

Beyond the tried and true: How virtual reality, dialog setups, and a focus on multimodality can take bilingual language production research forward

Alex Titus^{a,b,*}, Ton Dijkstra^c, Roel M. Willems^{a,1}, David Peeters^d

^a Radboud University, Centre for Language Studies, Nijmegen, the Netherlands

^b Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands

^c Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, the Netherlands

^d Tilburg University, Department of Communication and Cognition, TiCC, Tilburg, the Netherlands

ARTICLE INFO

Keywords:

Bilingualism
Virtual reality
Dialog
Multimodal communication
Ecological validity

ABSTRACT

Bilinguals possess the ability of expressing themselves in more than one language, and typically do so in contextually rich and dynamic settings. Theories and models have indeed long considered context factors to affect bilingual language production in many ways. However, most experimental studies in this domain have failed to fully incorporate linguistic, social, or physical context aspects, let alone combine them in the same study. Indeed, most experimental psycholinguistic research has taken place in isolated and constrained lab settings with carefully selected words or sentences, rather than under rich and naturalistic conditions. We argue that the most influential experimental paradigms in the psycholinguistic study of bilingual language production fall short of capturing the effects of context on language processing and control presupposed by prominent models. This paper therefore aims to enrich the methodological basis for investigating context aspects in current experimental paradigms and thereby move the field of bilingual language production research forward theoretically. After considering extensions of existing paradigms proposed to address context effects, we present three far-ranging innovative proposals, focusing on virtual reality, dialog situations, and multimodality in the context of bilingual language production.

1. Introduction

Imagine that you and a friend recently went on a trip abroad to London. It is not unlikely that you, like most people in the world, learned a second language such as English in school and would be able to interact with the local population (Grosjean, 1989). Visiting a market place, you may have said to your friend in your first language (L1) that you wished to buy some healthy snacks, then turning to the owner of a fruit and vegetable stand speaking in your non-dominant second language (L2) English, while pointing at the oranges at hand. Monolingual bystanders may have been amazed that you seemed to change languages effortlessly, as if you were simply flipping a mental switch. However, already a long time ago, research has shown that, at a cognitive level, switching between languages is actually quite a complicated matter (Kolers, 1966; Penfield and Roberts, 1959). It has become clear that many different aspects of a situation influence how bilinguals control their languages for production.

Indeed, from a general perspective, in daily interactions like the above, at least three types of contextual factor can be discerned. First, there are *linguistic* context aspects, such as which exact languages are spoken, the words and grammatical structures they provide to express one's thoughts, and the non-verbal communicative signals they are combined with. Second, there are *social* context aspects, such as the relationship between the interlocutors and the knowledge they have or assume of each other's language background and relative language proficiency. This social relation will clearly be different between you and your friend compared to you and the owner of the stand. Third, there is the *non-linguistic* physical context of the conversation, such as the market place in the example above and the context-appropriate language(s) it allows for.

The linguistic, social, and non-linguistic physical context all subtly affect which language a bilingual will use in a situation and whether a switch between languages will take place, within and across sentences, interlocutors, and settings. In fact, all these context aspects may

* Corresponding author. Center for Language Studies Radboud University P.O. Box 9102 NL-6500 HC, Nijmegen, The Netherlands.
E-mail address: alex.titus@ru.nl (A. Titus).

¹ Present address: The Netherlands Institute for Social Research (SCP), Den Haag, the Netherlands.

interactively influence the speakers' word choice and even their accent, register, or dialect. In light of this complexity, psycholinguistic researchers in the field of bilingual language production have begun to wonder how they should investigate these common daily-life situations of context-sensitive language production and language switching that, in spite of their complexity, appear to take place relatively flawlessly and with minimal apparent effort outside the lab.

So far, however, the use of context-sensitive bilingual language in naturalistic environments has remained under-investigated in psycholinguistics. For decades, researchers studying the psychology and neurobiology of language have instead concentrated on developing refined models of individual word and sentence production, basing themselves primarily on experimental studies conducted in well-controlled laboratory setups. The influence of different types of context on bilingual language production may have remained relatively underexplored in this field because of the methodological complexity that incorporating context brings. Indeed, the default (fairly unidimensional) environment in which experimental research typically occurs is quite distant in nature from rich and dynamic (multidimensional) real-life situations. Moreover, scientific practice in the domain of experimental psychology is typically restricted to manipulating a single or a couple of factors at a time (Donders, 1969; Roelofs, 2018b). The tacit assumption is that a thorough understanding of the constituent pieces of the overall situation should suffice to eventually lead to an integration and understanding of the whole.

In light of the above, finding a balance between experimental control and a high degree of real-life context within an experimental study has therefore sometimes been considered hard, unnecessary, or even impossible to achieve. It has been commonly assumed that an experiment must sacrifice some ecological validity to attain the degree of control required to collect reliable data (cf. Peeters, 2019). Nevertheless, the inherent risk of this approach is that we are completing a jigsaw that pictures what happens in the mind of a bilingual speaker in the lab, but not necessarily corresponds to what happens in real life situations (cf. Myers-Scotton, 2006). Therefore, given the considerable depth of our current understanding of the bilingual language production system based on existing laboratory studies, in the study of the (neuro)psychology of bilingual language use we should now focus on how contextual factors affect this system in daily life.

To assess the present research status, we will begin this review article by discussing leading psycholinguistic models in bilingual language production research. Interestingly, these models attribute a central role to the multidimensional context in which bilingual language production typically takes place. Next, we will see how the most popular experimental paradigms developed for laboratory research actually restrict further fine-tuning and testing of these theories and models of bilingual natural language use, in the sense that context is typically not given center stage. We therefore need new experimental paradigms that better reflect dynamic everyday situations. After having considered different existing adaptations that contextually enrich standard paradigms, we will propose three task-related and methodological advancements that should enable us to study face-to-face bilingual interactions more holistically, and under more natural conditions, in the lab. As a complementary and different way of studying bilingual language production, these advances should allow for taking the field of bilingual language production research forward theoretically.

We restrict the scope of this article to the psycholinguistic and neuropsychological study of bilingual language production and control. As such, we do not cover the vast literature on bilingual language comprehension in this article, although we hope that the two domains (production and comprehension) will eventually become one integrated topic of study. Some suggestions on how this can be achieved are put forward in Section 5.2. We also restrict our focus to the cognitive processes and representations involved in bilingual language production and control that precede overt articulation.

2. Theoretical models of bilingual language production and control

As the depth and scope of bilingual language production research have increased over the years, so have the (neuro)psychological models accounting for bilingual language use (e.g., Abutalebi and Green, 2016; Baus et al., 2015; Branzi et al., 2014; Costa et al., 1999, 2006; Dylman and Barry, 2018; Finkbeiner et al., 2006; Green, 1998; Kroll et al., 2010; La Heij, 2005; Philipp et al., 2007; Poulisse and Bongaerts, 1994; Roelofs, 1997, 2014; Runnqvist et al., 2012, 2019). In this section, we will discuss some of the most prominent models of bilingual language production and control. In doing so, it will become clear how, over the years, various types of context have continuously played a central role in these models. This theoretical basis will allow us, later in this article, to look at the manipulation of context aspects in experimental studies of bilingual language production.

2.1. Models of bilingual language production

A critical difference between monolinguals and bilinguals is that bilinguals may describe the same concept or non-linguistic thought using words from more than one language. It seems unlikely that bilinguals can completely switch off one of their languages, so words (and grammatical structures) from two active languages may compete for selection, and inhibition of words (and grammatical structures) from a context-irrelevant language may be required for a bilingual to correctly express their thoughts in the context-appropriate language (Green, 1998; Hatzidaki et al., 2011). Indeed, theoretical proposals commonly emphasize the co-activation, integration, and interaction of both languages in the bilingual mind, assuming that a cognitive control system, engaging in activities such as monitoring and inhibition, plays a crucial role in managing relative language activation and language competition (e.g., Kroll and Gollan, 2014).

Several influential models of bilingual language production rely strongly on the standard model of (primarily monolingual) language production (Levelt, 1989, 1992; Levelt et al., 1999), a descriptive account that has been computationally implemented in the WEAVER++ model (Roelofs, 1997, 2014). According to this model, the process of language production consists of four main processing stages: 1) conceptualization and language choice, 2) formulation, entailing grammatical encoding and lemma retrieval, 3) morphophonological encoding, and 4) articulation. In the first processing stage, a message is prepared for speaking. During the second stage, lemmas (i.e., word-like units without phonological specification) are selected. Together they cover the conceptual representation to be expressed, while at the same time allowing for the construction of a complete and correct syntactic structure. Full specification of the morphological, phonological, and phonetic realization of the words in that structure takes place next, after which the sentence is actually incrementally uttered. Although models sometimes use different names for terms and components, competing models, such as the interactive activation model of language production (Dell, 1986; Dell et al., 1997) roughly presume the same stages of processing.

In line with earlier proposals (e.g., de Bot and Schreuder, 1993), the WEAVER++ model has been extended to the bilingual domain (e.g., Roelofs and Verhoef, 2006). The bilingual model assumes that language-independent concept representations coactivate lemmas and word form units specific for different languages. As a consequence, a flexible language control system should then allow bilingual speakers to activate and select (words of) their languages depending on their language goals and the situational context. As such, bilingual language production is proposed to involve a dynamic interplay of the two languages with both competition and cooperation processes that are subject to the bilingual's relative proficiency level in each language, language switching mechanisms, and language selection strategies. Roelofs and colleagues consider especially the role of attention in applying

condition-action rules (e.g., Roelofs, 1998, 2018a). In their view, inhibition can modulate the efficiency of language switching, although it is not strictly necessary (Roelofs et al., 2011; Verhoef et al., 2009).

These views notably contrast with that of Costa (e.g., Costa, 2018; Costa et al., 1999), who proposes that bilingual speakers' two languages are represented rather separately, their words being selected especially using context-sensitive language cues. According to Costa, the bilinguals' language systems and language selection processes function without significant interference of cross-activation. Each language is activated according to its own rules and structures, regardless of other languages. On the basis of contextual and other language cues, one language is activated selectively, while the other is inhibited or ignored. For example, in a conversation with a monolingual, the language spoken by the monolingual might be more readily activated and used by the bilingual speaker than other languages. In addition, the relative activation and language dominance would depend on the speaker's proficiency in each language. When bilinguals, depending on the situation, selectively activate and use one language rather than another, they apply their knowledge of the language to which word forms belong (e.g., 'book' is an English word). This sense of 'language membership' is shaped by various factors, such as language proficiency, language use patterns, social and cultural context, and individual experience. Language membership also plays a role in code-switching during conversations. According to Costa, strategic code-switching is done on the basis of language membership, in order to use each language for specific functions or expressing particular social or cultural identities. In all, Costa therefore pays less attention to inhibitory mechanisms than the alternative models described above, although these mechanisms are not excluded (cf. Costa, 2018; Costa et al., 2006). Despite their intrinsic differences, it should be noted that all these models have in common that they highlight the importance of *context* for language selection and production.

In the various models, the input message for language production is assumed to specify both the concept for which a lemma must be found and the language in which it is intended to be uttered. Because lemmas are linked to representations of language membership, the correct word form can be retrieved in the correct language. However, this still begs the question of how precisely a particular language choice is made for the intended utterance. Clearly, both conceptualization and language choice are dependent on non-linguistic and linguistic external cues. Returning to the opening example, the geographical location and background language spoken at the market place might serve as cues for speakers to opt for using their non-dominant (e.g., English) rather than their dominant language (e.g., Spanish) when interacting with any stand owner they encounter. Until now, such context-sensitive aspects of language choice have been insufficiently specified, most likely because they have been difficult to manipulate in experimental research paradigms.

2.2. Models of bilingual language control

Acknowledging the importance of language monitoring and control in bilingual language production, theoretical models have been developed that specifically aim to describe how bilinguals manage to select the context-appropriate language for speaking. Here we will focus on the influential Inhibitory Control Model (Green, 1998) and the Adaptive Control Hypothesis (Green and Abutalebi, 2013). The development of these models paralleled changes in experimental paradigms and acted as a catalyst for developing novel ones (for a review, see Sánchez et al., 2023). Critically, the Inhibitory Control Model and Adaptive Control Hypothesis, respectively, make a case for the importance of reactive and proactive inhibitory control processes as supporting efficient and context-appropriate bilingual language production.

