specifically or more general language proficiency. It was therefore difficult to ascertain whether higher English proficiency suggest more L1-like planning.

I will conclude this discussion by comparing the experimental results to the general predictions. The first prediction was that performance in the L2 experimental task would be more native-like the higher the participant’s L2 proficiency level. However, due to the equivocal nature of the ‘English proficiency’ factor and the inconclusiveness of several of the findings when relating the bilingual profile factors to experimental performance, the relationship between English proficiency and planning strategies proved difficult to satisfactorily interpret.

The second prediction was that bilingual L2 planning scope and monolingual L1 planning scope would display a similar pattern of scope differences between coordinate and prepositional structures. However, experimental results revealed tangible scope differences between L1 and L2 planning. In the L1, speakers prioritised only the first functional phrase for grammatical encoding before speaking could begin, thus prepositional structures had shorter speech onsets than coordinate structures. In the L2, the entire subject phrase was grammatically encoded prior to speech onset, thus prepositional structures had longer speech onsets than coordinate structures. Neither did the prediction hold true that L2 speakers would have a longer SOL than L1 speakers. Although it may have been a practice effect, participants proved to be faster in the English experiment.

The third prediction was that adding modifiers to either noun phrase in a conjoined noun phrase would either break down planning scope and lead to faster speech onsets for sentences with a coordinate structure than for those with prepositional structures, or make no difference to scope and lead to an increase in speech onsets for each complexity level. The first alternative turned out to be accurate for L1 production. Adding complexity to the initial noun phrase did indeed break up the planning scope in conjoined noun phrases, such that planning scope became more similar to that of prepositional structures. However, in the L2 the effect of added complexity behaved more as predicted by the second alternative, as speech onset latencies instead increased consistently with each complexity level.

To conclude, the findings obtained by the present study suggest that the two languages of a bilingual operate with different scopes of grammatical encoding during the sentence planning process. Moreover, this study revealed that for sentence-initial conjoined noun phrases, adding complexity to the first noun phrase will break up planning scope, in that speakers will prioritise only the initial noun phrase for grammatical encoding prior to speech onset. That this effect only applies to the speaker's native language provides further evidence in favour of second language sentence planning differing significantly from first language sentence planning.

REFERENCES

Allum, P. H. & Wheeldon, L. (2007). Planning Scope in Spoken Sentence Production: The Role of Grammatical Units. *Journal of Experimental Psychology: Learning Memory and Cognition*, 33(4), 791-810. https://doi.org/10.1037/0278-7393.33.4.791

Allum, P. H. & Wheeldon, L. (2009). Scope of Lexical Access in Spoken Sentence Production: Implications for the Conceptual-Syntactic Interface. *Journal of Experimental Psychology: Learning Memory and Cognition*, 35(5), 1240–1255. https://doi.org/10.1037/a0016367

Bock, J. K. (1986). Syntactic Persistence in Language Production. *Cognitive Psychology*, 18, 355-387. https://doi.org/10.1016/0010-0285(86)90004-6

Bock, K., Dell, G. S., Chang, F. & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition*, 104(3), 437-458. https://doi.org/10.1016/j.cognition.2006.07.003

Bock, K. & Griffin, Z. M. (2000). The Persistence of Structural Priming: Transient Activation or Implicit Learning? *Journal of Experimental Psychology: General*, 129(2), 177-192. https://doi.org/10.1037//0096-3445.129.2.177

Bock, K., Irwin, D. E., Davidson, D. J. & Levelt, W. J. M. (2003). Minding the clock. *Journal of Memory and Language*, 48(4), 653-685. https://doi.org/10.1016/S0749-596X(03)00007-X

Bock, K. & Levelt, W. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). Academic Press.

Brown-Schmidt, S. & Konopka, A. E. (2008). Little houses and casas pequeñas: Message formulation and syntactic form in unscripted speech with speakers of English and Spanish. *Cognition*, 109(2), 274-280. https://doi.org/10.1016/j.cognition.2008.07.011

Brown-Schmidt, S. & Konopka, A. E. (2014). Processes of incremental message planning during conversation. *Psychonomic Bulletin & Review*, 22(3), 833-843. https://doi.org/10.3758/s13423-014-0714-2

Brown-Schmidt, S. & Tanenhaus, M. K. (2006). Watching the eyes when talking about size: An investigation of message formulation and utterance planning. *Journal of Memory and Language*, 54(4), 592-609. https://doi.org/10.1016/j.jml.2005.12.008

Chang, F. (2002). Symbolically speaking: a connectionist model of sentence production. *Cognitive Science*, 26(5), 609-651. https://doi.org/10.1016/S0364-0213(02)00079-4

Chang, F., Dell, G. S. & Bock, K. (2006). Becoming Syntactic. *Psychological Review*, 113(2), 234-272. https://doi.org/10.1037/0033-295X.113.2.234

Chang, F., Dell, G. S., Bock, K. & Griffin, Z. M. (2000). Structural Priming as Implicit Learning: A Comparison of Models of Sentence Production. *Journal of Psycholinguistic Research*, 29(2), 217-229. https://doi.org/10.1023/a:1005101313330

Cleland, A. A. & Pickering, M. J. (2003). The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. *Journal of Memory and Cognition*, 49(2), 214-230. https://doi.org/10.1016/S0749-596X(03)00060-3

Cleland, A. A. & Pickering, M. J. (2006). Do writing and speaking employ the same syntactic representations? *Journal of Memory and Language*, 54(1), 185-198. https://doi.org/10.1006/jmla.2001.279310.1016/j.jml.2005.10.003

Colomé, Á. (2001). Lexical Activation in Bilinguals' Speech Production: Language-Specific or Language-Independent? Journal of Memory and Language, 45(4), 721-736. https://doi.org/10.1006/jmla.2001.2793

Corley, M. & Scheepers, C. (2002). Syntactic priming in English sentence production: Categorical and latency evidence from an Internet-based study. *Psychonomic Bulletin & Review*, 9(1), 126-131. https://doi.org/10.3758/BF03196267

Costa, A. (2005). Lexical Access in Bilingual Production. In J. F. Kroll & A. M. B. de Groot (Eds.) *Handbook of Bilingualism – Psycholinguistic Approaches* (pp. 308-325). Oxford University Press.

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). https://doi.org/10.1037//0278-7393.26.5.1283

Costa, A., Miozzo, M. & Caramazza, A. (1999). Lexical Selection in Bilinguals: Do Words in the Bilingual's Two Lexicons Compete for Selection? *Journal of Memory and Language*, 41(3), 365-397. https://doi.org/10.1006/jmla.1999.2651

Damian, M. F., Bowers, J. S., Stadthagen-Gonzalez, H. & Spalek, K. (2010). Does Word Length Affect Speech Onset Latencies When Producing Single Words? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(4), 892-905. https://doi.org/10.1037/a0019446

de Bot, K. (1992). A Bilingual Production Model: Levelt's 'Speaking' Model Adapted. *Applied Linguistics*, 13(1), 1-24, https://doi.org/10.1093/applin/13.1.1

Dell, G. S. (1986). A Spreading-Activation Theory of Retrieval in Sentence Production. *Psychological Review*, 93(3), 283-321. https://doi.org/10.1037/0033-295X.93.3.283

Dell, G. S., Oppenheim, G. M. & Kittredge, A. K. (2008). Saying the right word at the right time: Syntagmatic and paradigmatic interference in sentence production. *Language, Cognition and Neuroscience*, 23(4), 583-608. https://doi.org/10.1080/01690960801920735

Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E. & Brysbaert, M. (2018). MultiPic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology*, 71(4), 808-816. https://doi.org/10.1080/17470218.2017.1310261

EF Education First. (2020). *EF English Proficiency Index*. https://www.ef.com/wwen/epi/#

Ferrand, L. & New, B. (2003). Syllabic length effects in visual word recognition and naming. *Acta Psychologica*, 113(2), 167-183. https://doi.org/10.1016/s0001-6918(03)00031-3

Ferreira, F. & Swets, B. (2002). How Incremental Is Language Production? Evidence from the Production of Utterances Requiring the Computation of Arithmetic Sums. *Journal of Memory and Language*, 46(1), 57-84. https://doi.org/10.1006/jmla.2001.2797

Ferreira, V. S. & Pashler, H. (2002). Central Bottleneck Influences on the Processing Stages of Word Production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28(6), 1187-1199. https://doi.org/10.1037//0278-7393.28.6.1187

Fromkin, V. A. (1971). The Non-Anomalous Nature of Anomalous Utterances. *Language*, 47(1), 27-52. https://doi.org/10.2307/412187

Gleitman, L. R., January, D., Nappa, R. & Trueswell, J. C. (2007) On the *give* and *take* between event apprehension and utterance formulation. *Journal of Memory and Language*, 57(4), 544-569. https://doi.org/10.1016/j.jml.2007.01.007

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1(2), 67-81, https://doi.org/10.1017/S1366728998000133

Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition*, 82(1), B1-B14. https://doi.org/10.1016/S0010-0277(01)00138-X

Griffin, Z. M. & Bock, K. (2000). What the Eyes Say About Speaking. *Psychological Science*, 11(4), 274-279. https://doi.org/10.1111/1467-9280.00255

Griffin, Z. M. & Ferreira, V. S. (2006). Properties of Spoken Language Production. In M. J. Traxler & M. A. Gernsbacher (Eds.) *Handbook of Psycholinguistics – Second Edition* (pp. 21-59). Academic Press. https://doi.org/10.1016/B978-012369374-7/50003-1

Guevara, E. (2010). NoWaC: a large web-based corpus for Norwegian. *Proceedings of the NAACL HLT 2010 Sixth Web as Corpus Workshop*, 1–7.

