

The Complex Adaptive System Principles model for bilingualism: Language interactions within and across bilingual minds

Luna Filipović

University of East Anglia, UK

John A Hawkins

University of California Davis, USA

Abstract

Aims and objectives/purpose/research question: We propose a model that captures general patterns in bilingual language processing, based on empirical evidence elicited in a variety of experimental studies. We begin by considering what linguistic outputs are logically possible when bilingual speakers communicate based on the typological features of two languages in the bilingual mind. Our aim is to explain why some outputs are more frequent or more likely than others in bilingual language use.

Design/methodology/approach: Our empirically derived multi-factor model combines insights from various empirical studies of different bilingual populations and it includes a variety of methodologies and approaches, such as lexical categorisation, lexical priming, syntactic priming, event verbalisation and memory, historical language change, grammaticality judgments and observational reports.

Data and analysis: We critically discuss both lexical and syntactic processing data, as well as data that reflect bilingual type differences and different communicative situations (i.e. who the bilingual speakers are talking to and for what purpose). Crucially, we explain when the relevant factors collaborate and when they compete.

Originality: There are three main reasons why this paper can be deemed original: 1) it offers a unified model for understanding bilingual language processing that is not focused on a single factor or a single linguistic level, as has most often been the case in the past; 2) it brings together the study of bilingualism from both psycholinguistic and sociolinguistic perspectives and in a unifying fashion, which is rare in the literature; and 3) it creates a platform for testing numerous predictions that are not dependent on any one theory.

Significance/implications: This new model opens up new avenues for research into bilingual language processing for all types of bilingual speakers and in different communicative situations. It captures and explains the variety of outputs in bilingual communication and enables us to make predictions about communicative outcomes.

Corresponding author:

Luna Filipović, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, UK.

Email: l.filipovic@uea.ac.uk

Keywords

Bilingual language processing, bilingualism, complex adaptive system principles (CASP), efficiency, language typology

Introduction

Bilingualism has been extensively researched in a number of disciplines over many years, yet we still do not have a holistic model of how bilingual speakers use and process their respective languages. Studies are often based on a single theoretical approach or methodological perspective, or they examine limited numbers of languages and language types. Research is also compartmentalised by linguistic level, for example, word processing versus syntactic processing. There is also the problem that different disciplines do not ‘talk to each other’ sufficiently, principally linguistics, psychology and sociology, even though psycholinguistics and sociolinguistics exist as cross-disciplinary bridges that should make bilingualism more amenable to interdisciplinarity. Language is both a psychological and a sociological phenomenon, so it is legitimate to frame the study of bilingual language use within either domain. However, we believe the time has come to try and unify the findings from these fields and to capture the many invaluable insights from different strands of bilingualism research. A general model is now needed that has overarching explanatory power and that can help us understand why we get disparate and sometimes conflicting empirical results in different studies. This paper makes a first attempt in this direction.

The biggest challenge for any overall model of bilingualism is that many factors are in play at the same time, including age of acquisition, proficiency, language dominance, frequency of use, language mode, and power status, to name but a few. It is difficult to capture their respective roles and their different strengths in different communicative situations within a single model in a way that is compatible with the highly variable bilingual speaker profiles that have been observed. Our model incorporates a proposal for how to do this.

We begin with a theoretical discussion of what we term the *logic of bilingualism*, by which we mean the identification of possible outcomes that may manifest themselves in bilingual communication when two linguistic systems are active in one mind. We then seek confirmation for these outcomes from reported empirical studies, our own as well as others, with both a psycholinguistic and a sociolinguistic focus. We propose a Complex Adaptive System Principles (CASP) Model for Bilingualism, which builds on and extends our earlier Complex Adaptive System Principles for second language acquisition (CASP for SLA; see Filipović & Hawkins, 2013), and we show how it can account for the empirical observations and provide theoretical underpinnings for bilingualism. Finally, we indicate some ways in which testable predictions can be formulated and investigated in the future.

Some background: Why this model is needed

A multi-factor model is needed that can provide a unifying platform for past and future research. Our aim is to capture, within a single model, the bilingual language phenomena that have traditionally been studied under different disciplinary and subdisciplinary umbrellas (SLA, L1 vs. L2 storage and access, heritage languages, language contact, historical language change, etc.) and described using different constructs (e.g. transfer, convergence, bidirectional influence, etc.). Many previous studies have focused on just a single factor at a time (e.g. proficiency, or type of acquisition) and tested predictions related to it, without examining the other factors that play a role in all situations of bilingual language use (such as e.g. who the bilingual person is speaking to or how formal the communicative situation is). There is now a substantial body of research supporting the

kind of multi-factor model we outline here. It gives a unifying and forest view of this interdisciplinary field (see Hulstijn, 2013, 2015 for insightful discussion of the plethora of perspectives in SLA research alone).