So how do bilinguals manage to select the correct language for speech production? The Inhibitory Control Model assumes that complex tasks such as selecting a language for speaking are performed by

following a series of processing steps required by the task at hand, a so-called 'task schema'. For instance, in a picture naming experiment, the participant's task might be to 'name the picture in Language A' or to 'name the picture in Language B'. The task schema would include several processing steps for selecting the picture's name in Language A or B, depending on the task instruction and/or a given language cue. When Language A is the target language, naming in Language B should be inhibited, and vice versa. If, later on, a cue indicates that a word in the other language should be produced ('switching'), this may momentarily reactivate the previously inhibited task schema for picture naming in Language B, and inhibit the currently active task schema for picture naming in Language A. Overcoming the inhibition of a previously suppressed task schema (and the language it corresponds to) takes time and effort, and the Inhibitory Control Model therefore predicts that switching languages comes at a cost. Importantly, the inhibitory process at work is proposed to be *reactive* here: It plays a role only after a certain external stimulus triggers the language control mechanism.

Going beyond the Inhibitory Control Model and its theoretical focus on reactive inhibitory control, the Adaptive Control Hypothesis assumes that what exact control processes are required and implemented differs as a function of aspects of the broader context in which the speaking event takes place (Abutalebi and Green, 2016). In *single-language contexts*, such as when speaking with a monolingual friend, the choice for what (single) expected language to be spoken is straightforward and can be anticipated. In *dual-language contexts*, a bilingual may be expected to switch between languages, for instance as a function of whom they interact with, as in our opening example. In *dense code-switching contexts*, the bilingual may even voluntarily switch languages at will, for instance when interacting with another bilingual that masters the same language pair. Critically, what cognitive control processes (e.g., goal maintenance, suppression of interference) are active in the background, and the extent to which they are required and involved, differs as a function of the context. While a single-language context demands effective suppression of a non-target language for communication to be successful, a dense code-switching context actually makes efficient use of the parallel activation of both languages, and suppressing a language may not be necessary (Abutalebi and Green, 2016).

The Adaptive Control Hypothesis therefore proposes that inhibitory control processes during bilingual language production may not only be reactive, but also *proactive* in nature. The idea is that, as a function of the context at hand, a given language may in certain environments be inhibited in a proactive and sustained fashion to facilitate language production in another language. For instance, unbalanced bilinguals may proactively slightly inhibit their stronger L1 in a dual-language context to reduce competition and facilitate language production in their non-dominant and weaker L2 (Peeters and Dijkstra, 2018). Returning to our opening example, if you are visiting a market place in London, the likelihood of having to use your L2 English is quite high. In general terms, there is an enhanced probability of needing one language compared to another as a function of context. As such, proactively activating or inhibiting one language over the other may facilitate the efficiency of some of the upcoming interactions you may have.

In sum, models of language production and control have time and again ascribed to the importance of *context* when attempting to describe and explain the workings of the bilingual language production system. In the next section, we will see that the main experimental paradigms in this domain, however, have largely ignored the important role context has to play.

3. Common experimental practice in bilingual language production research

In this section, we will discuss two long-established and widely used experimental paradigms in the field of bilingual language production research: the *picture-word interference paradigm* and the *cued language-switching task*. This overview will provide us with a solid basis for

zooming in on adaptations of these traditional paradigms in the next section, as a first step towards empirically addressing linguistic, social, and non-linguistic context effects on bilingual language production.

3.1. The picture-word interference paradigm

What are the cognitive processes involved in (bilingual) language production? To study this issue in more detail, researchers have adapted the original picture naming paradigm (Cattell, 1886; Glaser, 1992) into a picture-word interference paradigm, initially applying it in the monolingual and later also in the bilingual domain (e.g., Damian and Bowers, 2003; Ehri and Ryan, 1980; Glaser and Dünghoff, 1984; Lupker, 1979; Miozzo and Caramazza, 2003; Rosinski et al., 1975; Schriefers et al., 1990). In a picture-word interference paradigm, participants are typically instructed to name pictures presented on a computer screen as quickly and accurately as possible while ignoring superimposed distractor words. As an example from the monolingual domain, a picture of a cat may be named while it is accompanied by a visual distractor word that is unrelated (e.g., *bin*), semantically related (e.g., *dog*), phonologically related (e.g., *hat* or *cap*), or even both semantically and phonologically related (e.g., *rat*) to the target picture (see Fig. 1). In an auditory variant of the paradigm, participants must ignore auditory rather than visual distractors while they are naming the target picture or planning to do so (e.g., Damian and Bowers, 2009; Schriefers et al., 1990). Any (typically semantic) interference or (typically phonological) facilitation relative to the unrelated condition provides insights into the cognitive representations and processing stages involved in word retrieval and production (for a review, see Hall, 2011). For instance, by varying the onset of the distractor word versus the onset of the picture on the screen, clever manipulations have made clear that semantic activation typically precedes phonological activation during

the production of individual words (e.g., Schriefers et al., 1990).

The monolingual picture-word interference paradigm has been adapted to address fundamental issues in the domain of bilingualism (e.g., Costa et al., 1999, 2005; Costa and Caramazza, 1999; Ehri and Ryan, 1980; Emmorey et al., 2021; Giezen and Emmorey, 2016; Gollan and Acenas, 2004; Guo et al., 2011; Hermans, 2004; Mahon et al., 2007). The bilingual version typically presents one image at a time on a computer screen to be named by the bilingual participant in a target language (e.g., English), while a written distractor word from the same language (e.g., English) or from another language (e.g., Dutch) is simultaneously presented on the image. Taking the paradigm into the bilingual realm thus allows for additional possible relations between picture and word, and questions of *cross-linguistic* cognitive representation and processing can be addressed. For example, when a Dutch-English bilingual is asked to name target images in their L2 English, for instance, a picture of a shark can be paired with a direct translation of the word corresponding to that image from the L1 Dutch (i.e., *haai*, the Dutch word for 'shark'), a semantically related word from the L1 (e.g., *dolfijn*, the Dutch word for 'dolphin') or the L2 (e.g., *dolphin*), or a phonologically related word from the L1 (e.g., *hark*, the Dutch word for 'rake') or the L2 (e.g., the English word *bark*). These manipulations allow for testing at what stages of language production the bilingual's different languages interact, and whether or not lexico-semantic and phonological representations from the two languages are stored together or separately in bilingual long term memory.

Experimental studies using the picture-word interference paradigm have shown at least three distinct effects in bilinguals. First, a *translation equivalent facilitation effect* has been observed (Costa et al., 1999; Costa and Caramazza, 1999; Hall, 2011; Hermans, 2004). For instance, if an image of a horse is to be named by an unbalanced Dutch-English bilingual in their L2 English, RTs will typically be shorter when the direct

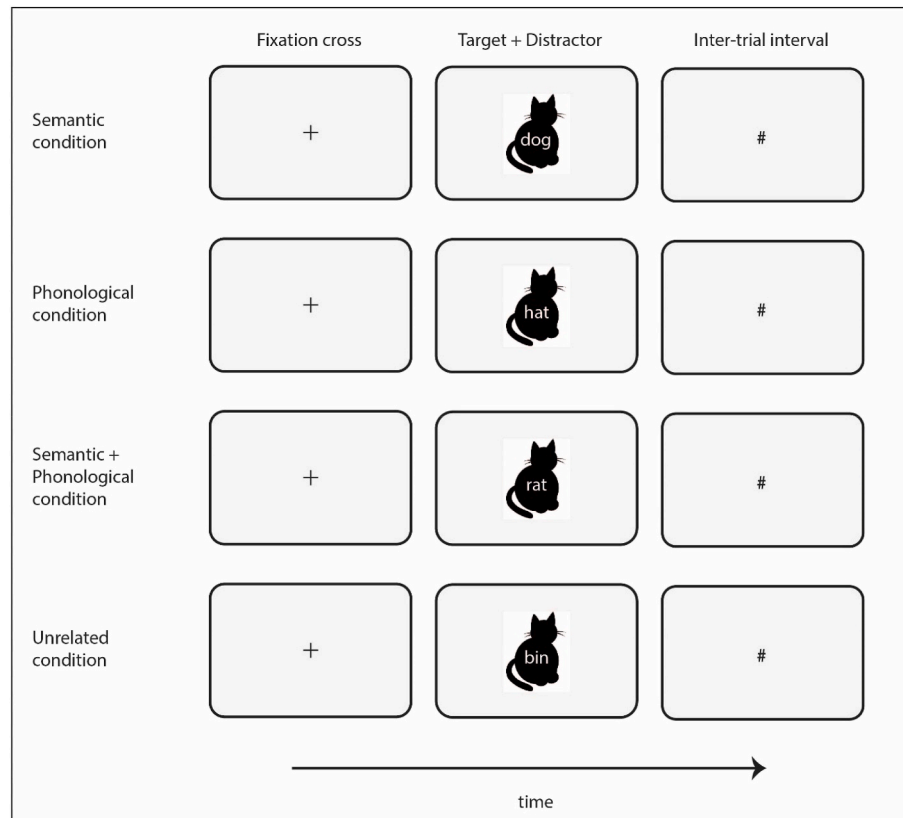


Fig. 1. Different conditions in the monolingual picture-word interference paradigm as a function of the relation between target image and super-imposed word. Participants are commonly instructed to name each image in a single word as quickly and accurately as possible. The paradigm can easily be extended to the bilingual domain (see main text for examples).

translation of the target word (the Dutch word *paard*, meaning ‘horse’) is shown as a distractor compared to when an unrelated distractor word (e.g., the Dutch word *stoel*, meaning ‘chair’) is shown. Second, response times are commonly longer for semantically related word-image pairs than for semantically unrelated word-image pairs. This *semantic interference effect* (Caramazza, 1997; Dell, 1986; Levelt et al., 1999; Roelofs, 1992) is typically prevalent regardless of the naming language (Belke et al., 2005; Costa et al., 2009; Howard et al., 2006; Roelofs, 2018a). The *phonological facilitation* effect has also been replicated in bilinguals, meaning that cross-linguistic phonological overlap between picture name and distractor word speeds up naming compared to an unrelated cross-linguistic control condition, though the effect is typically smaller than for within-language phonological overlap conditions (see Hall, 2011, for review).

In sum, experimental research using the picture-word interference paradigm has considerably increased our knowledge of how bilinguals store and select words from their two languages. These words are not stored in separate databases, but are part of a shared lexicon and may compete for selection (Costa et al., 2003; Ehri and Ryan, 1980; Gauvin et al., 2018; Hermans et al., 1998; Kroll and Stewart, 1994; Mahon et al., 2007). Although this insight is very valuable from a theoretical perspective, the experimental paradigm is limited in including aspects of the real-life communicative situations in which bilinguals typically find themselves. Obviously, bilinguals hardly ever encounter situations where images and words of different languages are superimposed. Moreover, they typically do not speak in single-word utterances and in the absence of an actual addressee. As such, the PWI paradigm commonly does not capture all the richness of bilingual language production in everyday situations. This leads to the possibility that existing models of bilingual language production – which are strongly based on experimental studies using this paradigm – may not fully generalize to bilingual language production in (everyday) context and need to be adapted or extended to do so.

3.2. The cued language-switching paradigm

Aspects of bilingual language production have also been studied widely by means of the *cued language-switching paradigm* (see Fig. 2). In this task, bilinguals are typically presented with individual digits or pictures on a computer screen and asked to name each stimulus in one language or another as a function of a cue. This cue can have an arbitrary

relation to the language it cues, for instance when colors are used, or be in some ‘motivated’ way related to it, such as when flags, culturally iconic images, or faces cue what language the bilingual should use on a given trial (e.g., Costa and Santesteban, 2004; Declerck and Philipp, 2015; Macnamara et al., 1968; Meuter and Allport, 1999). Responses are typically single-word utterances recorded by a microphone that are analyzed for response speed and accuracy between conditions. In a standard 2x2-design, performance on non-switch trials (eliciting the same language as on the previous trial) can be compared to performance on switch trials (eliciting a different language as on the previous trial). In addition, performance on L1 trials can be compared to performance on L2 trials. Finally, the design allows for testing whether potential costs of switching languages are larger in one direction (e.g., switching from L2 to L1) than in the other (e.g., from L1 to L2), for bilinguals with different degrees of overlap between the languages they master and different degrees of proficiency in the languages at hand. As a first proxy of the influence of context on language use, some studies have added single language blocks (in which all trials require the use of the same language) to the experiment, in order to test to what extent language context (i.e., the use of two vs one languages required during the block) influences performance.