Hermans, D., Bongaerts, T., de Bot, K. & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition*, 1(3), 213-229. https://doi.org/10.1017/S1366728998000364

Kempen, G. & Hoenkamp, E. (1987). An Incremental Procedural Grammar for Sentence Formulation. *Cognitive Science*, 11(2), 201-258. https://doi.org/10.1016/S0364-0213(87)80006-X

Kempen, G. & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition*, 14(2), 185-209. https://doi.org/10.1016/0010-0277(83)90029-X

Konopka, A. E. & Brown-Schmidt, S. (2014). Message Encoding. In M. Goldrick, Ferreira, V. S. & Miozzo, M. (Eds.) *Oxford library of psychology. The Oxford Handbook of Language Production* (pp. 3-20). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199735471.013.010

Konopka, A. E. & Meyer, A. S. (2014). Priming sentence planning. *Cognitive Psychology*, 73, 1-40. https://doi.org/10.1016/j.cogpsych.2014.04.001

Konopka, A. E., Meyer, A. S. & Forest, T. A. (2018). Planning to speak in L1 and L2. *Cognitive Psychology*, 108, 72-104. https://doi.org/10.1016/j.cogpsych.2017.12.003

Kuchinsky, S. E., Bock, K. & Irwin, D. E. (2011). Reversing the hands of time: Changing the mapping from seeing to saying. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 37(3), 748-756. https://doi.org/10.1037/a0022637.

La Heij, W. (1988). Components of Stroop-like interference in picture naming. *Memory and Cognition*, 16(3), 400-410. https://doi.org/10.3758/BF03214220

La Heij, W. (2005). Selection Processes in Monolingual and Bilingual Lexical Access. In J. F. Kroll & A. M. B. de Groot (Eds.) *Handbook of Bilingualism – Psycholinguistic Approaches* (pp. 289-307). Oxford University Press.

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. The MIT Press.

Levelt, W. J. M. & Maassen, B. (1981). Lexical Search and Order of Mention in Sentence Production. In W. Klein & W. J. M. Levelt (Eds.) *Crossing the Boundaries in Linguistics* (pp. 221-252). D. Reidel Publishing Company. DOI: 10.1007/978-94-009-8453-0_12

Levelt, W. J. M., Roelofs, A. & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences,* 22(1), 1-75. https://doi.org/10.1017/S0140525X99001776

Levelt, W. J. M. & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition*, 50(1-3), 239-269. https://doi.org/10.1016/0010-0277(94)90030-2

MacWhinney, B. (2005). A Unified Model of Language Acquisition. In J. F. Kroll & A. M. B. de Groot (Eds.) *Handbook of Bilingualism – Psycholinguistic Approaches* (pp. 49-67). Oxford University Press.

Mahowald, K., James, A., Futrell, R. & Gibson, E. (2016). A meta-analysis of syntactic priming in language production. *Journal of Memory and Language*, 91(3), 5-27. https://doi.org/10.1016/j.jml.2016.03.009

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), 940-67. https://doi.org/10.1044/1092-4388(2007/067)

Melinger, A. & Dobel, C. (2005). Lexically-driven syntactic priming. *Cognition*, 98(1), B11-20. https://doi.org/10.1016/j.cognition.2005.02.001

Meuter, R. F. I. & Allport, A. (1999). Bilingual Language Switching in Naming: Asymmetrical Costs of Language Selection. *Journal of Memory and Language*, 40(1), 25-40. https://doi.org/10.1006/jmla.1998.2602

Meyer, A. S. (1996). Lexical Access in Phrase and Sentence Production: Results from Picture–Word Interference Experiments. *Journal of Memory and Language*, 35(4), 477-496. https://doi.org/10.1006/jmla.1996.0026

Meyer, A. S., Belke, E., Häcker, C. & Mortensen, L. (2007). Use of word length information in utterance planning. *Journal of Memory and Language*, 57(2), 210-231. https://doi.org/10.1016/j.jml.2006.10.005

Meyer, A. S., Roelofs, A. & Levelt, W. J. M. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language*, 48(1), 131-147. https://doi.org/10.1016/S0749-596X(02)00509-0

Meyer, A. S. & Schriefers, H. (1991). Phonological Facilitation in Picture-Word Interference Experiments: Effects of Stimulus Onset Asynchrony and Types of Interfering Stimuli. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 17(6), 1146-1160. https://doi.org/10.1037/0278-7393.17.6.1146

Meyer, A. S., Sleiderink, A. M. & Levelt, W. J. M. (1998). Viewing and naming objects: eye movements during noun phrase production. *Cognition*, 66(2), B25-B33. https://doi.org/10.1016/S0010-0277(98)00009-2

Meyer, A. S., Wheeldon, L., van der Meulen, F. & Konopka, A. (2012). Effects of speech rate and practice on the allocation of visual attention in multiple object naming. *Frontiers in Psychology*, 3:39. https://doi.org/10.3389/fpsyg.2012.00039

Oppermann, F., Jescheniak, J. D., Schriefers, H. (2010). Phonological advance planning in sentence production. *Journal of Memory and Language*, 63(4), 526-540. https://doi.org/10.1016/j.jml.2010.07.004

Pashler, H. (1994). Dual-Task Interference in Simple Tasks: Data and Theory. *Psychological Bulletin*, 116(2), 220-244. https://doi.org/10.1037//0033-2909.116.2.220

Pickering, M. J. & Branigan, H. P. (1998). The Representation of Verbs: Evidence from Syntactic Priming in Language Production. *Journal of Memory and Language*, 39(4), 633-651. https://doi.org/10.1006/jmla.1998.2592

Pickering, M. J. & Ferreira, V. S. (2009). Structural Priming: A Critical Review. *Psychological Bulletin*, 134(3), 427-459. https://doi.org/10.1037/0033-2909.134.3.427

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42(1-3), 107-142. https://doi.org/10.1016/0010-0277(92)90041-F

Roelofs, A. (1998). Lemma selection without inhibition of languages in bilingual speakers. *Bilingualism: Language and Cognition*, 1(02), 94-95. https://doi.org/10.1017/S1366728998000194

Roelofs, A. (2015). Modeling of phonological encoding in spoken word production: From Germanic languages to Mandarin Chinese and Japanese. *Japanese Psychological Research*, 57(1), 22-37. https://doi.org/10.1111/jpr.12050

Santiago, J., MacKay, D. G., Palma, A. & Rho, C. (2000). Sequential activation processes in producing words and syllables: Evidence from picture naming. *Language and Cognitive Processes*, 15(1), 1-44. https://doi.org/10.1080/016909600386101

Schnur, T. T., Costa, A. & Caramazza, A. (2006). Planning at the Phonological Level during Sentence Production. *Journal of Psycholinguistic Research*, 35(2), 189-213. https://doi.org/10.1007/s10936-005-9011-6

Schriefers, H., Meyer, A. S. & Levelt, W. J. M. (1990). Exploring the Time Course of Lexical Access in Language Production: Picture-Word Interference Studies. *Journal of Memory and Language*, 29(1), 86-102. https://doi.org/10.1016/0749-596X(90)90011-N

Slevc, L. R. (2011). Saying what's on your mind: Working memory effects on sentence production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 37(6), 1503-1514. https://doi.org/10.1037/a0024350

Smith, M. & Wheeldon, L. (1999). High level processing scope in spoken sentence production. *Cognition*, 73(3), 205-246. https://doi.org/10.1016/S0010-0277(99)00053-0

Smith, M. & Wheeldon, L. (2001). Syntactic priming in spoken sentence production - an online study. *Cognition*, 78(2), 123-164. https://doi.org/10.1016/S0010-0277(00)00110-4

Swets, B., Jacovina, M. E. & Gerrig, R. J. (2014). Individual Differences in the Scope of Speech Planning: Evidence from Eye Movements. *Language and Cognition*, 6(1), 12-44. https://doi.org/10.1017/langcog.2013.5

van Hell, J. & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review*, 9(4), 780-789. https://doi.org/10.3758/BF03196335

van Heuven, W. J. B., Mandera, P., Keuleers, E. & Brysbaert, M. (2006). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, 67(6), 1176-1190. http://dx.doi.org/10.1080/17470218.2013.850521

Wagner, V., Jescheniak, J. D. & Schriefers, H. (2010). On the Flexibility of Grammatical Advance Planning During Sentence Production: Effects of Cognitive Load on

Multiple Lexical Access. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 36(2), 423-440. https://doi.org/10.1037/a0018619

Ward, N. (1992). A parallel approach to syntax for generation. *Artificial Intelligence*, 57(2-3), 183-225. https://doi.org/10.1016/0004-3702(92)90017-R

Wheeldon, L. (2011). Generating spoken sentences: The relationship between words and syntax. *Language and Linguistics Compass*, 5(6), 310-321. https://doi.org/10.1111/j.1749-818X.2011.00281.x

Wheeldon, L. (2012). Producing spoken sentences: The scope of incremental planning. In S. Fuchs, M. Weirich, D. Pape & P. Perrier (Eds.) *Speech Production and Perception Volume 1: Speech Planning and Dynamics* (pp. 93-114). Peter Lang.