Research on bilingualism is difficult because of the sheer nature of the variation it must tackle, in the language systems being combined, speakers' acquisition histories and proficiency levels, and because of the social aspects of specific communicative situations. Bilingual studies have examined speakers of various competences and acquisition histories, and this makes it hard to find a single and consistent way in which to interpret the findings. Athanasopoulos (2011) formulates the problem very clearly:

For example, deviation from monolingual patterns in early bilinguals may be seen as *convergence to a unitary form*, whereas deviation from monolingual patterns in late bilinguals may be seen as '*failure*' to acquire or use the target construction (Athanasopoulos, 2011: p. 29, our emphasis).

Thus, the same experimental outcomes may be interpreted and labeled differently based on when and how the two languages were acquired by the bilingual populations in individual studies. There is a growing awareness of the need for different disciplines to look at each other's findings. Fields that are ultimately concerned with bilingualism (contact linguistics, heritage linguistics, sociolinguistic typology, translation, language development, bilingual memory, etc.) have been developed largely in isolation from one another. Muysken (2013, p. 710) advocates the need for more unification:

Most researchers will acknowledge that these fields are closely interconnected, through their focus on interaction between languages and its outcomes; however, many of the topics listed have been studied in relative isolation from one another. In actual academic practice, separate conferences are held for most of these sub-disciplines, with different journals and debates, and apparently unrelated conceptual frameworks and terminologies.

Similarly, Jarvis and Pavlenko (2008, pp. 234–235) suggest that 'there is a clear need for more interaction and dialogue between researchers who study cross-linguistic influence phenomena in the areas of language contact, childhood bilingualism, child and adult SLA, and first language attrition.' They go on to say that it would be interesting to compare findings on shared phenomena (e.g. syntactic transfer or borrowing), which are studied from different angles in different disciplines (e.g. SLA research vs. language contact and historical language change).

This is exactly our goal here. We steer clear of conflicting labels, such as 'convergence' versus 'failure to acquire', because they have been used to refer to similar outcomes from two different perspectives (early bilingual speakers vs. L2 learners). Our model aims to capture all bilingual behavior under one roof, so to speak, and show that it is driven by common processing mechanisms that we define here (see **General Principles**). Various social and individual factors then modulate these processing mechanisms accordingly (as we explain in **Discussion**). We explain what performance outcomes are most likely depending on the interaction between two linguistic systems, and when we can expect collaboration between different processing principles versus when they are more likely to compete. Different outcomes depend crucially on the language combinations in question and on the general processing mechanisms that characterise bilingual language use and that are flexible and adjustable depending on the specifics of the communicative situation (i.e. *who speaks what language to whom and when*, as Fishman, 1965 put it).

There have been some prior attempts to model the general ways in which bilingual speakers learn and use their languages. They have mainly been concerned with a single linguistic level at a

time (most often the lexicon; see the Revised Hierarchical Model, Kroll, van Hell, Tokowicz, & Green, 2010; van Hell and Kroll, 2013) or their aim has been to extend insights about monolingual language use to bilingual speakers (see e.g. de Bot's [1992] extension of Levelt's [1989] Speaking Model). In this paper we view bilingual language behavior as a form of *complex adaptive system* in the sense of Gell-Mann (1992) in which multiple factors interact to produce a range of observable outcomes and different kinds of interlanguages. Some other proposals within this general emergentist framework have also been put forward (see e.g. Ellis, 1998; Ellis & Larsen-Freeman, 2009; Larsen-Freeman, 1997; Mellow, 2008; O'Grady, 2005, 2008).¹ More recently, Muysken (2013) has put forward a proposal for a model in the context of historical language change within the framework of Optimality Theory. This model addresses language contact phenomena and the different bilingual optimisation strategies in specific contact situations. Four different strategies are proposed to explain outcomes of language contact, namely:

- (a) maximise L1 structural coherence;
- (b) maximise L2 structural coherence;
- (c) match L1 and L2 patterns where possible; and
- (d) rely on universal principles of language processing.

Muysken (2013, p. 709) contends that 'different outcomes correspond to different interactions of these strategies in bilingual speakers and their communities.'

All these works recognise the need for multiple interacting factors in understanding SLA and bilingualism. They differ in the number and nature of the principles proposed, in their precise formulation, in the predictions they make for interlanguage data, and in the range of data on which they have actually been tested. The empirical support offered hitherto is encouraging but still limited. Because our approach in this paper is a multi-factor and integrative one, similar in spirit to Muysken (2013) and not an oppositional one, as in for example, Long (1993), we do not engage in a critical discussion of the shortcomings of other theories and approaches. Rather, we propose a model that is theory-free, as it were, because it is driven primarily by what the empirical data in numerous methodologically varied studies have shown.