The use of the cued language-switching paradigm has resulted in at least three theoretically important findings. First, bilinguals typically take longer to switch between languages than to stay in the same language (e.g., Bobb and Wodniecka, 2013; Costa et al., 2006; Kleinman and Gollan, 2016; Meuter and Allport, 1999; Peeters et al., 2014). In other words, *switch costs* have been observed. Second, and perhaps counter-intuitively, bilinguals sometimes respond more quickly in their non-dominant language compared to their dominant language, giving them a temporarily *reversed language dominance* (e.g., Christoffels et al., 2007; Costa and Santesteban, 2004; Kleinman and Gollan, 2016; Liu et al., 2019; Peeters and Dijkstra, 2018; Verhoef et al., 2009, 2010; for reviews, see Baus et al., 2015; Declerck and Philipp, 2015; Goldrick and Gollan, 2023). Third, unbalanced bilinguals typically name digits or pictures faster in a single-language context than on non-switch trials in a mixed language context: There are *language mixing costs* (e.g., Peeters and Dijkstra, 2018; Prior and Gollan, 2013; Segal et al., 2021; Timmer et al., 2019a,b).

These three types of findings have at least two important theoretical implications. First, in line with the Inhibitory Control Model, switch costs have been taken as evidence for the presence of reactive inhibitory

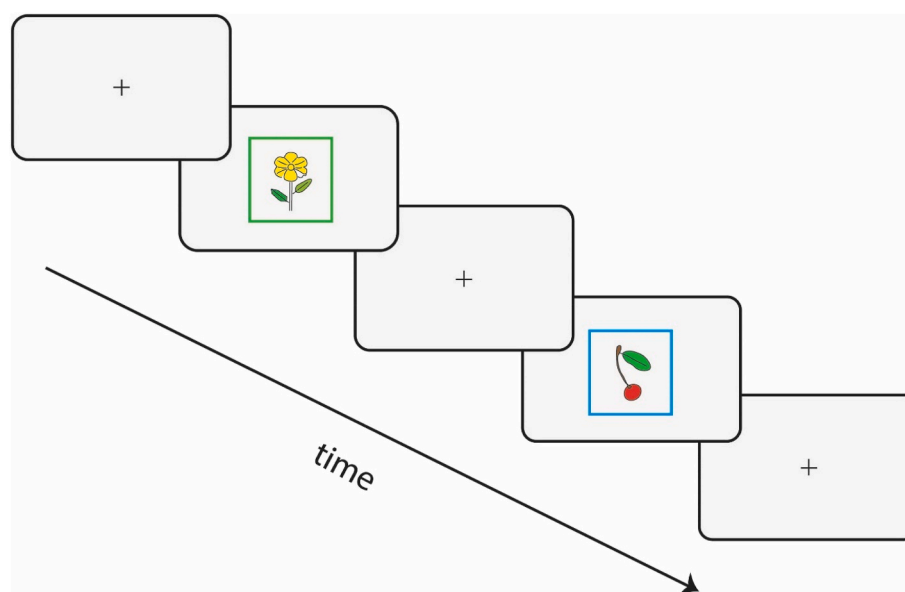


Fig. 2. Example order of events during a typical cued-language switching experiment. Participants name pictures or digits in their L1 or their L2 as a function of a color cue, here a colored rectangle around the pictures (cf. Meuter and Allport, 1999).

control processes in the bilingual mind. Indeed, when they are cued to switch languages, bilinguals are assumed to inhibit the presently active task schema and (re)activate the competing task schema that corresponds to the language at hand. The observation of *asymmetrical* switch costs in unbalanced bilingual participants is strongly in line with this account, because overcoming the temporary (substantial) inhibition of a stronger language should take longer than overcoming the temporary (limited) inhibition of a language that is not that strong anyway (Green, 1998; Meuter and Allport, 1999). Second, the presence of reversed language dominance and mixing costs have been taken to indicate that bilinguals are capable of proactively and adaptively activating or inhibiting one language over another, in a sustained rather than a trial-by-trial manner, in line with the Adaptive Control Hypothesis. As such, findings obtained with the cued language-switching paradigm are nicely in line with prominent theoretical models of bilingual language production and control.

Nevertheless, like the picture-word interference paradigm, the cued language-switching task raises the issue of ecological validity, because bilinguals in everyday life do not use (arbitrary) cues for selecting what language to express their thoughts in via single-word utterances. Rather, they often base this decision on contextual cues (e.g., the overall setting, or the language background of their interlocutor) or internal motivations (e.g., Molnar et al., 2015; Peeters, 2020; Woumans et al., 2015) and produce full sentences. It is therefore a valid question under what circumstances switch costs, reversed language dominance, and mixing costs may occur in the everyday life of a bilingual outside the lab (cf. Myers-Scotton, 2006).

4. Adding context to experimental bilingual language production research

To what extent do results from strict experimental laboratory paradigms generalize to more natural circumstances? It is increasingly suggested that the over-reliance on experimental control has made researchers ignore crucial aspects of natural interactive discourse in their experimental paradigms (Blanco-Elorrieta and Pyllkänen, 2018; Hamilton and Huth, 2020; Hasson et al., 2018; Peeters, 2019; Willems and Peelen, 2021). As illustrated by the previous section, for a long time, psycholinguistic studies have only minimally incorporated context aspects into their experimental paradigms, whether they are linguistic (e.g., discourse or full-sentence stimuli), social (e.g., looking at language production in the presence of an addressee), or non-linguistic (e.g., varying the physical context in which a cognitive process takes place) in nature. Nevertheless, over the last decades, some researchers in psycholinguistics have started to enhance the ecological validity of experimental paradigms for bilingual language production by means of clever adaptations of the existing paradigms discussed above. We will discuss examples of such studies in this section as a basis for our main proposal in the next.

4.1. Extensions of the picture-word interference paradigm

The original picture-word interference paradigm is contextually limited as it reduces the rich and dynamic process of everyday language production to a situation where individual participants produce single-word utterances in the absence of an actual addressee. Not surprisingly, once the paradigm became well-established, researchers have therefore begun to extend it in various ways. In the following, we will consider extensions in linguistic, social, and meaningful non-linguistic physical context aspects.

The *linguistic context* in the picture-word interference paradigm was initially enriched in the form of the so-called *multiword interference paradigm* (e.g., Meyer, 1996; Schriefers et al., 1998). In the monolingual domain, Meyer (1996) presented participants with images of two objects next to each other while they heard an auditory distractor word that could be semantically or phonologically related to either object name or

unrelated to both. Participants were instructed to produce phrases (e.g., ‘the snail and the hill’) or sentences (e.g., ‘the snail is next to the hill’) describing the object pair. Whenever the distractor word was semantically related to one of the elicited nouns, it delayed the onset of participants’ speech compared to the unrelated condition. When the word was phonologically related to the first noun in the sentence, phonological facilitation was observed. No such effect was observed for a phonological relation between distractor word and the second noun in the sentence. These findings are taken to indicate that the lemma representations of both nouns and the phonological form of the first (but not the second) noun in the phrase were already selected prior to the start of the produced utterance, demonstrating a limit to how much phonological form information is planned in advance at an early stage prior to the onset of articulation of a multi-word utterance (Meyer, 1996). Other ‘extended picture-word interference paradigms’ have used displays of two or four images with visually superimposed words to investigate the planning and production of different types of utterances including both nouns and verbs in a monolingual sentence context (e.g., Momma and Ferreira, 2019; Momma et al., 2015; also see Mädebach et al., 2020).

In the bilingual domain, the extended picture-word interference paradigm was recently adopted and adapted in a first study to examine the planning and production of multi-word utterances in multilingual participants (Ahn et al., 2021). Specifically, the extended paradigm has allowed for an investigation of how the syntax of multiple languages is stored, activated, and used, for instance in the case of languages with varying word orders. The study focused on Korean-English bilinguals, exploiting the substantially different word order characteristics in these two languages in the description of spatial relations (Ahn et al., 2021). Specifically, while English speakers would describe a scene of a lemon below a lobster (‘object-first’) using the word order [lemon][below] [lobster], speakers of Korean would have to use the order [lobster] [below][lemon] to indicate (‘location-first’) that the lemon is below the lobster. As such, compared to English monolinguals, concurrent activation of structural word order properties of both languages in the mind of a Korean-English bilingual might lead to interference effects caused by the mismatch in word order preferences across languages. Interestingly, by analyzing the effect of a superimposed visual distractor word that was semantically related to one of the two nouns (e.g., ‘apple’ related to ‘lemon’; ‘crab’ related to ‘lobster’), the authors were able to conclude that cross-linguistic syntactic interference was actually minimal. Specifically, relatively proficient Korean-English bilinguals performed similarly to English monolinguals when producing sentences in English and to Korean speakers with minimal knowledge of English when producing sentences in Korean. These findings were even observed in a context in which regular switches between the two languages were induced. Apparently, this group of Korean-English speakers activated only the syntactic structure of the language they were expected to use on a given trial, suggesting separately stored noun phrase (word order) representations or restricted dual-language activation (Ahn et al., 2021). In sum, the degree of cross-linguistic interference observed using the extended picture-word interference paradigm differed substantially from the observations taken from the traditional picture-word interference paradigm as summarized above.

In terms of extending the *social* and the *non-linguistic context* present in the picture-word interference paradigm, most progress has been made in the monolingual domain. Specifically, the traditional paradigm was adapted for use in a dialog setup in which two participants jointly take part in the experiment and produce both the target word and the distractor themselves (Kuhlen and Abdel Rahman, 2022). In this setup, mimicking a card game, one participant would be instructed to name a target sentence containing the distractor word as its final item (e.g., ‘Which word comes on apple?’), after which the other participant named a picture that could be semantically related (e.g., ‘pear’) or not (e.g., ‘chair’) to that distractor word. Clearly, this experimental setup enhances the social context in which language is produced, as participants produce language in the presence of a listener while having a

communicative motive. The non-linguistic physical context is enhanced in that participants play a game in a room that is larger and richer in nature compared to the traditional individual sound-proof booth. Interestingly, in this contextually enriched dialog setup, the classic semantic interference effect is not observed (Kuhlen and Abdel Rahman, 2022). The authors argue that, in their more social setup, participants may have processed their partner's speech primarily at a conceptual level. The conceptual facilitation between concepts such as APPLE and PEAR may have cancelled out any interference as observed in the traditional, non-social experimental setup. While these methodological advances yet await adaptations to the bilingual domain, they do show once more that well-established result patterns may have limited generalizability once context is taken into account.

In sum, language production studies in the bilingual domain have attempted to extend the well-established picture-word interference paradigm by enriching the linguistic context and having bilingual participants produce multiword utterances rather than single words. Studies from the monolingual domain show that aspects of a broader social and non-linguistic context can be included in the paradigm as well. Together, this handful of studies confirms that including a richer context in the lab may lead to theoretically and methodologically valuable results.

4.2. Extensions of the language-switching paradigm

As we have seen above, the traditional cued language-switching paradigm has yielded several theoretically interesting insights. Over the past decade, similar to adaptations of the picture-word interference task described above, the language-switching paradigm has been adapted and extended to better include context in the experimental equation. Specifically, first steps have been taken to include aspects of the linguistic, social, and non-linguistic context in which bilinguals typically switch languages in the experimental manipulations (e.g., Blanco-Elorrieta and Pykkänen, 2015; Hartsuiker, 2015; Martin et al., 2016; Molnar et al., 2015; Peeters, 2020; Peeters and Dijkstra, 2018; Smith et al., 2020; Timmer et al., 2017, 2019; Woumans et al., 2015; Zhang et al., 2015).

In terms of enriching the *linguistic context* in the language-switching paradigm, in various studies stimuli have now been used that allow participants to produce phrases or sentences rather than single words (e.g., Declerck et al., 2021; Declerck and Philipp, 2015; Gollan and Goldrick, 2016, 2017; Gullifer et al., 2013; Johns and Steuck, 2021; Li et al., 2022; Sánchez et al., 2022). For instance, one early study in this domain had unbalanced Polish-English bilinguals describe scenes using a simple progressive or perfective phrase in one of their two languages (Tarlowski et al., 2013). Based on an auditory language cue, participants described pictures of completed scenes (e.g., someone sitting with an empty glass) or pictures of scenes still in action (e.g., someone with a half-raised arm holding a full glass) in their L1 Polish or their L2 English. Comparable switch costs in both directions (from L1 to L2 and vice versa) and a reversed language dominance (faster RTs in L2 English compared to L1 Polish) were observed when participants described images of ongoing actions using a progressive phrase (e.g., 'He is drinking'). For completed actions and use of the perfective (e.g., 'He has drunk'), however, switching from the non-dominant English to the more dominant Polish was more costly than the other way around. The authors explain this finding by referring to the variable relative difficulty of acquiring different L2 English syntactic structures by L1 Polish speakers, where there may be more (as in the case of the progressive) or less (as in the case of the perfective) cross-linguistic overlap between L1 and L2 in terms of how aspect is expressed, with downstream consequences on language switching difficulties between phrases. These results hence indicate that context may modulate traditional findings from studies eliciting one-word utterances.