Wheeldon, L. R. & Konopka, A. E. (2018). Spoken Word Production: Representation, retrieval, and integration. In S. Rueschemeyer & M. G. Gaskell (Eds.) *The Oxford Handbook of Psycholinguistics, Second Edition* (pp. 337-373). Oxford University Press. DOI: 10.1093/oxfordhb/9780198786825.013.15

Wheeldon, L. R. & Lahiri, A. (1997). Prosodic Units in Speech Production. *Journal of Memory and Language*, 37(3), 356-381. https://doi.org/10.1006/jmla.1997.2517

Wheeldon, L. R. & Lahiri, A. (2002). The minimal unit of phonological encoding: prosodic or lexical word. *Cognition*, 85(2), B31-41. https://doi.org/10.1016/S0010-0277(02)00103-8

Wheeldon, L. R. & Monsell, S. (1992). The Locus of Repetition Priming of Spoken Word Production. *The Quarterly Journal of Experimental Psychology*, 44A(4), 723-61. https://doi.org/10.1080/14640749208401307

Wheeldon, L. R. & Smith, M. (2003). Phrase structure priming: A short-lived effect. *Language and Cognitive Processes*, 18(4), 431-442. https://doi.org/10.1080/01690960244000063

Wheeldon, L. R., Smith, M. & Apperly, I. A. (2011). Repeating Words in Sentences: Effects of Sentence Structure. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 37(5), 1051-64. https://doi.org/10.1037/a0023435

APPENDICES

Appendix A: Bilingual Profiling Questionnaire (BPQ)

Part 1

Participant number: ______ **Date of testing:** ______

SCREENING QUESTIONNAIRE

Experimenter: Ask participant the following questions and fill in the yellow boxes with their responses.

1	What is your age? (in years)			
2	What is your gender?			
3	Are you a native speaker of Norwegian?			
4	Is Norwegian the only language you speak at home (aside from English)?		*If no, please specify other home language*	
5	Are you a reasonably good speaker of English?			
6	Do you have normal vision or vision that is corrected to normal with glasses or contact lenses?			
7	Can you confirm that you have no language impairments such as dyslexia, stuttering etc.?			
8	Do you have normal hearing or hearing that is corrected to normal?			
9	Are you left or right handed?			
10	What is your country of birth?			
11	What is your current country of residence?			
12	How many years of education do you have?			
13	What is the highest education level you have? (Select from the drop-down options)		*If other, please specify*	
14	Have you participated in any experiments here before?			

Part 2

2. LANGUAGE BACKGROUND

Participant: please answer these questions below about the different languages you speak.
Please fill in your responses in the appropriate yellow boxes, and ask the experimenter if you have any questions.

Q1 Please list all the languages you speak in order of DOMINANCE (up to 5).

1
2
3
4
5

Q2 Please list all the languages you speak in order of ACQUISITION (up to 5).

1
2
3
4
5

Q3 Please list what percentage of the time you are on average exposed to each language (e.g. exposure in terms of talking, listening, and reading, including TV, films and music).

(All your answers should add up to 100%)

	Language	%
1		
2		
3		
4		
5		
	Total:	0

Please make sure your answer adds up to 100%

Q4 Please list what percentage of the time you spend speaking each language.

(All your answers should add up to 100%)

	Language	%
1		
2		
3		
4		
5		
	Total:	0

Please make sure your answer adds up to 100%

Q5 Please list what percentage of the time you typically spend reading in each language.

(All your answers should add up to 100%)

	Language	%
1		
2		
3		
4		
5		
	Total:	0

Please make sure your answer adds up to 100%

Q6 When choosing a language to speak, with a person who is equally fluent in all your languages, what percentage of time would you choose to speak each language? Please report percentage of total time.

(All your answers should add up to 100%)

	Language	%
1		
2		
3		
4		
5		
	Total:	0

Please make sure your answer adds up to 100%

Q7 What cultures do you identify with (e.g., Norwegian, British, American, etc)? Please list each culture below (up to 5) and use the scale from 0-10 to rate the degree of identification, whereby 0 = no identification, 5 = moderate identification, 10 = complete identification.

	Culture	Identification
1		
2		
3		
4		
5		

Q8 Do you feel that you were once better in one of your languages and that you have become less fluent?

If yes, which one? *And at what age did you become less fluent?*

Q9 In which language do you usually do the following tasks?

Task	Language
Simple maths (count, add)	
Dream	
Express anger or affection	
Talk to yourself	

Part 3

3. NORWEGIAN AND ENGLISH PROFICIENCY

Participant: please answer these questions below about your experience with Norwegian and English.
Please fill in your responses in the appropriate yellow boxes, and ask the experimenter if you have any questions.

Q1 Please list the number of years and months you have spent in each language environment.

	Norwegian		English	
	Years	Months	Years	Months
A country where this language is spoken				
A family where this language is spoken				
A school where this language is spoken ALL of the time				
A school where this language is spoken SOME of the time				
A workplace where this language is spoken ALL of the time				
A workplace where this language is spoken SOME of the time				

Q2 Please rate how much the following factors contributed to your learning of each language on a scale of 0-10 whereby 0 = not a contributor, 5 = moderate contributor and 10 = most important contributor.

	Norwegian	English
Interacting with friends / colleagues		
Interacting with family		
Reading (e.g., books, magazines, online material)		
School and education		
Self-instruction (e.g., language learning videos or apps)		
Watching TV / streaming		
Listening to music/media		

Q3 Please rate to what extent you are currently (e.g. in the last month or so) exposed to each language on a scale of 0-10 whereby 0 = never, 5 = half of the time and 10 = almost always.

	Norwegian	English
Interacting with friends		
Interacting with family		
Reading (e.g., books, magazines, online material)		
Self-instruction (e.g., language learning videos or apps)		
Watching TV / streaming		
Listening to music/media		

Q4 Please rate your level of proficiency in the following aspects of each language on a scale of 0-10 whereby: 0 = none; 1 = very low; 2 = low; 3 = fair; 4 = slightly less than adequate; 5 = adequate; 6 = slightly more than adequate; 7 = good; 8 = very good; 9 = excellent; 10 = perfect.

	Norwegian	English
Speaking (general fluency)		
Pronunciation (accent)		
Listening (understanding spoken language)		
Reading		
Writing		
Grammar		
Vocabulary		
Spelling		

Q5 Please list the AGE (in years) you were when the following occurred for each language.

	Norwegian	English
Started hearing this language on a regular basis		
Became fluent in speaking this language		
Started learning to read in this language		
Became fluent in reading this language		

Q6 Please rate your level of proficiency in switching between your languages when you need to, on a scale of 0-10 whereby: 0 = none; 1 = very low; 2 = low; 3 = fair; 4 = slightly less than adequate; 5 = adequate; 6 = slightly more than adequate; 7 = good; 8 = very good; 9 = excellent; 10 = perfect.

Q7 When you are speaking do you ever find yourself accidentally mixing words or sentences from Norwegian and English?

(a) If yes, how often does English accidentally intrude in your Norwegian on a scale of 0-10 (whereby 0 = never, 5 = half of the time, 10 = all of the time)?

(b) And how often does Norwegian accidentally intrude into your English on a scale of 0-10 (whereby 0 = never, 5 = half of the time, 10 = all of the time)?

Q8 When you are speaking with a person who also knows both Norwegian and English do you ever find yourself intentionally mixing words or sentences from Norwegian and English?

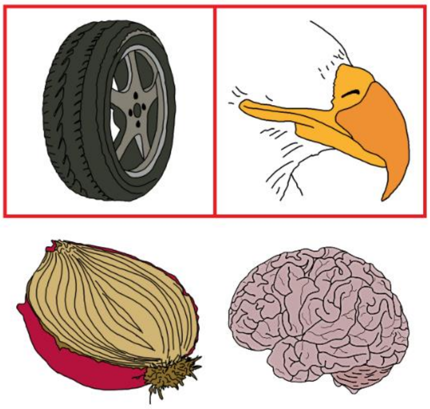
(a) If yes, how often do you intentionally use English words when speaking Norwegian on a scale of 0-10 (whereby 0 = never, 5 = half of the time, 10 = all of the time)?

(b) And how often do you intentionally use Norwegian words when speaking English on a scale of 0-10 (whereby 0 = never, 5 = half of the time, 10 = all of the time)?

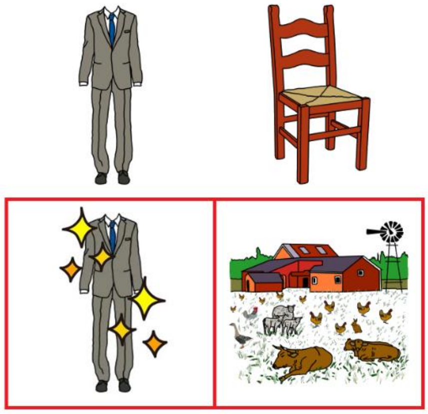
END OF QUESTIONNAIRE - THANK YOU FOR YOUR TIME!

Appendix B: Examples of experimental stimuli and fillers

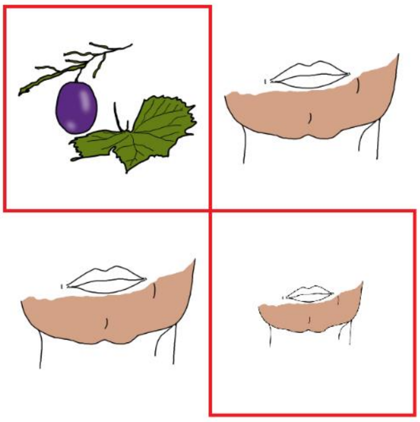
Coordinate structures



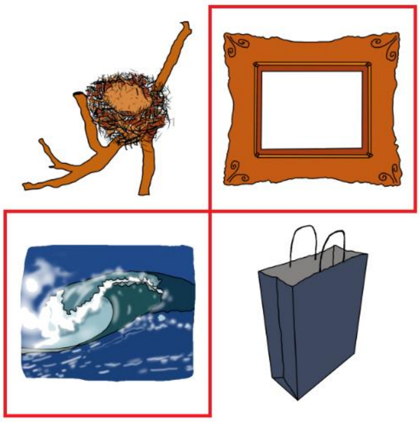
Layout: 1, top horizontal
Sentence type 1: “The tyre and the beak”
Complexity level: zero



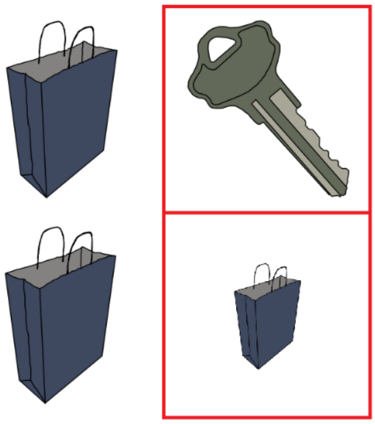
Layout: 2, bottom horizontal
Sentence type 3: “The sparkly suit and the farm”
Complexity level: NP1



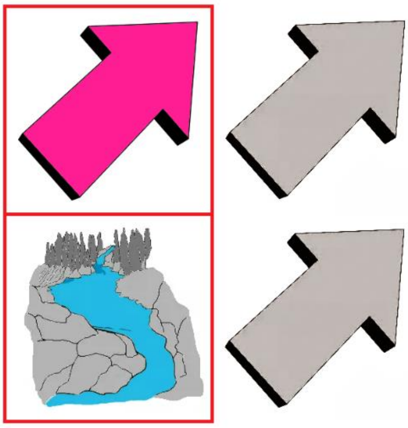
Layout: 3, downward diagonal
Sentence type 5: “The grape and the small chin”
Complexity level: NP2



Layout: 4, upward diagonal
Sentence type 1: “The frame and the wave”
Complexity level: zero

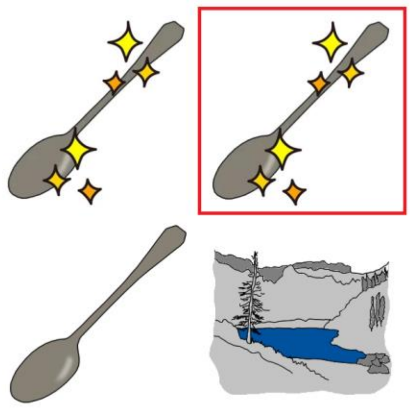
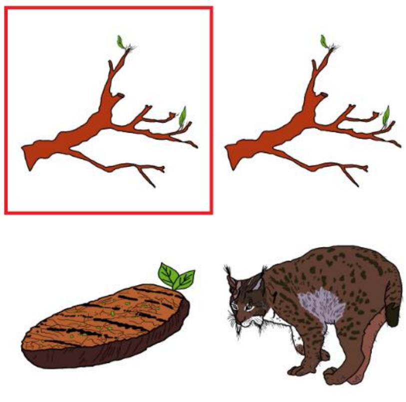
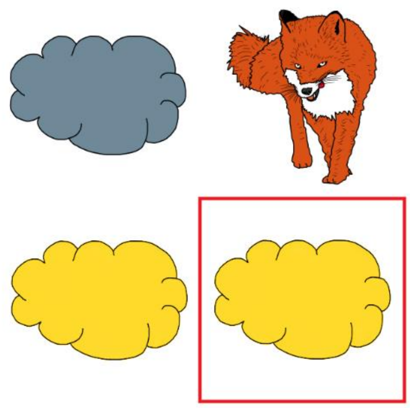
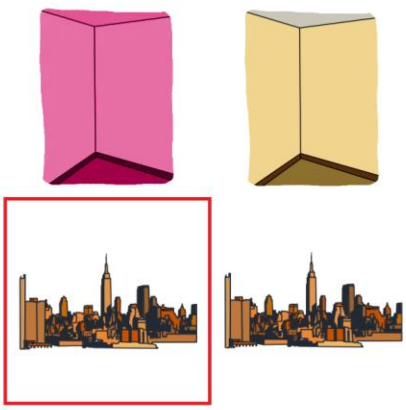


Layout: 5, right vertical
Sentence type 5: “The key and the small bag”
Complexity level: NP2

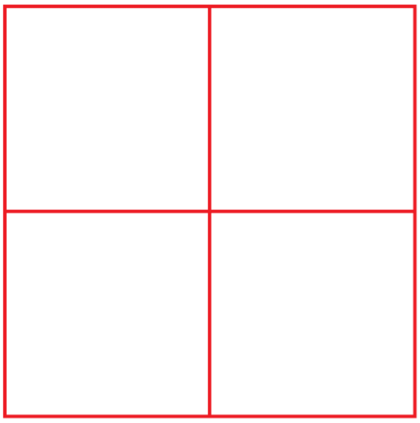
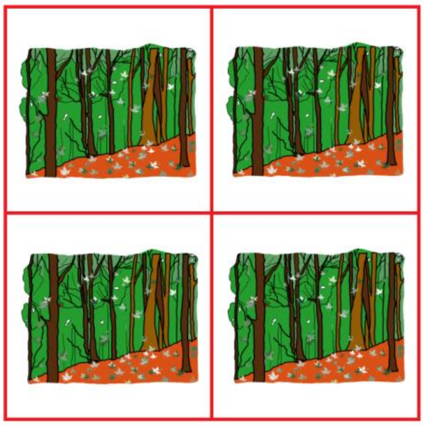
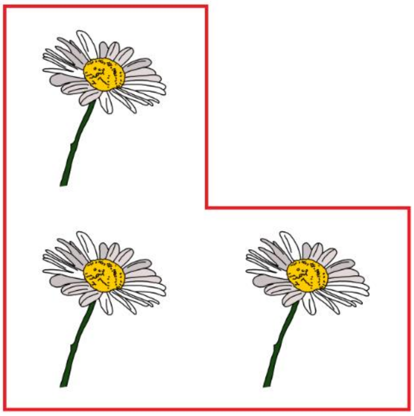
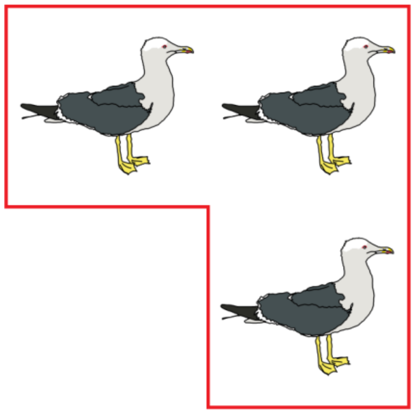