In terms of enriching the *social* and the *non-linguistic context* in the language-switching paradigm, an increasing number of studies have

turned to using more naturalistic and 'motivated' language cues. The use of well-known iconic cultural artifacts (e.g., the Statue of Liberty or the Great Wall of China) as cues has been observed to facilitate speaking in the language associated with that cue (Zhang et al., 2013). Also faces have been shown to prime a language, such that bilinguals are faster in producing words to a listener in a language that matches the presumed language identity of that listener (Woumans et al., 2015; cf. Blanco-Elorrieta and Pykkänen, 2015, 2017; Molnar et al., 2015). Indeed, for people we have met before, we typically know in which language we can communicate with them, and an individual interlocutor may serve as its own natural language cue, even in the case of famous individuals we do not know personally (Hartsuiker, 2015). Importantly, the use of more naturalistic language cues has been shown to modulate switch costs. For instance, symmetrical switch costs turned into asymmetrical switch costs for Chinese-English bilinguals when languages were cued by language-congruent facial cues compared to artificial color cues (Liu et al., 2019). At the same time, a reversed language dominance was observed irrespective of whether the language cue was artificial or motivated (Liu et al., 2019).

Not just the type of cue used, but also the overall physical context in which a language switching experiment is taking place may influence the results one observes. For example, in a study by Peeters and Dijkstra (2018), unbalanced Dutch-English bilinguals met two life-size virtual agents with a different monolingual language background (Dutch vs English). The virtual agents introduced themselves in either Dutch or English, after which participants named pictures in a large 3D virtual environment as a function of which of the two agents looked at them during picture presentation. This setup, which increased the social and non-linguistic context in comparison to traditional cued language-switching studies, reliably yielded symmetrical switch costs and reversed language dominance as in traditional studies testing this bilingual population. However, when the experiment was taken into a visually rich marketplace, and bilingual participants acted as stand owners informing monolingual market visitors about the price of their fruits and vegetables, reversed asymmetrical rather than symmetrical switch costs were found (Titus and Peeters, under review).

Finally, several studies have attempted to examine language switching and mixing in the absence of any language cue, by allowing participants to voluntarily switch between languages throughout (a part of) an experiment (e.g., de Bruin et al., 2018; Gollan and Ferreira, 2009; Jevtović et al., 2020; Sánchez et al., 2022). The voluntary language-switching paradigm typically follows the same procedure as the original cued language-switching task, but adds another block of mixed-language trials to the experiment. In this block of 'free-choice switching', participants may respond with whatever language comes to mind first rather than using an external cue. As in certain everyday situations, bilinguals may naturally choose which language to respond with, which makes a switch between languages potentially less effortful and may even provide a benefit if they pick the word that comes to mind first regardless of the language it belongs to (de Bruin et al., 2018; Gollan and Ferreira, 2009; Jevtović et al., 2020). A range of findings has been observed, with some studies showing similar switch costs in voluntary and cued tasks (e.g., de Bruin et al., 2018), other studies showing smaller voluntary than cued switch costs (e.g., Jevtović et al., 2020), and yet some other studies showing no switching costs at all when switches are voluntary (e.g., Blanco-Elorrieta and Pykkänen, 2017). In addition, voluntary language choice has sometimes made mixing costs disappear or even turn into mixing benefits (e.g., de Bruin et al., 2018, 2020; Gollan and Ferreira, 2009; Jevtović et al., 2020). As such, language-switching studies using (artificial) exogenous cues may have overestimated the cost of language switching and mixing compared to naturally occurring, endogenously motivated switches in a bilingual's everyday life (Peeters et al., 2014). Nevertheless, studies on spontaneous language switching in naturalistic conversations do suggest that switching may come at a cost (Faroqi-Shah and Wereley, 2022; Fricke et al., 2016) and the presence of 'motivated' language cues such as flags

does influence what language a bilingual will naturally use (de Bruin and Martin, 2022; see also Vaughan-Evans, 2023).

In sum, a broad variety of studies have expanded the original language-switching paradigm by increasing its linguistic, social, and/or non-linguistic richness, as inspired by the multidimensional nature of everyday bilingual interactions. These adaptations have led to result patterns that are not always in line with findings obtained using the traditional cued language-switching paradigm. As such, it seems fruitful to further investigate the extent to which cognitive processes involved in bilingual language production differ as a function of context. Indeed, while important first steps have now been taken, there are still several discrepancies between the behavior elicited from bilingual participants in lab setups as compared to the rich and dynamic settings in which bilinguals produce language in everyday life.

5. Promising future avenues for bilingual language production research

As we have seen above, theoretical models typically ascribe an important role to context effects in bilingual language processing. At the same time, however, standard experimental paradigms in this domain commonly do not consider the aspects of the broader context in which bilinguals typically communicate in any detail. It is only in recent adaptations that these paradigms have started to do so. In the present section, we will provide three proposals intended to advance the field from traditional lab studies towards more contextually relevant experimental paradigms. The first proposal (discussed in Section 5.1) focuses on simultaneously enriching the linguistic, social, and non-linguistic context by using virtual reality technology as a new mode of stimulus display in bilingual language production research. Analogously, the second proposal (Section 5.2) considers how the experimental study of bilingual language production could be enriched by studying the phenomenon in the context of bilingual language comprehension in dialog setups. Finally, our third proposal (Section 5.3) argues in favor of considering bilingual language production first and foremost from a multimodal perspective. In this perspective, bilingual speech production is studied in the context of the other bodily signals - such as facial expressions and co-speech hand gestures - that bilinguals naturally transmit while speaking.

5.1. Use of virtual reality to study bilingual language production

Typically, bilingual speakers in everyday life produce multi-word utterances in rich and dynamic social contexts, like a market place, bar, or university. In contrast, bilingual participants in a standard language production experiment are usually positioned in front of a computer screen to name digits or pictures in one language or another as a function of task instructions and/or language cues. In this perceptually poor and static environment, they are commonly asked to speak only single words into a microphone. In a sense, it is difficult to come up with a setting that is further removed from real life than the lab, given that the core aspects of everyday communication have been taken out. So how can we reproduce the richness of everyday bilingual communication in the lab?

To tackle this issue, we have distinguished between the linguistic, social, and non-linguistic physical context in which bilingual language production typically takes place. In the last decade, psycholinguistic work has increasingly applied immersive virtual reality technology for creating such rich settings in the lab while maintaining the experimental control required to collect meaningful data (see Fig. 3). Studies have now used virtual reality to investigate practically all aspects of the psychology of language, from the acquisition and use of spatial language (Nölle et al., 2020) to the production of pointing gestures and their synchronization with speech production processes (Chu and Hagoort, 2014; Raghavan et al., 2023), to the role of prediction in language comprehension (Heyselaar et al., 2021; Huizeling et al., 2022), and even

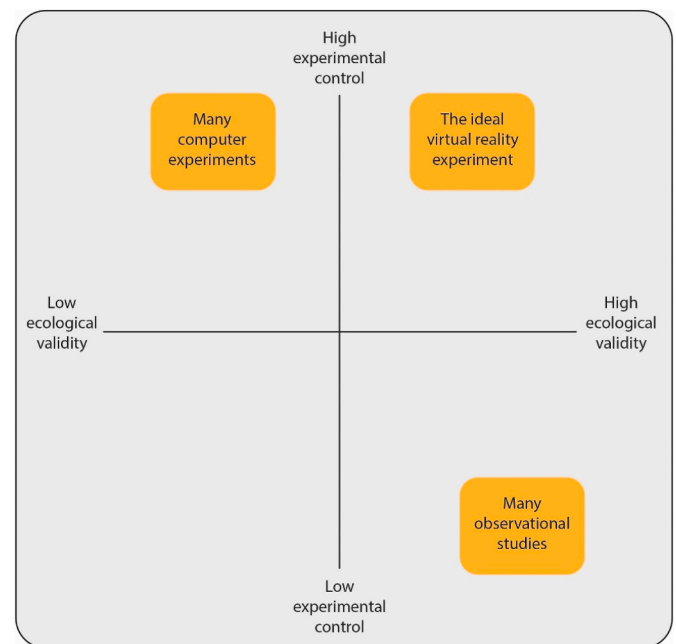


Fig. 3. Experimental control and ecological validity when considered as two orthogonal factors rather than two ends on a continuum (adapted from Peeters, 2019). While many non-experimental observational studies do not require and lack experimental control, and for many standard computer experiments it remains unclear to what extent they are ecologically valid, virtual reality holds the promise to combine the best of both worlds.

to the cognitive processes underlying reading (Mirault et al., 2020; Pianzola et al., 2019). As we saw above, also a first study on bilingual language production turned to using virtual reality technology as a means of more real-life stimulus display. This allowed bilingual participants to produce language for life-size and 3D virtual listeners in the lab (Peeters and Dijkstra, 2018).

Critically, the immersive stimulus display provided by virtual reality technology requires and allows the incorporation of linguistic, social, and physical context aspects. Who will the (bilingual) participant interact with in the virtual world? What should this virtual world actually look like? What linguistic input is the participant presented with and what linguistic output do we expect the participant to produce? What should the virtual agents' non-verbal behavior look like? Importantly, we know that participants typically interact with virtual agents as they would with actual humans (Heyselaar et al., 2017; Gijssels et al., 2016). They even take into account the virtual agents' non-verbal bodily signals during language comprehension (Hömke et al., 2018). At the same time, obtained data are as reliable as those collected using computer paradigms in the language sciences (Huizeling et al., 2022; Peeters and Dijkstra, 2018; Tromp et al., 2018). Participants commonly feel present in the virtual environment and suspend their disbelief (Cummins and Bailenson, 2016; Slater, 2009). As such, the method can reliably be applied to study bilingual language production in context as well.

Virtual reality cannot only be used to answer old questions in new ways, but also generates new theoretical questions. Peeters (2020) investigated to what extent switching *between listeners* during bilingual language production yields similar behavioral and neurophysiological switch costs as in switching *between languages*. Unbalanced Dutch-English bilinguals were immersed in a virtual environment where they met two Dutch and two English virtual agents. Depending on whom of these four addressees looked at them, they described pictures in either Dutch or English. Switching between listeners, even within the same language, was found to slow down response times compared to addressing the same listener twice in a row. When this *listener switch* was

at the same time a *language switch*, response times slowed down even more. Thus, switching languages comes at a cost over and above the cost that is induced by switching between (naturalistic) language cues. Interestingly, listener switches within and across languages yielded similar amplitude deviations in the bilinguals' event-related potentials prior to speaking. This finding suggests that any potential cognitive benefits of bilingualism are not caused by bilinguals' switching between languages, given that switching between listeners within the same language elicits similar neurophysiological activity compared to switching between listeners across languages (Peeters, 2020).

In this virtual reality application, bilinguals switch for specific listeners in a 3D environment. As a consequence, the social and non-linguistic physical context in the experiment is contextually enriched relative to traditional studies that present stimuli on a computer screen. Nevertheless, the linguistic context in this study was still limited, because participants responded only with single-word utterances. Ongoing work therefore puts bilingual participants in an immersive virtual market place where they act as store owners of a fruit and vegetable stand (Titus and Peeters, under review). In this environment, life-size virtual agents meet the participant, showing them a piece of fruit or a vegetable while wanting to know its cost (see Fig. 4). Participants have been made aware that these virtual customers function as Dutch or English monolinguals. They can therefore answer the customers with full sentences, such as *'the pineapple costs 90 cents'* or its Dutch equivalent. A change of customer may or may not correspond to a switch between languages. This allows the experiment not only to make a distinction between interlocutor and language switches, but also to establish whether any switch costs or reversed language dominance surface under lifelike conditions (Titus and Peeters, under review).