Layout: 6, left vertical
Sentence type 3: “The pink arrow and the river”
Complexity level: NP1

Prepositional structures

 Layout: 7, top right Sentence type 4: “The sparkly spoon above the pond” Complexity level: NP1	 Layout: 8, top left Sentence type 2: “The branch above the steak” Complexity level: zero
 Layout: 9, bottom right Sentence type 4: “The yellow cloud below the fox” Complexity level: NP1	 Layout: 10, bottom left Sentence type 6: “The city below the pink corner” Complexity level: NP2

Fillers

 Filler type 1: “There are no pictures”	 Filler type 2: “All the pictures are the same”
 Filler type 3: “There are three flowers”	 Filler type 3: “There are three seagulls”

Appendix C: Experimental lists (English versions)

English list 1

TrialType	Block	TRIAL	COND	Layout	Item	Sentence	NamingLanguage
	1	1	1	3	19	19.1 the spoon and the pond	EN
	1	2	2	7	6	06.2 the bottle above the pocket	EN
	1	3	3	3	11	11.3 the small desert and the hunter	EN
	1	4	0		100	1.1 there are no pictures	EN
	1	5	2	9	18	18.2 the tray below the shark	EN
	1	6	1	2	13	13.1 the arrow and the river	EN
	1	7	4	8	10	10.4 the pink turkey above the carrot	EN
	1	8	1	2	7	07.1 the cone and the shrimp	EN
	1	9	5	1	3	03.5 the pharmacy and the big wheelbarrow	EN
	1	10	0		101	2.2 all the pictures are the same	EN
	1	11	6	9	8	08.6 the grape below the small chin	EN
	1	12	1	1	1	01.1 the city and the corner	EN
	1	13	0		102	3.2 there are 3 newspapers	EN
	1	14	6	9	14	14.6 the suit below the sparkly farm	EN
	1	15	3	6	5	05.3 the sparkly frame and the wave	EN
	1	16	6	7	20	20.6 the duck above the yellow smoke	EN
	1	17	0		100	1.1 there are no pictures	EN
	1	18	4	9	4	04.4 the yellow hippo below the apron	EN
	1	19	5	1	15	15.5 the cloud and the yellow fox	EN
	1	20	2	7	12	12.2 the rabbit above the barrel	EN
	1	21	0		102	3.1 there are 3 seagulls	EN
	1	22	4	10	16	16.4 the big tyre below the beak	EN
	1	23	3	1	17	17.3 the small branch and the steak	EN
	1	24	0		101	2.1 all the pictures are the same	EN
	1	25	5	2	9	09.5 the lemon and the big trophy	EN
pause	1	26	6	8	2	02.6 the key above the small bag	EN
	2	1	2	7	13	13.2 the arrow above the river	EN
	2	2	3	1	18	18.3 the pink tray and the shark	EN
	2	3	2	10	7	07.2 the cone below the shrimp	EN
	2	4	0		101	2.3 all the pictures are the same	EN
	2	5	4	9	17	17.4 the small branch below the steak	EN
	2	6	3	3	12	12.3 the yellow rabbit and the barrel	EN
	2	7	6	8	15	15.6 the cloud above the yellow fox	EN
	2	8	0		102	3.4 there are 3 fairies	EN
	2	9	2	9	19	19.2 the spoon below the pond	EN
	2	10	5	4	10	10.5 the turkey and the pink carrot	EN
	2	11	4	7	5	05.4 the sparkly frame above the wave	EN
	2	12	0		100	1.1 there are no pictures	EN
	2	13	1	4	14	14.1 the suit and the farm	EN
	2	14	2	7	1	01.2 the city above the corner	EN
	2	15	3	4	6	06.3 the big bottle and the pocket	EN
	2	16	6	10	3	03.6 the pharmacy below the big wheelbarrow	EN
	2	17	1	2	2	02.1 the key and the bag	EN
	2	18	0		100	1.1 there are no pictures	EN
	2	19	5	4	4	04.5 the hippo and the yellow apron	EN
	2	20	4	10	11	11.4 the small desert below the hunter	EN
	2	21	5	2	16	16.5 the tyre and the big beak	EN
	2	22	0		102	3.3 there are 3 flowers	EN
	2	23	1	2	20	20.1 the duck and the smoke	EN
	2	24	6	8	9	09.6 the lemon above the big trophy	EN
	2	25	0		101	2.4 all the pictures are the same	EN
pause	2	26	1	3	8	08.1 the grape and the chin	EN
	3	1	2	9	20	20.2 the duck below the smoke	EN
	3	2	3	5	7	07.3 the sparkly cone and the shrimp	EN
	3	3	4	8	18	18.4 the pink tray above the shark	EN
	3	4	0		101	2.6 all the pictures are the same	EN
	3	5	5	6	17	17.5 the branch and the small steak	EN
	3	6	6	7	4	04.6 the hippo above the yellow apron	EN

	3	7	3	3	1	01.3 the pink city and the corner	EN
	3	8	0		102	3.5 there are 3 lions	EN
	3	9	4	10	12	12.4 the yellow rabbit below the barrel	EN
	3	10	1	6	9	09.1 the lemon and the trophy	EN
	3	11	6	9	10	10.6 the turkey below the pink carrot	EN
	3	12	3	6	13	13.3 the pink arrow and the river	EN
	3	13	2	8	8	08.2 the grape above the chin	EN
	3	14	5	1	11	11.5 the desert and the small hunter	EN
	3	15	0		100	1.1 there are no pictures	EN
	3	16	1	6	3	03.1 the pharmacy and the wheelbarrow	EN
	3	17	4	9	6	06.4 the big bottle below the pocket	EN
	3	18	0		101	2.5 all the pictures are the same	EN
	3	19	2	10	2	02.2 the key below the bag	EN
	3	20	1	5	15	15.1 the cloud and the fox	EN
	3	21	3	6	19	19.3 the sparkly spoon and the pond	EN
	3	22	0		100	1.1 there are no pictures	EN
	3	23	5	3	5	05.5 the frame and the sparkly wave	EN
	3	24	6	8	16	16.6 the tyre above the big beak	EN
	3	25	0		102	3.6 there are 3 forks	EN
pause	3	26	2	7	14	14.2 the suit above the farm	EN
	4	1	5	5	18	18.5 the tray and the pink shark	EN
	4	2	4	9	13	13.4 the pink arrow below the river	EN
	4	3	0		100	1.1 there are no pictures	EN
	4	4	5	5	12	12.5 the rabbit and the yellow barrel	EN
	4	5	4	9	7	07.4 the sparkly cone below the shrimp	EN
	4	6	1	3	4	04.1 the hippo and the apron	EN
	4	7	0		101	2.8 all the pictures are the same	EN
	4	8	2	7	3	03.2 the pharmacy above the wheelbarrow	EN
	4	9	3	3	20	20.3 the yellow duck and the smoke	EN
	4	10	2	10	15	15.2 the cloud below the fox	EN
	4	11	0		102	3.7 there are 3 diamonds	EN
	4	12	6	8	11	11.6 the desert above the small hunter	EN
	4	13	1	1	16	16.1 the tyre and the beak	EN
	4	14	0		100	1.1 there are no pictures	EN
	4	15	3	2	8	08.3 the small grape and the chin	EN
	4	16	2	9	9	09.2 the lemon below the trophy	EN
	4	17	3	4	2	02.3 the small key and the bag	EN
	4	18	0		101	2.7 all the pictures are the same	EN
	4	19	5	6	6	06.5 the bottle and the big pocket	EN
	4	20	6	7	17	17.6 the branch above the small steak	EN
	4	21	3	2	14	14.3 the sparkly suit and the farm	EN
	4	22	4	8	1	01.4 the pink city above the corner	EN
	4	23	0		102	3.8 there are 3 horses	EN
	4	24	6	10	5	05.6 the frame below the sparkly wave	EN
	4	25	1	1	10	10.1 the turkey and the carrot	EN
pause	4	26	4	7	19	19.4 the sparkly spoon above the pond	EN
	5	1	1	4	11	11.1 the desert and the hunter	EN
	5	2	2	9	16	16.2 the tyre below the beak	EN
	5	3	0		101	2.9 all the pictures are the same	EN
	5	4	5	3	13	13.5 the arrow and the pink river	EN
	5	5	6	8	12	12.6 the rabbit above the yellow barrel	EN
	5	6	3	1	9	09.3 the big lemon and the trophy	EN
	5	7	0		100	1.1 there are no pictures	EN
	5	8	5	4	7	07.5 the cone and the sparkly shrimp	EN
	5	9	2	10	10	10.2 the turkey below the carrot	EN
	5	11	4	7	8	08.4 the small grape above the chin	EN
	5	12	0		102	3.10 there are 3 balloons	EN
	5	13	4	8	2	02.4 the small key above the bag	EN
	5	14	1	4	5	05.1 the frame and the wave	EN
	5	15	4	10	14	14.4 the sparkly suit below the farm	EN
	5	16	0		101	2.10 all the pictures are the same	EN
	5	17	6	8	18	18.6 the tray above the pink shark	EN
	5	18	1	3	17	17.1 the branch and the steak	EN
	5	19	4	10	20	20.4 the yellow duck below the smoke	EN

	5	20	5	4	19	19.5 the spoon and the sparkly pond	EN
	5	20	3	1	3	03.3 the big pharmacy and the wheelbarrow	EN
	5	21	0		100	1.1 there are no pictures	EN
	5	22	2	9	4	04.2 the hippo below the apron	EN
	5	23	3	4	15	15.3 the yellow cloud and the fox	EN
	5	24	6	7	6	06.6 the bottle above the big pocket	EN
	5	25	0		102	3.9 there are 3 stars	EN
pause	5	26	5	2	1	01.5 the city and the pink corner	EN
	6	1	6	9	13	13.6 the arrow below the pink river	EN
	6	2	3	5	16	16.3 the big tyre and the beak	EN
	6	3	2	7	5	05.2 the frame above the wave	EN
	6	4	1	6	18	18.1 the tray and the shark	EN
	6	5	0		100	1.1 there are no pictures	EN
	6	6	4	10	9	09.4 the big lemon below the trophy	EN
	6	7	5	6	14	14.5 the suit and the sparkly farm	EN
	6	8	0		102	3.12 there are 3 swords	EN
	6	9	4	9	3	03.4 the big pharmacy below the wheelbarrov	EN
	6	10	1	5	12	12.1 the rabbit and the barrel	EN
	6	11	6	7	19	19.6 the spoon above the sparkly pond	EN
	6	12	3	2	4	04.3 the yellow hippo and the apron	EN
	6	13	0		100	1.1 there are no pictures	EN
	6	14	2	8	17	17.2 the branch above the steak	EN
	6	15	5	3	8	08.5 the grape and the small chin	EN
	6	16	0		101	2.12 all the pictures are the same	EN
	6	17	6	10	1	01.6 the city below the pink corner	EN
	6	18	3	5	10	10.3 the pink turkey and the carrot	EN
	6	19	2	7	11	11.2 the desert above the hunter	EN
	6	20	0		102	3.11 there are 3 hands	EN
	6	21	1	5	6	06.1 the bottle and the pocket	EN
	6	22	4	9	15	15.4 the yellow cloud below the fox	EN
	6	23	5	5	2	02.5 the key and the small bag	EN
	6	24	6	8	7	07.6 the cone above the sparkly shrimp	EN
	6	25	0		101	2.11 all the pictures are the same	EN
	6	26	5	1	20	20.5 the duck and the yellow smoke	EN

English list 2

TrialType	Block	TRIAL	COND	Layout	Item	Sentence	NamingLanguage
	1	1	1	6	19	19.1 the mushroom and the candle	EN
	1	2	6	9	2	02.6 the sheep below the small map	EN
	1	3	0		101	2.1 all the pictures are the same	EN
	1	4	5	2	3	03.5 the onion and the big basket	EN
	1	5	2	7	12	12.2 the hedgehog above the mermaid	EN
	1	6	3	6	17	17.3 the yellow plane and the screen	EN
	1	7	6	10	20	20.6 the leaf below the sparkly fence	EN
	1	8	0		102	3.1 there are 3 seagulls	EN
	1	9	6	9	8	08.6 the witch below the sparkly coin	EN
	1	10	1	2	1	01.1 the drawer and the cabin	EN
	1	11	4	7	16	16.4 the sparkly chair above the meat	EN
	1	12	1	3	13	13.1 the road and the bed	EN
	1	13	2	10	18	18.2 the butterfly below the strawberry	EN
	1	14	3	2	5	05.3 the big umbrella and the pineapple	EN
	1	15	0		100	1.1 there are no pictures	EN
	1	16	3	6	11	11.3 the pink pig and the brain	EN
	1	17	6	8	14	14.6 the goal above the big cheese	EN
	1	18	0		101	2.2 all the pictures are the same	EN
	1	19	4	9	10	10.4 the yellow car below the dog	EN
	1	20	5	2	15	15.5 the bird and the yellow leg	EN
	1	21	2	8	6	06.2 the blanket above the mirror	EN
	1	22	0		102	3.2 there are 3 newspapers	EN
	1	23	5	1	9	09.5 the ant and the sparkly cage	EN
	1	24	4	8	4	04.4 the small squirrel above the pumpkin	EN
	1	25	1	2	7	07.1 the suitcase and the pencil	EN
pause	1	26	0		100	1.1 there are no pictures	EN
	2	1	6	10	3	03.6 the onion below the big basket	EN
	2	2	3	6	6	06.3 the pink blanket and the mirror	EN
	2	3	6	7	15	15.6 the bird above the yellow leg	EN
	2	4	0		102	3.3 there are 3 flowers	EN
	2	5	2	10	19	19.2 the mushroom below the candle	EN
	2	6	1	3	20	20.1 the leaf and the fence	EN
	2	7	4	8	11	11.4 the pink pig above the brain	EN
	2	8	0		101	2.4 all the pictures are the same	EN
	2	9	5	6	16	16.5 the chair and the sparkly meat	EN
	2	10	4	9	5	05.4 the big umbrella below the pineapple	EN
	2	11	0		100	1.1 there are no pictures	EN
	2	12	1	5	2	02.1 the sheep and the map	EN
	2	13	2	8	13	13.2 the road above the bed	EN
	2	14	3	2	18	18.3 the yellow butterfly and the strawberry	EN
	2	15	0		102	3.4 there are 3 fairies	EN
	2	16	1	6	8	08.1 the witch and the coin	EN
	2	17	6	9	9	09.6 the ant below the sparkly cage	EN
	2	18	3	2	12	12.3 the pink hedgehog and the mermaid	EN
	2	19	2	7	7	07.2 the suitcase above the pencil	EN
	2	20	0		100	1.1 there are no pictures	EN
	2	21	5	1	10	10.5 the car and the yellow dog	EN
	2	22	2	10	1	01.2 the drawer below the cabin	EN
	2	23	0		101	2.3 all the pictures are the same	EN
	2	24	1	2	14	14.1 the goal and the cheese	EN
	2	25	4	8	17	17.4 the yellow plane above the screen	EN
pause	2	26	5	3	4	04.5 the squirrel and the small pumpkin	EN
	3	1	0		102	3.5 there are 3 lions	EN
	3	2	1	3	3	03.1 the onion and the basket	EN
	3	3	4	9	18	18.4 the yellow butterfly below the strawberry	EN
	3	4	1	1	15	15.1 the bird and the leg	EN
	3	5	4	7	6	06.4 the pink blanket above the mirror	EN
	3	6	0		101	2.6 all the pictures are the same	EN
	3	7	3	4	13	13.3 the small road and the bed	EN
	3	8	2	9	20	20.2 the leaf below the fence	EN