In sum, the rich environments that can be created in the lab by virtual reality technology allow one to realistically study the role of context in bilingual language production outside of the lab. For instance, it is possible now to immerse bilingual participants in different cultural contexts, such as a French café or an English pub, and test how the bilingual's presence in one or the other culture-specific environment affects their language switching behavior. Although designing a virtual

reality study is graphically more challenging than setting up a picture naming study, databases with validated 3D assets are now available (Hein et al., 2022; Peeters, 2018; Tromp et al., 2020). Furthermore, virtual reality headsets have become relatively affordable. In all, the virtual reality research method effectively allows an important enhancement of the various contexts in which bilingual language production occurs.

5.2. Focus on bilingual dialog situations in the lab

While the shift from single-word picture naming to the production of sentences enriched the linguistic context in which bilingual language production research took place, existing studies of bilingual sentence production have commonly focused on one side of linguistic interaction, namely language production. Obviously, language use commonly occurs in a broader context that is often bidirectional: We speak *and* listen to an interlocutor in our everyday communication, and align what we say to the person we speak with (Brennan and Clark, 1996; Garrod and Pickering, 2004). A conversation can be thought of as a joint action, in which the interlocutors constantly take each other's viewpoint into account and adapt what they say to the assumed communicative needs of their partner (Clark, 1996). Would one scientifically investigate dancing the tango by asking individuals to dance on their own in a research laboratory (cf. Sebanz et al., 2006)?

In Sections 4.1 and 4.2 we saw that both the standard cued language-switching paradigm and the standard picture-word interference paradigm have been extended to focus on bilinguals' production of multi-word utterances (i.e., phrases, sentences). Then, in Section 5.1, we saw that virtual reality can be used to have bilinguals switch between languages in a rich physical context in which they produce full sentences for an addressee. Although the linguistic and non-verbal behavior of virtual agents is more easily replicable across experimental sessions and different labs than the behavior of human confederates (Kuhlen and Brennan, 2013), having a naturalistic and full-fledged dialog with a virtual agent or avatar still remains a technical challenge (Lugrin et al., 2022). Nevertheless, it is possible to study bilinguals' production of



Fig. 4. A virtual reality CAVE setup, as seen from the control room situated behind it. In the CAVE, the participant wears 3-D shutter glasses to become immersed in a virtual environment that is projected around them on large digital screens. In this example, bilingual participants are stand owners at a 3D market place. They communicate with life-size virtual customers in one language or another as a function of the customer's language background (cf. Titus and Peeters, under review).

sentences in *dialog* settings in the lab.

The switching between languages in a bilingual dialog context has recently been studied experimentally by Kootstra et al. (2020). In two experiments, unbalanced Dutch-English bilinguals played a game with a confederate in which participant and confederate took turns describing line drawings of events in single, active sentences. Confederates could switch between Dutch and English within a sentence ('code-switching') or not, and pictures could elicit words that overlapped between Dutch and English (e.g., Dutch *roos*, corresponding to the English word *rose*) or not (e.g., Dutch *fiets*, corresponding to the English word *bike*). Participants were allowed to use Dutch, English, or a combination of both (i.e., code-switching) to describe target pictures. Interestingly, participants code-switched more often when the confederate just code-switched compared to when the confederate just produced a sentence in Dutch. They particularly did so when the picture for the participant could be described using a cognate (e.g., Dutch *roos*, English *rose*). Thus, participants aligned their choice of language to the behavior of their conversational partner, and lexical overlap between languages could be an additional trigger to do so.

By considering sentence production in a dialog setup, it becomes clear that the bilingual language production process is strongly dependent on social and linguistic context. Specifically, recently encountered linguistic input activating the language comprehension system affects the language choice a bilingual makes for production. Bilinguals commonly align their speech with their dialog partner. Such central properties of bilingual language production cannot be investigated in paradigms in which bilingual participants name pictures or digits in an isolated booth into a microphone. Indeed, theoretical models such as the Adaptive Control Hypothesis can be tested and further developed only by manipulating the (social, linguistic, physical) context in which bilingual language production takes place. Therefore, the psycholinguistic study of bilingual language production from a dialog perspective is a promising and timely way to go. In turn, the sociolinguistic analysis of naturalistic bilingual language production in speech corpora may serve as a well-informed foundation for the development of targeted psycholinguistic experiments to test and extend existing cognitive models of bilingual language use (e.g., Fricke and Kootstra, 2016; Fricke et al., 2016).

5.3. Focus on multimodal bilingual communication in the lab

Human communication is intrinsically multimodal in nature, in that we typically use multiple bodily channels in parallel to get our communicative messages across. In face-to-face situations, we often combine a spoken signal with a meaningful facial expression in tight synchronization with the co-speech hand gestures we concurrently produce (Holler and Levinson, 2019; Mondada, 2016). Extensions of the standard model of language production, such as the Interface Hypothesis, consider the production of such multimodal messages (Kita and Özyürek, 2003). Surprisingly, in the experimental study of bilingual language production, 'language' has often been fully equated with 'speech'. Common laboratory tasks such as picture-word interference or cued-language switching ignore any non-spoken bodily signals produced by the speaker. Research on bilingual language acquisition in children often does look into communicative non-verbal signals (e.g., Nicoladis et al., 1999) and research on bimodal bilingualism naturally focuses on the body as a communicative channel (e.g., Emmorey et al., 2008). In contrast, psycholinguistic studies on bilingual spoken language production tend to focus on the unimodal speech stream alone.

In the literature on hand gestures and speech-gesture integration, it is commonly accepted that speech and gesture form one integrated system (e.g., Kendon, 2004; McNeill, 1992; Özyürek, 2014). We therefore miss out on understanding a significant part of our communicative abilities if we reduce language to speech. Moreover, what and how many bodily signals bilinguals transmit may differ as a function of what language they concurrently speak in. For instance, bilinguals may produce more

gestures in their weaker compared to their stronger language (Gullberg, 1998; Marcos, 1979). Approaching bilingual language production from a multimodal perspective raises many new and original theoretical questions (also see Gullberg, 2012). How do bilinguals synchronize their hand gestures and facial expressions with concurrently produced speech in their two languages? To what extent do language-specific co-speech gestures require inhibition when communicating in the other language? Are co-speech gestural representations, like lexical items, stored with a particular language tag in long term memory, or are they linked to relatively language-independent conceptual representations instead? How should models of bilingual language production and control be extended to account for the bodily signals bilinguals produce when they speak? How does language proficiency modulate aspects of bilingual non-verbal communication?

Over the past decade, only a few studies have started looking at bilinguals' gestural behavior in the lab. Azar et al. (2020) compared the gestural behavior of Turkish-Dutch bilinguals to that of Dutch and Turkish monolinguals in an event description task. Participants watched silent videos on a computer screen and were subsequently asked to explain to a naïve addressee what they just saw in each video. Bilingual participants performed the task once in Turkish and once in Dutch, with a two-week interval in between. Overall, it was observed that monolingual Turkish speakers gestured more than Dutch monolinguals, and that bilingual co-speech gesture rate matched these cross-linguistic differences dependent on the task language. In other words, Turkish-Dutch bilinguals gestured as much as Turkish monolinguals when speaking Turkish, and as little as Dutch monolinguals when speaking Dutch. Therefore, when switching from one language to another, these bilinguals seem to adapt their overall gestural behavior as well. In a different study, Özçalışkan (2016) actually observed that the co-speech gestures of Turkish-English bilinguals resembled those of Turkish monolinguals, regardless of whether the bilinguals described motion events in their L1 Turkish or their L2 English. These studies deserve follow-up work, not only because their findings stand in contrast, but also and mainly because they enrich our understanding of the bilingual language production process.

In sum, focusing on language as a unimodal phenomenon, it is impossible to achieve a full understanding of the cognitive processes supporting bilingual language production. When speakers are embedded in a social context, they will naturally co-produce meaningful signals via bodily channels such as the hands and face. In fact, by approaching language as an integrated multimodal system of closely synchronized streams of (visual, auditory) information, the bodily signals themselves can be seen as enriching the linguistic context in which speech is produced (Dijkstra and Peeters, 2023).

6. Conclusion

While laboratory experiments allow us to determine the basic mechanisms underlying language processing under empirically well-controlled conditions, the ultimate goal of psycholinguistics is to understand real-life human language use. A comprehensive analysis of complex language activities such as bilingual language production requires therefore research paradigms that combine experimental vigor with ecological validity. We have described the possibilities but also the limitations of the paradigms that have dominated the field, as well as the attempts to extend them to richer and more naturalistic, contextually-sensitive circumstances of bilingual language use. More recently, paradigms have shifted their focus from processing mechanisms in individual language users to aspects of dialog, allowing the investigation of the interaction of speech production and comprehension, and the interaction of speech and non-verbal bodily signals. Although this has clarified some aspects of speaker-internal contextual effects, the effects of physical and social contextual influences remain relatively little understood.

With the arrival of advanced virtual reality technology, the experimenter's goal of performing ecologically valid research while keeping

experimental control comes within reach. This is especially important for the study of bilinguals, because their language use may be considered as one of the most complex human feats. In particular, virtual reality introduces the possibility of generating complex physical and social contexts at relatively low cost. Importantly, use of the method allows the simultaneous manipulation and combination of multiple contextual and language factors. This opens up a vision of future research in which multiple disciplines contribute to bilingual language production research by investigating the interactions between multiple modalities, including the consideration of gesture, facial cues, visual search, emotion, and complex decision making. Participants can, for instance, be immersed in realistic and dynamic non-linguistic contexts such as a Saturday marketplace, workplace, or café in the home country or abroad. In terms of linguistic contexts, participants could speak with monolingual and multilingual virtual agents, while incorporating a greater depth of interlocutor background knowledge and various discourse scenarios.

In sum, we have argued that recent advances allow the researcher to add a variety of linguistic and non-linguistic context aspects to the experimental study of bilingual language production. The resulting multimodal perspective allows the investigation of language use in rich and broad social and interactive natural discourse contexts. Because of its immersive character and the relatively easy manipulation of a variety of context factors, virtual reality allows the study of natural bilingual language production in dialogs including multiple full-sentence utterances. As such, it will bring the researchers' dream of understanding actual day-to-day language production, in monolinguals and bilinguals alike, closer to reality.

CRedit authorship contribution statement

Alex Titus: Conceptualization, Investigation, Writing - original draft. **Ton Dijkstra:** Conceptualization, Investigation, Writing - original draft, Writing - review & editing. **Roel M. Willems:** Conceptualization, Writing - review & editing. **David Peeters:** Conceptualization, Writing - original draft, Writing - review & editing, Visualization, Supervision.

Declaration of competing interest

None.

Data availability

No data was used for the research described in the article.

Acknowledgments

We are greatly indebted to Angela de Bruin and one anonymous reviewer for their helpful suggestions on an earlier draft of this article.