	3	9	3	5	19	19.3 the big mushroom and the candle	EN
	3	10	0		100	1.1 there are no pictures	EN
	3	11	2	8	2	02.2 the sheep above the map	EN
	3	12	3	3	7	07.3 the pink suitcase and the pencil	EN
	3	13	6	10	16	16.6 the chair below the sparkly meat	EN
	3	14	1	2	9	09.1 the ant and the cage	EN
	3	15	0		102	3.6 there are 3 forks	EN
	3	16	3	3	1	01.3 the small drawer and the cabin	EN
	3	17	6	8	4	04.6 the squirrel above the small pumpkin	EN
	3	18	5	5	5	05.5 the umbrella and the big pineapple	EN
	3	19	0		101	2.5 all the pictures are the same	EN
	3	20	5	1	17	17.5 the plane and the yellow screen	EN
	3	21	6	7	10	10.6 the car above the yellow dog	EN
	3	22	4	10	12	12.4 the pink hedgehog below the mermaid	EN
	3	23	5	4	11	11.5 the pig and the pink brain	EN
	3	24	2	7	14	14.2 the goal above the cheese	EN
	3	25	0		100	1.1 there are no pictures	EN
pause	3	26	2	9	8	08.2 the witch below the coin	EN
	4	1	5	6	18	18.5 the butterfly and the yellow strawberry	EN
	4	2	0		101	2.8 all the pictures are the same	EN
	4	3	6	10	17	17.6 the plane below the yellow screen	EN
	4	4	3	1	20	20.3 the sparkly leaf and the fence	EN
	4	5	2	8	9	09.2 the ant above the cage	EN
	4	6	5	2	6	06.5 the blanket and the pink mirror	EN
	4	7	0		100	1.1 there are no pictures	EN
	4	8	5	3	12	12.5 the hedgehog and the pink mermaid	EN
	4	9	4	7	19	19.4 the big mushroom above the candle	EN
	4	10	1	5	16	16.1 the chair and the meat	EN
	4	11	0		102	3.7 there are 3 diamonds	EN
	4	12	3	5	14	14.3 the big goal and the cheese	EN
	4	13	6	9	11	11.6 the pig below the pink brain	EN
	4	14	4	8	7	07.4 the pink suitcase above the pencil	EN
	4	15	3	1	2	02.3 the small sheep and the map	EN
	4	16	0		101	2.7 all the pictures are the same	EN
	4	17	4	8	1	01.4 the small drawer above the cabin	EN
	4	18	1	3	4	04.1 the squirrel and the pumpkin	EN
	4	19	6	10	5	05.6 the umbrella below the big pineapple	EN
	4	20	3	1	8	08.3 the sparkly witch and the coin	EN
	4	21	0		100	1.1 there are no pictures	EN
	4	22	2	7	3	03.2 the onion above the basket	EN
	4	23	1	4	10	10.1 the car and the dog	EN
	4	24	4	9	13	13.4 the small road below the bed	EN
	4	25	0		102	3.8 there are 3 horses	EN
pause	4	26	2	10	15	15.2 the bird below the leg	EN
	5	1	1	1	5	05.1 the umbrella and the pineapple	EN
	5	2	4	9	2	02.4 the small sheep below the map	EN
	5	3	1	4	17	17.1 the plane and the screen	EN
	5	4	0		102	3.10 there are 3 balloons	EN
	5	5	4	8	8	08.4 the sparkly witch above the coin	EN
	5	6	5	5	13	13.5 the road and the small bed	EN
	5	7	0		100	1.1 there are no pictures	EN
	5	8	3	4	3	03.3 the big onion and the basket	EN
	5	9	2	10	10	10.2 the car below the dog	EN
	5	10	1	1	11	11.1 the pig and the brain	EN
	5	11	0		101	2.10 all the pictures are the same	EN
	5	12	5	4	19	19.5 the mushroom and the big candle	EN
	5	13	6	7	18	18.6 the butterfly above the yellow strawberry	EN
	5	14	0		102	3.9 there are 3 stars	EN
	5	15	2	9	4	04.2 the squirrel below the pumpkin	EN
	5	16	3	5	9	09.3 the sparkly ant and the cage	EN
	5	17	2	8	16	16.2 the chair above the meat	EN
	5	18	3	3	15	15.3 the yellow bird and the leg	EN
	5	19	0		100	1.1 there are no pictures	EN

	5	20	6	10	6	06.6 the blanket below the pink mirror	EN
	5	21	5	4	7	07.5 the suitcase and the pink pencil	EN
	5	22	4	7	20	20.4 the sparkly leaf above the fence	EN
	5	23	5	6	1	01.5 the drawer and the small cabin	EN
	5	24	6	7	12	12.6 the hedgehog above the pink mermaid	EN
	5	25	0		101	2.9 all the pictures are the same	EN
pause	5	26	4	10	14	14.4 the big goal below the cheese	EN
	6	1	6	8	13	13.6 the road above the small bed	EN
	6	2	3	4	4	04.3 the small squirrel and the pumpkin	EN
	6	3	4	7	15	15.4 the yellow bird above the leg	EN
	6	4	0		101	2.11 all the pictures are the same	EN
	6	5	1	6	12	12.1 the hedgehog and the mermaid	EN
	6	6	6	10	19	19.6 the mushroom below the big candle	EN
	6	7	3	1	16	16.3 the sparkly chair and the meat	EN
	6	8	0		102	3.11 there are 3 hands	EN
	6	9	2	9	11	11.2 the pig below the brain	EN
	6	10	1	4	6	06.1 the blanket and the mirror	EN
	6	11	4	8	3	03.4 the big onion above the basket	EN
	6	12	6	9	1	01.6 the drawer below the small cabin	EN
	6	13	0		100	1.1 there are no pictures	EN
	6	14	4	10	9	09.4 the sparkly ant below the cage	EN
	6	15	5	6	14	14.5 the goal and the big cheese	EN
	6	16	2	8	5	05.2 the umbrella above the pineapple	EN
	6	17	3	2	10	10.3 the yellow car and the dog	EN
	6	18	0		102	3.12 there are 3 swords	EN
	6	19	5	3	8	08.5 the witch and the sparkly coin	EN
	6	20	6	10	7	07.6 the suitcase below the pink pencil	EN
	6	21	0		100	1.1 there are no pictures	EN
	6	22	5	6	20	20.5 the leaf and the sparkly fence	EN
	6	23	2	7	17	17.2 the plane above the screen	EN
	6	24	5	4	2	02.5 the sheep and the small map	EN
	6	25	0		101	2.12 all the pictures are the same	EN
	6	26	1	5	18	18.1 the butterfly and the strawberry	EN

Appendix D: List of trial picture names

Word EN	Word NO	S. length EN	S. length NO	P. length EN	P. length NO	EN frequency	NO frequency	Assignment
ant	maur	1	1	3	3	3.95	3.28	English List 1
apron	forkle	2	2	5	6	3.48	2.53	English List 2
arrow	pil	2	1	3	3	3.78	3.4	English List 2
bag	pose	1	2	3	4	4.89	3.86	English List 2
barrel	tønne	2	2	5	4	4.06	3.37	English List 2
basket	kurv	2	1	6	4	4.16	3.23	English List 1
beak	nebb	1	1	3	3	3.83	2.93	English List 2
bed	seng	1	1	3	3	5.12	4.08	English List 1
bird	fugl	1	1	3.4	3	4.85	4.33	English List 1
blanket	teppe	2	2	7	4	3.93	3.59	English List 1
bottle	flaske	2	2	4	6	4.65	4.08	English List 2
brain	hjerne	1	2	4	4.5	4.84	3.75	English List 1
branch	grein	1	1	5	4	4.1	3.41	English List 2
butterfly	sommerfugl	3	3	8	8	4.03	3.08	English List 1
cabin	hytte	2	2	5	4	3.9	4.23	English List 1
cage	bur	1	1	3	3	4.03	4.02	English List 1
candle	lys	2	1	5	3	3.79	N/A	English List 1
car	bil	1	1	3.2	3	5.44	4.99	English List 1
carrot	gulrot	2	2	5	6.7	4.08	3.54	English List 2
chair	stol	1	1	2.3	4	4.66	3.95	English List 1
cheese	ost	1	1	3	3	4.81	3.97	English List 1
chin	hake	1	2	3	4	3.93	2.6	English List 2
city	by	2	1	4	2	5.4	5.36	English List 2
cloud	sky	1	1	4	2	4.77	3.17	English List 2
coin	mynt	1	1	3	4	4.05	3.99	English List 1
cone	kjegle	2	2	3	5	3.67	2.3	English List 2
corner	hjørne	2	2	4.6	4.5	4.96	4.02	English List 2
desert	ørken	2	2	5.6	5	4.3	3.2	English List 2
dog	hund	1	1	3	3	5.17	4.88	English List 1
drawer	skuff	1	1	5	4	3.98	3.22	English List 1
duck	and	1	1	3	2	4.53	5.71	English List 2
farm	gård	1	1	3.4	3	4.81	4.26	English List 2
fence	gjerde	1	2	4	4	4.31	3.44	English List 1
fox	rev	1	1	4	3	4.46	3.72	English List 2
frame	ramme	1	2	4	4	4.76	4.03	English List 2
goal	mål	1	1	3	3	4.97	5.41	English List 1
grape	drue	1	2	4	4	3.54	2.76	English List 2
hedgehog	pinnsvin	2	2	6	7	3.72	3	English List 1
hippo	flodhest	2	2	4	8	3.58	2.41	English List 2
hunter	jeger	2	2	5.6	5	4.18	3.9	English List 2
key	nøkkel	1	2	2	5	5.07	3.65	English List 2
leaf	blad	1	1	3	3	4.29	3.97	English List 1
leg	bein	1	1	3	3	4.88	4.49	English List 1
lemon	sitron	2	2	5	6	4.46	2.72	English List 2
map	kart	1	1	3	3.4	4.52	4.41	English List 1
meat	kjøtt	1	1	3	3	4.8	4.38	English List 1
mermaid	havfrue	2	3	7	7	3.41	2.45	English List 1
mirror	speil	2	1	4.5	4	4.45	3.81	English List 1
mushroom	sopp	2	1	6	3	3.87	3.79	English List 1
onion	løk	2	1	5	3	4.28	3.83	English List 1
pencil	blyant	2	2	6	6	3.98	3.2	English List 1
pharmacy	apotek	3	3	7	6	3.44	4.07	English List 2
pig	gris	1	1	3	4	4.51	3.72	English List 1
pineapple	ananas	3	3	6	6	3.84	3.15	English List 1
plane	fly	1	1	4	3	4.58	4.7	English List 1
pocket	lomme	2	2	5	4	4.63	3.43	English List 2
pond	dam	1	1	4	3	4.2	3.34	English List 2
pumpkin	gresskar	2	2	7	7	3.79	2.91	English List 1
rabbit	kanin	2	2	5	5	4.39	3.73	English List 2
river	elv	2	1	4.5	3	4.93	3.33	English List 2
road	vei	1	1	3	2	5.28	5.26	English List 1
screen	skjerm	1	1	5	4	4.61	4.44	English List 1
shark	hai	1	1	4	2	4.17	4.14	English List 2
sheep	sau	1	1	3	2	4.67	3.86	English List 1
shrimp	reke	1	2	5	4	3.63	3.31	English List 2
smoke	røyk	1	1	4	3	4.53	3.99	English List 2

spoon	skje	1	1	4	2.3	4.29	3.58	English List 2
squirrel	ekorn	2	2	7	4.5	3.94	3.04	English List 1
steak	biff	1	1	4	3	4.05	3.6	English List 2
strawberry	jordbær	3	2	7	6	3.99	3.71	English List 1
suit	dress	1	1	3	4	4.59	3.69	English List 2
suitcase	koffert	2	2	6	5.6	3.78	3.34	English List 1
tie	slips	1	1	2	5	4.58	3.33	English List 1
tray	brett	1	1	3	4	4.15	3.26	English List 2
trophy	pokal	2	2	5	5	4.37	3.01	English List 2
turkey	kalkun	2	2	5	6	4.35	3.14	English List 2
tyre	dekk	2	1	4	3	3.83	4.11	English List 2
umbrella	paraply	3	3	7	7	3.89	3.26	English List 1
wave	bølge	1	2	3	5	4.55	2.99	English List 2
wheelbarrow	trillebår	3	3	7	8	3.39	2.56	English List 2
witch	heks	1	1	3	4	4.06	3.21	English List 1

Note: frequencies calculated as Zipf-values

Appendix E: Instructions booklet (English version)

Sentence Production Experiment

Thank you for participating in our study of bilingual language processing. In this sentence production experiment, you will be shown a grid of pictures on the screen and asked to name one or two of them.

For the upcoming experiment, you will speak in English only.

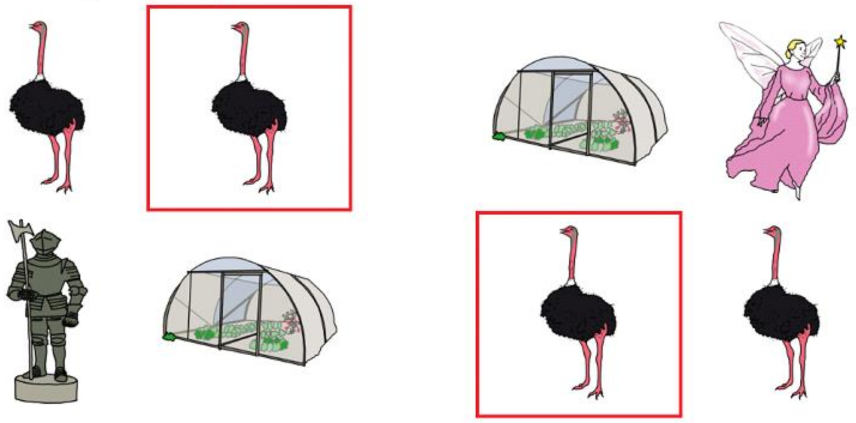
We ask you to **uniquely identify the framed pictures**.
This means that you must describe the pictures in such a way that the description could not apply to any of the other pictures in the grid.
You must also describe the pictures in relation to the other pictures, and *not* in terms of spatial locations, such as “in the top right corner.”
If there are more than one framed picture, the order in which they are to be named is **top to bottom, left to right**.

We would like you to produce different kinds of descriptions or sentences, depending on the kind of picture you see. The exact nouns to be used and the placement of the pictures will vary from one picture grid to another.

Sometimes, two pictures will have a red frame around them, like in these pictures:

All of these pictures would be “the ostrich and the greenhouse,” even though the layout is very different from one picture to another.

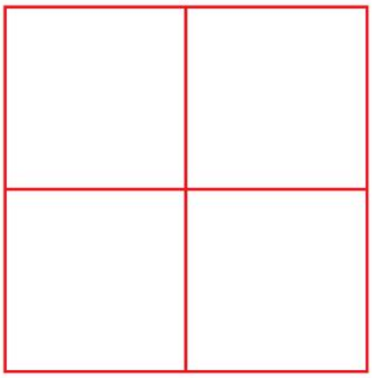
Other times, only one picture will be framed:



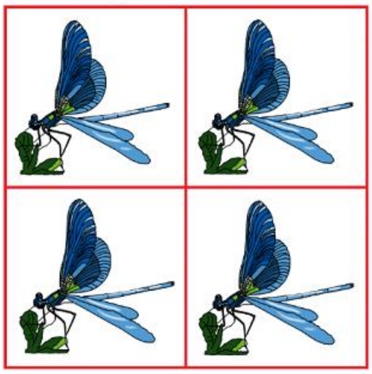
In these cases, there will be another identical picture within the grid, so it will be necessary to define the framed picture in relation to the picture above or below. So, it will be "the ostrich above the greenhouse" or "the ostrich below the greenhouse."

Examples of other picture types:

"There are no pictures"



"All the pictures are the same"



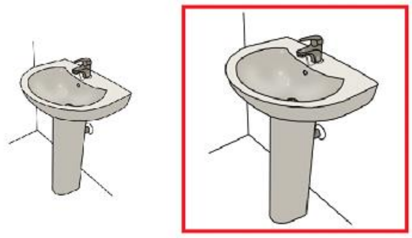
"There are three parachutes/snails etc."

Adjectives

Sometimes, there will be several pictures of the same type within the grid. You must then use an adjective to distinguish the picture within the red frame, e.g. 'small.'

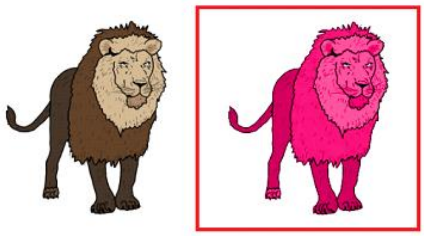
For example, the description here would be: "the small ostrich and the greenhouse."

These are examples of the adjectives you will need to use:

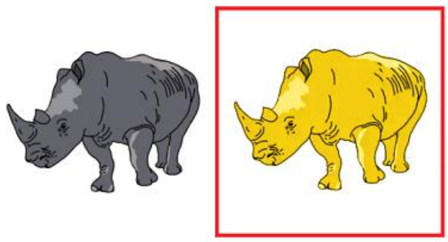


'big'

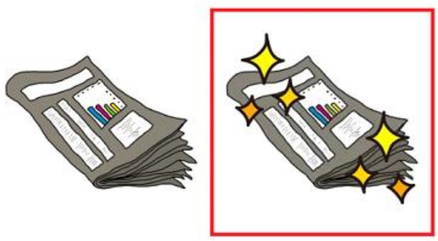
'small'



'pink'



'yellow'



'sparkly'

Appendix F: Variables removed from the factor analysis

Removed variable	*Reason for removal*
‘Q8b Intentional mixing NO to EN’	Insufficient co-variation
‘Q3 Time exposed to L1’	> 0.9 correlation with ‘Q3 Time exposed to L2’
‘Q4 Time spent speaking L1’	> 0.9 correlation with ‘Q4 Time spent speaking L2’
‘Q5 Time spent reading L1’	> 0.9 correlation with ‘Q5 Time spent reading L2’
‘Q6 Choosing to speak L1’	> 0.9 correlation with ‘Q6 Choosing to speak L1’
‘Q4 EN proficiency: spelling’	> 0.8 correlation with ‘Q4 EN proficiency: grammar’
‘Q2 NO learning: music’	> 0.8 correlation with ‘Q2 NO learning: TV/streaming’
‘Q3 NO exposure: music’	> 0.8 correlation with ‘Q3 NO exposure: TV/streaming’

Appendix G: Data loss and mean error rate by participant

Data loss per participant:

Participant no.	1	3	4	6	8	10	12	13	15	16	17	21	22	23	24	26	29	30	39	50	51	53	54
Data loss (%)	57	15	29	29	16	2	61	5	1	13	3	2	48	14	11	2	66	1	16	21	1	7	4

Mean error rate per participant:

Participant no.	1	3	4	6	8	9	10	11	12
Mean error rate	21.868852	26.222222	29.383886	30.331754	16.964286	12.5	10.084034	43.75	10.614525
Participant no.	13	15	16	17	19	20	21	22	23
Mean error rate	9.787234	41.841004	32.142857	19.915254	22.5	19.166667	37.815126	21.052632	17.333333
Participant no.	24	26	29	30	39	50	51	53	54
Mean error rate	51.528384	41.772152	62.068966	16.317992	11.160714	41.552511	14.225941	49.568966	19.915254