References

- Abutalebi, J., Green, D.W., 2016. Neuroimaging of language control in bilinguals: neural adaptation and reserve. *Biling. Lang. Cognit.* 19 (4), 689–698. <https://doi.org/10.1017/S1366728916000225>.
- Ahn, D., Ferreira, V.S., Gollan, T.H., 2021. Selective activation of language specific structural representations: evidence from extended picture-word interference. *J. Mem. Lang.* 120, 104249. <https://doi.org/10.1016/j.jml.2021.104249>.
- Azar, Z., Özyürek, A., Backus, A., 2020. Turkish-Dutch bilinguals maintain language-specific reference tracking strategies in elicited narratives. *Int. J. Biling.* 24 (2), 376–409. <https://doi.org/10.1177/1367006919838375>.
- Baus, C., Branzi, F., Costa, A., 2015. On the mechanism and scope of language control in bilingual speech production. In: Schwieter, J.W. (Ed.), *The Cambridge Handbook of Bilingual Processing*. Cambridge University Press, pp. 508–526. <https://doi.org/10.1017/CBO9781107447257.022>.
- Belke, E., Meyer, A.S., Damian, M.F., 2005. Refractory effects in picture naming as assessed in a semantic blocking paradigm. *Q. J. Exp. Psychol.* 58 (4), 667–692. <https://doi.org/10.1080/02724980443000142>.
- Blanco-Elorrieta, E., Pylkkänen, L., 2015. Brain bases of language selection: MEG evidence from Arabic-English bilingual language production. *Front. Hum. Neurosci.* 9. <https://doi.org/10.3389/fnhum.2015.00027>.
- Blanco-Elorrieta, E., Pylkkänen, L., 2017. Bilingual Language switching in the laboratory versus in the wild: the spatiotemporal dynamics of adaptive language control. *J. Neurosci.* 37 (37), 9022–9036. <https://doi.org/10.1523/JNEUROSCI.0553-17.2017>.
- Blanco-Elorrieta, E., Pylkkänen, L., 2018. Ecological validity in bilingualism research and the bilingual advantage. *Trends Cognit. Sci.* 22 (12), 1117–1126. <https://doi.org/10.1016/j.tics.2018.10.001>.
- Bobb, S.C., Wodniecka, Z., 2013. Language switching in picture naming: what asymmetric switch costs (do not) tell us about inhibition in bilingual speech planning. *J. Cognit. Psychol.* 25 (5), 568–585. <https://doi.org/10.1080/20445911.2013.792822>.
- Branzi, F.M., Martin, C.D., Abutalebi, J., Costa, A., 2014. The after-effects of bilingual language production. *Neuropsychologia* 52, 102–116. <https://doi.org/10.1016/j.neuropsychologia.2013.09.022>.
- Brennan, S.E., Clark, H.H., 1996. Conceptual pacts and lexical choice in conversation. *J. Exp. Psychol. Learn. Mem. Cognit.* 22, 1482–1493. <https://doi.org/10.1037/0278-7393.22.6.1482>.
- Caramazza, A., 1997. How many levels of processing are there in lexical access? *Cogn. Neuropsychol.* 14 (1), 177–208. <https://doi.org/10.1080/026432997381664>.
- Cattell, J.M., 1886. The time it takes to see and name objects. *Mind* 11 (41), 63–65.
- Christoffels, I.K., Firk, C., Schiller, N.O., 2007. Bilingual language control: an event-related brain potential study. *Brain Res.* 1147, 192–208. <https://doi.org/10.1016/j.brainres.2007.01.137>.
- Chu, M., Hagoort, P., 2014. Synchronization of speech and gesture: evidence for interaction in action. *J. Exp. Psychol. Gen.* 143, 1726–1741. <https://doi.org/10.1037/a0036281>.
- Clark, H.H., 1996. *Using Language*. Cambridge University Press.
- Costa, A., 2018. *The Bilingual Brain: and what it Tells Us about the Science of Language*. Allen Lane.
- Costa, A., Alario, F.-X., Caramazza, A., 2005. On the categorical nature of the semantic interference effect in the picture-word interference paradigm. *Psychonomic Bull. Rev.* 12 (1), 125–131. <https://doi.org/10.3758/BF03196357>.
- Costa, A., Caramazza, A., 1999. Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish-English and English-Spanish bilinguals. *Biling. Lang. Cognit.* 2 (3), 231–244. <https://doi.org/10.1017/S1366728999000334>.
- Costa, A., Mahon, B., Savova, V., Caramazza, A., 2003. Level of categorisation effect: a novel effect in the picture-word interference paradigm. *Lang. Cognit. Process.* 18 (2), 205–234. <https://doi.org/10.1080/01690960143000524>.
- Costa, A., Miozzo, M., Caramazza, A., 1999. Lexical selection in bilinguals: do words in the bilingual's two lexicons compete for selection? *J. Mem. Lang.* 41 (3), 365–397. <https://doi.org/10.1006/jmla.1999.2651>.
- Costa, A., Santesteban, M., 2004. Lexical access in bilingual speech production: evidence from language switching in highly proficient bilinguals and L2 learners. *J. Mem. Lang.* 50 (4), 491–511. <https://doi.org/10.1016/j.jml.2004.02.002>.
- Costa, A., Santesteban, M., Ivanova, I., 2006. How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *J. Exp. Psychol. Learn. Mem. Cognit.* 32 (5), 1057–1074. <https://doi.org/10.1037/0278-7393.32.5.1057>.
- Costa, A., Strijkers, K., Martin, C., Thierry, G., 2009. The time course of word retrieval revealed by event-related brain potentials during overt speech. *Proc. Natl. Acad. Sci. USA* 106 (50), 21442–21446. <https://doi.org/10.1073/pnas.0908921106>.
- Cummings, J.J., Bailenson, J.N., 2016. How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychol.* 19 (2), 272–309. <https://doi.org/10.1080/15213269.2015.1015740>.
- Damian, M.F., Bowers, J.S., 2003. Locus of semantic interference in picture-word interference tasks. *Psychonomic Bull. Rev.* 10 (1), 111–117. <https://doi.org/10.3758/BF03196474>.
- Damian, M.F., Bowers, J.S., 2009. Assessing the role of orthography in speech perception and production: evidence from picture-word interference tasks. *Eur. J. Cognit. Psychol.* 21 (4), 581–598. <https://doi.org/10.1080/09541440801896007>.
- de Bot, K., Schreuder, R., 1993. Word production and the bilingual lexicon. In: Schreuder, R., Weltens, B. (Eds.), *The Bilingual Lexicon*. John Benjamins, pp. 191–214. <https://doi.org/10.1075/sibil.6.10bot>.
- de Bruin, A., Martin, C.D., 2022. Perro or txakur? Bilingual language choice during production is influenced by personal preferences and external primes. *Cognition* 222, 104995. <https://doi.org/10.1016/j.cognition.2021.104995>.
- de Bruin, A., Samuel, A.G., Duñabeitia, J.A., 2018. Voluntary language switching: when and why do bilinguals switch between their languages? *J. Mem. Lang.* 103, 28–43. <https://doi.org/10.1016/j.jml.2018.07.005>.
- de Bruin, A., Samuel, A.G., Duñabeitia, J.A., 2020. Examining bilingual language switching across the lifespan in cued and voluntary switching contexts. *J. Exp. Psychol. Hum. Percept. Perform.* 46 (8), 759–788. <https://doi.org/10.1037/xhp0000746>.
- Declerck, M., Meade, G., Midgley, K.J., Holcomb, P.J., Roelofs, A., Emmorey, K., 2021. On the connection between language control and executive control—an ERP study. *Neurobiology of Language* 1–47. https://doi.org/10.1162/nol_a_00032.
- Declerck, M., Philipp, A.M., 2015. A sentence to remember: instructed language switching in sentence production. *Cognition* 137, 166–173. <https://doi.org/10.1016/j.cognition.2015.01.006>.
- Dell, G.S., Burger, L.K., Svec, W.R., 1997. Language production and serial order. *Psychol. Rev.* 104 (1), 123–147.

- Dell, G.S., 1986. A spreading-activation theory of retrieval in sentence production. *Psychol. Rev.* 93 (3), 283–321. <https://doi.org/10.1037/0033-295X.93.3.283>.
- Dijkstra, T., Peeters, D., 2023. *The New Psychology of Language. From Body to Mental Model and Back*. Routledge. <https://doi.org/10.4324/9781003326274>.
- Donders, F.C., 1969. On the speed of mental processes. *Acta Psychol.* 30, 412–431. [https://doi.org/10.1016/0001-6918\(69\)90065-1](https://doi.org/10.1016/0001-6918(69)90065-1).
- Dylman, A.S., Barry, C., 2018. When having two names facilitates lexical selection: similar results in the picture-word task from translation distractors in bilinguals and synonym distractors in monolinguals. *Cognition* 171, 151–171. <https://doi.org/10.1016/j.cognition.2017.09.014>.
- Ehri, L.C., Ryan, E.B., 1980. Performance of bilinguals in a picture-word interference task. *J. Psycholinguist. Res.* 9 (3), 285–302. <https://doi.org/10.1007/BF01067243>.
- Emmorey, K., Borinstein, H.B., Thompson, R., Gollan, T.H., 2008. Bimodal bilingualism. *Biling. Lang. Cognit.* 11 (1), 43–61. <https://doi.org/10.1017/S1366728907003203>.
- Emmorey, K., Mott, M., Meade, G., Holcomb, P.J., Midgley, K.J., 2021. Lexical selection in bimodal bilinguals: ERP evidence from picture-word interference. *Language, Cognition and Neuroscience* 36 (7), 840–853. <https://doi.org/10.1080/23273798.2020.1821905>.
- Farooqi-Shah, Y., Werede, S., 2022. Investigation of code-switching cost in conversation and self-paced reading tasks. *Int. J. BiLing.* 26 (3), 308–333. <https://doi.org/10.1177/13670069211056438>.
- Finkbeiner, M., Gollan, T.H., Caramazza, A., 2006. Lexical access in bilingual speakers: what's the (hard) problem? *Biling. Lang. Cognit.* 9 (2), 153–166. <https://doi.org/10.1017/S1366728906002501>.
- Fricke, M., Kootstra, G.J., 2016. Primed codeswitching in spontaneous bilingual dialogue. *J. Mem. Lang.* 91, 181–201. <https://doi.org/10.1016/j.jml.2016.04.003>.
- Fricke, M., Kroll, J.F., Dussias, P.E., 2016. Phonetic variation in bilingual speech: a lens for studying the production-comprehension link. *J. Mem. Lang.* 89, 110–137. <https://doi.org/10.1016/j.jml.2015.10.001>.
- Garrod, S., Pickering, M.J., 2004. Why is conversation so easy? *Trends Cognit. Sci.* 8 (1), 8–11. <https://doi.org/10.1016/j.tics.2003.10.016>.
- Gauvin, H.S., Jonen, M.K., Choi, J., McMahon, K., de Zubicaray, G.L., 2018. No lexical competition without priming: evidence from the picture-word interference paradigm. *Q. J. Exp. Psychol.* 71 (12), 2562–2570. <https://doi.org/10.1177/1747021817747266>.
- Giezen, M.R., Emmorey, K., 2016. Semantic integration and age of acquisition effects in code-blend comprehension. *J. Deaf Stud. Deaf Educ.* 21 (2), 213–221. <https://doi.org/10.1093/deafed/env056>.
- Gijssels, T., Casasanto, L.S., Jasmin, K., Hagoort, P., Casasanto, D., 2016. Speech accommodation without priming: the case of pitch. *Discourse Process* 53 (4), 233–251. <https://doi.org/10.1080/0163853X.2015.1023965>.
- Glaser, W.R., 1992. Picture naming. *Cognition* 42 (1), 61–105. [https://doi.org/10.1016/0010-0277\(92\)90040-O](https://doi.org/10.1016/0010-0277(92)90040-O).
- Glaser, W.R., Döngelhoff, F.-J., 1984. The time course of picture-word interference. *J. Exp. Psychol. Hum. Percept. Perform.* 10 (5), 640–654. <https://doi.org/10.1037/0096-1523.10.5.640>.
- Goldrick, M., Gollan, T.H., 2023. Inhibitory control of the dominant language: reversed language dominance is the tip of the iceberg. *J. Mem. Lang.* 130, 104410. <https://doi.org/10.1016/j.jml.2023.104410>.
- Gollan, T.H., Acenas, L.-A.R., 2004. What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and Tagalog-English bilinguals. *J. Exp. Psychol. Learn. Mem. Cognit.* 30 (1), 246–269. <https://doi.org/10.1037/0278-7393.30.1.246>.
- Gollan, T.H., Ferreira, V.S., 2009. Should I stay or should I switch? A cost-benefit analysis of voluntary language switching in young and aging bilinguals. *J. Exp. Psychol. Learn. Mem. Cognit.* 35 (3), 640–665. <https://doi.org/10.1037/a0014981>.
- Gollan, T.H., Goldrick, M., 2016. Grammatical constraints on language switching: language control is not just executive control. *J. Mem. Lang.* 90, 177–199. <https://doi.org/10.1016/j.jml.2016.04.002>.
- Gollan, T.H., Goldrick, M., 2017. A switch is not a switch: syntactically-driven bilingual language control. *J. Exp. Psychol. Learn. Mem. Cognit.* 44 (1), 143. <https://doi.org/10.1037/xlm0000462>.
- Green, D.W., 1998. Mental control of the bilingual lexico-semantic system. *Biling. Lang. Cognit.* 1 (2), 67–81. <https://doi.org/10.1017/S1366728998000133>.
- Green, D.W., Abutalebi, J., 2013. Language control in bilinguals: the adaptive control hypothesis. *J. Cognit. Psychol.* 25 (5), 515–530. <https://doi.org/10.1080/20445911.2013.796377>.
- Grosjean, F., 1989. Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain Lang.* 36 (1), 3–15. [https://doi.org/10.1016/0093-934X\(89\)90048-5](https://doi.org/10.1016/0093-934X(89)90048-5).
- Gullberg, M., 1998. Gesture as a communication strategy in second language discourse: a study of learners of French and Swedish [PhD thesis]. In: *Travaux de l'Institut de Linguistique de Lund*, vol. 35. Lund University. <http://lup.lub.lu.se/record/18684>.
- Gullberg, M., 2012. Bilingualism and gesture. In: Bhatia, T.K., Ritchie, W.C. (Eds.), *The Handbook of Bilingualism and Multilingualism*. John Wiley & Sons, pp. 417–437. <https://doi.org/10.1002/9781118332382.ch17>.
- Gullifer, J.W., Kroll, J.F., Dussias, P.E., 2013. When Language switching has No apparent cost: lexical access in sentence context. *Front. Psychol.* 4. <https://doi.org/10.3389/fpsyg.2013.00278>.
- Guo, T., Liu, H., Misra, M., Kroll, J.F., 2011. Local and global inhibition in bilingual word production: fMRI evidence from Chinese-English bilinguals. *Neuroimage* 56 (4), 2300–2309. <https://doi.org/10.1016/j.neuroimage.2011.03.049>.
- Hall, M., 2011. Bilingual picture-word studies constrain theories of lexical selection. *Front. Psychol.* 2. <https://www.frontiersin.org/articles/10.3389/fpsyg.2011.00381>.
- Hamilton, L.S., Huth, A.G., 2020. The revolution will not be controlled: natural stimuli in speech neuroscience. *Language, Cognition and Neuroscience* 35 (5), 573–582. <https://doi.org/10.1080/23273798.2018.1499946>.
- Hartsuiker, R.J., 2015. Visual cues for language selection in bilinguals. In: Mishra, R.K., Srinivasan, N., Huettig, F. (Eds.), *Attention and Vision in Language Processing*. Springer, pp. 129–145. https://doi.org/10.1007/978-81-322-2443-3_8.
- Hasson, U., Egidio, G., Marelli, M., Willems, R.M., 2018. Grounding the neurobiology of language in first principles: the necessity of non-language-centric explanations for language comprehension. *Cognition* 180, 135–157. <https://doi.org/10.1016/j.cognition.2018.06.018>.
- Hatzidakis, A., Branigan, H.P., Pickering, M.J., 2011. Co-activation of syntax in bilingual language production. *Cognit. Psychol.* 62 (2), 123–150. <https://doi.org/10.1016/j.cogpsych.2010.10.002>.
- Hein, R.M., Latoshchik, M.E., Wienrich, C., 2022. Inter- and transcultural learning in social virtual reality: a proposal for an inter- and transcultural virtual object database to be used in the implementation, reflection, and evaluation of virtual encounters. *Multimodal Technologies and Interaction* 6 (7). <https://doi.org/10.3390/mti6070050>. Article 7.
- Hermans, D., 2004. Between-language identity effects in picture-word interference tasks: a challenge for language-nonspecific or language-specific models of lexical access? *Int. J. BiLing.* 8 (2), 115–125. <https://doi.org/10.1177/13670069040080020101>.
- Hermans, D., Bongaerts, T., Bot, K.D., Schreuder, R., 1998. Producing words in a foreign language: can speakers prevent interference from their first language? *Biling. Lang. Cognit.* 1 (3), 213–229. <https://doi.org/10.1017/S1366728998000364>.
- Heyselaar, E., Hagoort, P., Segaert, K., 2017. In dialogue with an avatar, language behavior is identical to dialogue with a human partner. *Behav. Res. Methods* 49 (1), 46–60. <https://doi.org/10.3758/s13428-015-0688-7>.
- Heyselaar, E., Peeters, D., Hagoort, P., 2021. Do we predict upcoming speech content in naturalistic environments? *Language, Cognition and Neuroscience* 36 (4), 440–461. <https://doi.org/10.1080/23273798.2020.1859568>.
- Holler, J., Levinson, S.C., 2019. Multimodal Language processing in human communication. *Trends Cognit. Sci.* 23 (8), 639–652. <https://doi.org/10.1016/j.tics.2019.05.006>.
- Hömke, P., Holler, J., Levinson, S.C., 2018. Eye blinks are perceived as communicative signals in human face-to-face interaction. *PLoS One* 13 (12), e0208030. <https://doi.org/10.1371/journal.pone.0208030>.
- Howard, D., Nickels, L., Coltheart, M., Cole-Virtue, J., 2006. Cumulative semantic inhibition in picture naming: experimental and computational studies. *Cognition* 100 (3), 464–482. <https://doi.org/10.1016/j.cognition.2005.02.006>.
- Huizeling, E., Peeters, D., Hagoort, P., 2022. Prediction of upcoming speech under fluent and disfluent conditions: eye tracking evidence from immersive virtual reality. *Language, Cognition and Neuroscience* 37 (4), 481–508. <https://doi.org/10.1080/23273798.2021.1994621>.
- Jevtović, M., Duñabeitia, J.A., de Bruin, A., 2020. How do bilinguals switch between languages in different interactional contexts? A comparison between voluntary and mandatory language switching. *Biling. Lang. Cognit.* 23 (2), 401–413. <https://doi.org/10.1017/S1366728919000191>.
- Johns, M.A., Steuck, J., 2021. Is codeswitching easy or difficult? Testing processing cost through the prosodic structure of bilingual speech. *Cognition* 211, 104634. <https://doi.org/10.1016/j.cognition.2021.104634>.
- Kendon, A., 2004. *Gesture: Visible Action as Utterance*. Cambridge University Press.
- Kita, S., Özçürek, A., 2003. What does cross-linguistic variation in semantic coordination of speech and gesture reveal? evidence for an interface representation of spatial thinking and speaking. *J. Mem. Lang.* 48 (1), 16–32. [https://doi.org/10.1016/S0749-596X\(02\)00505-3](https://doi.org/10.1016/S0749-596X(02)00505-3).
- Kleinman, D., Gollan, T.H., 2016. Speaking two languages for the price of one: bypassing language control mechanisms via accessibility-driven switches. *Psychol. Sci.* 27 (5), 700–714. <https://doi.org/10.1177/0956797616634633>.
- Kolers, P.A., 1966. Reading and talking bilingually. *Am. J. Psychol.* 79 (3), 357–376. <https://doi.org/10.2307/1420877>.
- Kootstra, G.J., Dijkstra, T., van Hell, J.G., 2020. Interactive alignment and lexical triggering of code-switching in bilingual dialogue. *Front. Psychol.* 11. <https://www.frontiersin.org/article/10.3389/fpsyg.2020.01747>.
- Kroll, J.F., Gollan, T.H., 2014. Speech planning in two languages: what bilinguals tell us about language production. In: Goldrick, M., Ferreira, V.S., Miozzo, M. (Eds.), *The Oxford Handbook of Language Production*. Oxford University Press, pp. 165–181. <https://doi.org/10.1093/oxfordhb/9780199735471.001.0001>.
- Kroll, J.F., Hell, J.G.V., Tokowicz, N., Green, D.W., 2010. The Revised Hierarchical Model: a critical review and assessment. *Biling. Lang. Cognit.* 13 (3), 373–381. <https://doi.org/10.1017/S136672891000009X>.
- Kroll, J.F., Stewart, E., 1994. Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *J. Mem. Lang.* 33 (2), 149–174.
- Kuhlen, A.K., Abdel Rahman, R., 2022. Mental chronometry of speaking in dialogue: semantic interference turns into facilitation. *Cognition* 219, 104962. <https://doi.org/10.1016/j.cognition.2021.104962>.
- Kuhlen, A.K., Brennan, S.E., 2013. Language in dialogue: when confederates might be hazardous to your data. *Psychonomic Bull. Rev.* 20 (1), 54–72. <https://doi.org/10.3758/s13423-012-0341-8>.
- La Heij, W., 2005. Selection processes in monolingual and bilingual lexical access. In: Kroll, J.F., de Groot, A.M.B. (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches*. Oxford University Press, pp. 289–307.
- Levelt, W.J.M., 1989. *Speaking: from Intention to Articulation*. The MIT Press. <https://doi.org/10.7551/mitpress/6393.001.0001>.
- Levelt, W.J.M., 1992. Accessing words in speech production: stages, processes and representations. *Cognition* 42 (1), 1–22. [https://doi.org/10.1016/0010-0277\(92\)90038-J](https://doi.org/10.1016/0010-0277(92)90038-J).

- Levelt, W.J.M., Roelofs, A., Meyer, A.S., 1999. A theory of lexical access in speech production. *Behav. Brain Sci.* 22 (1), 1–38. <https://doi.org/10.1017/S0140525X99001776>.
- Li, C., Ferreira, V.S., Gollan, T.H., 2022. Language control after phrasal planning: playing Whack-a-mole with language switch costs. *J. Mem. Lang.* 126, 104338. <https://doi.org/10.1016/j.jml.2022.104338>.
- Liu, C., Timmer, K., Jiao, L., Yuan, Y., Wang, R., 2019. The influence of contextual faces on bilingual language control. *Q. J. Exp. Psychol.* 72 (9), 2313–2327. <https://doi.org/10.1177/1747021819836713>.
- Lugrin, B., Pelachaud, C., Traum, D., 2022. The handbook on socially interactive agents: 20 Years of research on embodied conversational agents, intelligent virtual agents, and social robotics. In: *Interactivity, Platforms, Application*, vol. 2. Morgan & Claypool.
- Lupker, S.J., 1979. The semantic nature of response competition in the picture-word interference task. *Mem. Cognit.* 7 (6), 485–495. <https://doi.org/10.3758/BF03198265>.
- Macnamara, J., Krauthammer, M., Bolgar, M., 1968. Language switching in bilinguals as a function of stimulus and response uncertainty. *J. Exp. Psychol.* 78 (2), 208–215. <https://doi.org/10.1037/h0026390>.
- Mädebach, A., Kurtz, F., Schriefers, H., Jescheniak, J.D., 2020. Pragmatic constraints do not prevent the co-activation of alternative names: evidence from sequential naming tasks with one and two speakers. *Language, Cognition and Neuroscience* 35 (8), 1073–1088. <https://doi.org/10.1080/23273798.2020.1727539>.
- Mahon, B.Z., Costa, A., Peterson, R., Vargas, K.A., Caramazza, A., 2007. Lexical selection is not by competition: a reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *J. Exp. Psychol. Learn. Mem. Cognit.* 33 (3), 503–535. <https://doi.org/10.1037/0278-7393.33.3.503>.
- Marcos, L.R., 1979. Nonverbal behavior and thought processing. *Arch. Gen. Psychiatr.* 36 (9), 940–943. <https://doi.org/10.1001/archpsyc.1979.01780090026003>.
- Martin, C.D., Molnar, M., Carreiras, M., 2016. The proactive bilingual brain: using interlocutor identity to generate predictions for language processing. *Sci. Rep.* 6, 26171. <https://doi.org/10.1038/srep26171>.
- McNeill, D., 1992. *Hand and Mind: what Gestures Reveal about Thought*. University of Chicago Press.
- Meuter, R.F.I., Allport, A., 1999. Bilingual Language switching in naming: asymmetrical costs of language selection. *J. Mem. Lang.* 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. *J. Mem. Lang.* 35 (4), 477–496. <https://doi.org/10.1006/jmla.1996.0026>.
- Miozzo, M., Caramazza, A., 2003. When more is less: a counterintuitive effect of distractor frequency in the picture-word interference paradigm. *J. Exp. Psychol. Gen.* 228–252. <https://doi.org/10.1037/0096-3445.132.2.228>.
- Mirault, J., Guerre-Genton, A., Dufau, S., Grainger, J., 2020. Using virtual reality to study reading: an eye-tracking investigation of transposed-word effects. *Methods in Psychology* 3, 100029. <https://doi.org/10.1016/j.metip.2020.100029>.
- Molnar, M., Ibáñez-Molina, A., Carreiras, M., 2015. Interlocutor identity affects language activation in bilinguals. *J. Mem. Lang.* 81, 91–104. <https://doi.org/10.1016/j.jml.2015.01.002>.
- Momma, S., Ferreira, V.S., 2019. Beyond linear order: the role of argument structure in speaking. *Cognit. Psychol.* 114, 101228. <https://doi.org/10.1016/j.cogpsych.2019.101228>.
- Momma, S., Slevc, L.R., Phillips, C., 2015. The timing of verb selection in Japanese sentence production. *J. Exp. Psychol. Learn. Mem. Cognit.* 42 (5), 813. <https://doi.org/10.1037/xlm0000195>.
- Mondada, L., 2016. Challenges of multimodality: language and the body in social interaction. *J. Sociolinguistics* 20 (3), 336–366. <https://doi.org/10.1111/josl.112177>.
- Myers-Scotton, C., 2006. Natural codeswitching knocks on the laboratory door. *Biling. Lang. Cognit.* 9 (2), 203–212. <https://doi.org/10.1017/S1366728906002549>.
- Nicoladis, E., Mayberry, R.I., Genesee, F., 1999. Gesture and early bilingual development. *Dev. Psychol.* 35, 514–526. <https://doi.org/10.1037/0012-1649.35.2.514>.
- Nölle, J., Kirby, S., Culbertson, J., Smith, K., 2020. Does environment shape spatial language? A virtual reality experiment. In: Ravignani, A., Barbieri, C., Martins, M., Flaherty, M., Jadoul, Y., Lattenkamp, E., Little, H., Mudd, K., Verhoef, T. (Eds.), *The Evolution of Language: Proceedings of the 13th International Conference (EvoLang13)*, pp. 321–323.
- Özçalışkan, Ş., 2016. Do gestures follow speech in bilinguals' description of motion? *Biling. Lang. Cognit.* 19 (3), 644–653. <https://doi.org/10.1017/S1366728915000796>.
- Özyürek, A., 2014. Hearing and seeing meaning in speech and gesture: insights from brain and behaviour. *Phil. Trans. Biol. Sci.* 369 (1651), 20130296. <https://doi.org/10.1098/rstb.2013.0296>.
- Peeters, D., 2018. A standardized set of 3-D objects for virtual reality research and applications. *Behav. Res. Methods* 50 (3), 1047–1054. <https://doi.org/10.3758/s13428-017-0925-3>.
- Peeters, D., 2019. Virtual reality: a game-changing method for the language sciences. *Psychonomic Bull. Rev.* 26 (3), 894–900. <https://doi.org/10.3758/s13423-019-01571-3>.
- Peeters, D., 2020. Bilingual switching between languages and listeners: insights from immersive virtual reality. *Cognition* 195, 104107. <https://doi.org/10.1016/j.cognition.2019.104107>.
- Peeters, D., Dijkstra, T., 2018. Sustained inhibition of the native language in bilingual language production: a virtual reality approach. *Biling. Lang. Cognit.* 21 (5), 1035–1061. <https://doi.org/10.1017/S1366728917000396>.
- Peeters, D., Runnqvist, E., Bertrand, D., Grainger, J., 2014. Asymmetrical switch costs in bilingual language production induced by reading words. *J. Exp. Psychol. Learn. Mem. Cognit.* 40 (1), 284–292. <https://doi.org/10.1037/a0034060>.
- Penfield, W., Roberts, L., 1959. *Speech and Brain Mechanisms*. Princeton University Press.
- Philipp, A.M., Gade, M., Koch, I., 2007. Inhibitory processes in language switching: evidence from switching language-defined response sets. *Eur. J. Cognit. Psychol.* 19 (3), 395–416. <https://doi.org/10.1080/09541440600758812>.
- Pianzola, F., Bálint, K., Weller, J., 2019. Virtual reality as a tool for promoting reading via enhanced narrative absorption and empathy. *Scientific Study of Literature* 9 (2), 163–194. <https://doi.org/10.1075/ssol.19013.pia>.
- Poullisse, N., Bongaerts, T., 1994. First Language use in second language production. *Appl. Ling.* 15 (1), 36–57. <https://doi.org/10.1093/applin/15.1.36>.
- Prior, A., Gollan, T.H., 2013. The elusive link between language control and executive control: a case of limited transfer. *J. Cognit. Psychol.* 25 (5), 622–645. <https://doi.org/10.1080/20445911.2013.821993>.
- Raghavan, R., Raviv, L., Peeters, D., 2023. What's your point? Insights from virtual reality on the relation between intention and action in the production of pointing gestures. *Cognition* 240, 105581. <https://doi.org/10.1016/j.cognition.2023.105581>.
- Roelofs, A., 1992. A spreading-activation theory of lemma retrieval in speaking. *Cognition* 42 (1), 107–142. [https://doi.org/10.1016/0010-0277\(92\)90041-F](https://doi.org/10.1016/0010-0277(92)90041-F).
- Roelofs, A., 1997. The WEAVER model of word-form encoding in speech production. *Cognition* 64 (3), 249–284. [https://doi.org/10.1016/S0010-0277\(97\)00027-9](https://doi.org/10.1016/S0010-0277(97)00027-9).
- Roelofs, A., 1998. Lemma selection without inhibition of languages in bilingual speakers. *Biling. Lang. Cognit.* 1, 94–95. <https://doi.org/10.1017/S1366728998000194>.
- Roelofs, A., 2014. A dorsal-pathway account of aphasic language production: the WEAVER++/ARC model. *Cortex* 59, 33–48. <https://doi.org/10.1016/j.cortex.2014.07.001>.
- Roelofs, A., 2018a. A unified computational account of cumulative semantic, semantic blocking, and semantic distractor effects in picture naming. *Cognition* 172, 59–72. <https://doi.org/10.1016/j.cognition.2017.12.007>.
- Roelofs, A., 2018b. One hundred fifty years after Donders: insights from unpublished data, a replication, and modeling of his reaction times. *Acta Psychol.* 191, 228–233. <https://doi.org/10.1016/j.actpsy.2018.10.002>.
- Roelofs, A., Piai, V., Garrido Rodriguez, G., 2011. Attentional inhibition in bilingual naming performance: evidence from delta-plot analyses. *Front. Psychol.* 2. <https://doi.org/10.3389/fpsyg.2011.00184> article 184.
- Roelofs, A., Verhoef, K., 2006. Modeling the control of phonological encoding in bilingual speakers. *Biling. Lang. Cognit.* 9 (2), 167–176. <https://doi.org/10.1017/S1366728906002513>.
- Rosinski, R.R., Golinkoff, R.M., Kukish, K.S., 1975. Automatic semantic processing in a picture-word interference task. *Child Dev.* 46 (1), 247–253. <https://doi.org/10.2307/1128859>.
- Runnqvist, E., Strijkers, K., Alario, F.-X., Costa, A., 2012. Cumulative semantic interference is blind to language: implications for models of bilingual speech production. *J. Mem. Lang.* 66 (4), 850–869. <https://doi.org/10.1016/j.jml.2012.02.007>.
- Runnqvist, E., Strijkers, K., Costa, A., 2019. Error-based learning and lexical competition in word production: evidence from multilingual naming. *PLoS One* 14 (3), e0213765. <https://doi.org/10.1371/journal.pone.0213765>.
- Sánchez, L.M., Philipp, A.M., Struys, E., Declerck, M., 2023. Bilingual Language production: a tale about interference resolution in different linguistic contexts. In: Hartsuiker, R.J., Strijkers, K. (Eds.), *Language Production*. Routledge. <https://doi.org/10.4324/9781003145790>.
- Sánchez, L.M., Struys, E., Declerck, M., 2022. Ecological validity and bilingual language control: voluntary language switching between sentences. *Language, Cognition and Neuroscience* 37 (5), 615–623. <https://doi.org/10.1080/23273798.2021.2016873>.
- Schriefers, H., Meyer, A.S., Levelt, W.J.M., 1990. Exploring the time course of lexical access in language production: picture-word interference studies. *J. Mem. Lang.* 29 (1), 86–102. [https://doi.org/10.1016/0749-596X\(90\)90011-N](https://doi.org/10.1016/0749-596X(90)90011-N).
- Schriefers, H., Teruel, E., Meinshausen, R.M., 1998. Producing simple sentences: results from picture-word interference experiments. *J. Mem. Lang.* 39 (4), 609–632. <https://doi.org/10.1006/jmla.1998.2578>.
- Sebanz, N., Bekkering, H., Knoblich, G., 2006. Joint action: bodies and minds moving together. *Trends Cognit. Sci.* 10 (2), 70–76. <https://doi.org/10.1016/j.tics.2005.12.009>.
- Segal, D., Prior, A., Gollan, T.H., 2021. Do all switches cost the same? Reliability of language switching and mixing costs. *Journal of Cognition* 4 (1), 3. <https://doi.org/10.5334/joc.140>.
- Slater, M., 2009. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Phil. Trans. Biol. Sci.* 364 (1535), 3549–3557. <https://doi.org/10.1098/rstb.2009.0138>.
- Smith, S.A., Seitz, S.R., Koutnik, K.H., Mckenna, M., Garcia, J.N., 2020. The “work” of being a bilingual: exploring effects of forced language switching on language production and stress level in a real-world setting. *Appl. Psycholinguist.* 41 (3), 701–725. <https://doi.org/10.1017/S0142716420000259>.
- Tarłowski, A., Wodniecka, Z., Marzecová, A., 2013. Language switching in the production of phrases. *J. Psycholinguist. Res.* 42 (2), 103–118. <https://doi.org/10.1007/s10936-012-9203-9>.
- Timmer, K., Calabria, M., Costa, A., 2019a. Non-linguistic effects of language switching training. *Cognition* 182, 14–24. <https://doi.org/10.1016/j.cognition.2018.09.001>.
- Timmer, K., Christoffels, I.K., Costa, A., 2019b. On the flexibility of bilingual language control: the effect of language context. *Biling. Lang. Cognit.* 22 (3), 555–568. <https://doi.org/10.1017/S1366728918000329>.

- Timmer, K., Grundy, J.G., Bialystok, E., 2017. The influence of contextual cues on representations in the mental lexicon for bilinguals. In: Libben, M., Goral, M., Libben, G. (Eds.), *Bilingual Processing and Acquisition*, vol. 6. John Benjamins, pp. 123–142. <https://doi.org/10.1075/bpa.6.06tim>.
- Titus, A., & Peeters, D. (under review). Multilingualism at the Market: A Pre-registered Immersive Virtual Reality Study of Bilingual Language Switching.
- Tromp, J., Klotzsche, F., Krohn, S., Akbal, M., Pohl, L., Quinque, E.M., Belger, J., Villringer, A., Gaebler, M., 2020. OpenVirtualObjects: an open set of standardized and validated 3D household objects for virtual reality-based research, assessment, and therapy. *Frontiers in Virtual Reality* 1. <https://www.frontiersin.org/articles/10.3389/frvir.2020.611091>.
- Tromp, J., Peeters, D., Meyer, A.S., Hagoort, P., 2018. The combined use of virtual reality and EEG to study language processing in naturalistic environments. *Behav. Res. Methods* 50 (2), 862–869. <https://doi.org/10.3758/s13428-017-0911-9>.
- Vaughan-Evans, A., 2023. External non-linguistic cues influence language selection during a forced choice task. *Biling. Lang. Cognit.* 26 (1), 193–201. <https://doi.org/10.1017/S136672892200044X>.
- Verhoef, K.M.W., Roelofs, A., Chwilla, D.J., 2009. Role of inhibition in language switching: evidence from event-related brain potentials in overt picture naming. *Cognition* 110 (1), 84–99. <https://doi.org/10.1016/j.cognition.2008.10.013>.
- Verhoef, K.M.W., Roelofs, A., Chwilla, D.J., 2010. Electrophysiological evidence for endogenous control of attention in switching between languages in overt picture naming. *J. Cognit. Neurosci.* 22 (8), 1832–1843. <https://doi.org/10.1162/jocn.2009.21291>.
- Willems, R.M., Peelen, M.V., 2021. How context changes the neural basis of perception and language. *iScience* 24 (5), 102392. <https://doi.org/10.1016/j.isci.2021.102392>.
- Woumans, E., Martin, C.D., Vanden Bulcke, C., Van Assche, E., Costa, A., Hartsuiker, R.J., Duyck, W., 2015. Can faces prime a language? *Psychol. Sci.* 26 (9), 1343–1352. <https://doi.org/10.1177/0956797615589330>.
- Zhang, S., Morris, M.W., Cheng, C.-Y., Yap, A.J., 2013. Heritage-culture images disrupt immigrants' second-language processing through triggering first-language interference. *Proc. Natl. Acad. Sci. USA* 110 (28), 11272–11277. <https://doi.org/10.1073/pnas.1304435110>.
- Zhang, Y., Wang, T., Huang, P., Li, D., Qiu, J., Shen, T., Xie, P., 2015. Free Language selection in the bilingual brain: an event-related fMRI study. *Sci. Rep.* 5 (1), 11704. <https://doi.org/10.1038/srep11704>.