We do not have space here to delve deeply into psycholinguistic and ultimately neurolinguistic issues such as how bilingual speakers store their languages and whether the store is separate or shared, or both separate and shared for different aspects of the lexicon and syntax (see de Groot, 2011 for insightful summaries). Many issues such as this are still unresolved in the field, as are central concepts such as, for example, what is an early bilingual speaker? As Myers-Scotton (2003) points out, different studies assume different cut-off points for early bilingual acquisition, ranging from 5 to 12 years of age. Recent research (e.g. Abutalebi, Cappa, & Perani, 2009) has shown that it is proficiency, rather than age of acquisition, that is the decisive factor in assessing bilingual competence. In addition, some studies have shown that for certain aspects of language, late acquisition can be better (see Dörnyei, 2009 for detailed discussion), so age of acquisition may be a less decisive factor. Furthermore, early acquisition is not much help if the relevant language is not reinforced constantly by frequent use. Incomplete acquisition or attrition can negatively impact proficiency despite early acquisition.

Given the near-impossibility of placing bilingual speakers into clear groups and categories, especially across different studies, we think that bilingualism is best conceived of as a *cline*. Thus, we analyse the linguistic outputs reported in a number of studies dealing with bilingual speakers who are equally or almost equally proficient in both languages,³ as well as with those that have better proficiency in one of their two languages (because e.g. they live in an environment where only one of their languages is spoken by the community, and/or because they are second language

learners at different stages of L2 acquisition⁴). We discuss mainly adult bilingualism studies and occasionally include some studies of child bilingualism in order to indicate how our findings apply there as well.

Introducing CASP for bilingualism

We build here on our earlier CASP for SLA (Filipović & Hawkins, 2013), which addressed SLA specifically. CASP for Bilingualism is a more general model of how bilingual speakers process and use their languages. We still include issues pertaining to second language, and also general issues of native language learning, because we aim to capture a number of different bilingualism types within a common model. We focus here on what the different types of outputs by different types of adult bilingual speakers tell us about their knowledge representations and processing mechanisms. We will refer to certain studies of child bilingualism to illustrate how our model can be supported and developed further.

CASP for Bilingualism, like CASP for SLA, consists of four very general principles that work in sync, sometimes competing and sometimes collaborating, enabling us to make predictions about possible and likely bilingual outcomes in different communicative situations. They are slightly redefined here to encompass all types of bilingual learning and processing. The main difference between CASP for Bilingualism and CASP for SLA lies in the formulation of an additional, fifth general principle, Maximise Common Ground. It subsumes previous instances of positive and negative transfer in the SLA context and what we see as the very similar instances of merger and accommodation between two linguistic systems in the bilingual mind. How exactly mergers and accommodations take place, and to what extent, depends on how our CASP principles work together in conjunction with the various social and individual circumstances that we refer to as *adjustability* and *variability* factors (see **Discussion**). We explain our **General Principles** in turn.

General principles

Here we define the five general principles, (A) to (E), that we believe underlie bilingual speakers' language behavior.

- (A) *Minimise learning effort.* Bilingual speakers prefer to minimise learning effort when they learn the grammatical and lexical properties of their languages, just as monolingual speakers do.

Learning effort can be minimised in a number of ways. It is minimised when grammatical and lexical properties are shared between L1a and L1b, and preexisting knowledge of items learned first in one language can be exploited when learning the other. Learning effort is also minimised when properties in both languages are frequently occurring in the input, which increases their exposure and with it the ease of learning. It is minimised when structural and semantic properties of the languages are simple rather than complex.

There is no suggestion here that learners consciously or intentionally strive to comply with (A), or are taught to comply with it, for this principle or any of the others in this article.

- (B) *Minimise processing effort.* Bilingual speakers prefer to minimise processing effort when using the grammatical and lexical properties of their languages, just as monolingual speakers do.

Even when more complex properties have been acquired, bilingual speakers will still prefer to use simpler properties, just like monolingual speakers do. Infrequent items in the input are not only harder to learn but also harder to access and process. Complex grammatical and lexical properties are also harder to both learn and process. This principle is, in essence, a principle of least effort.

Some specific principles following from these more general principles that are defined for SLA in Filipović & Hawkins (2013) apply here as well. *Maximise structurally and semantically simpler properties* means that we can expect simpler forms and meanings to be more readily learned and used in both languages compared to more complex forms and meanings. Similarly, *Maximise frequently occurring properties* drives the acquisition and use of more frequent forms in both languages more than less frequent ones.

(C) *Maximise expressive power.* Bilingual learners prefer to maximise their expressive power, that is, to be able to formulate in each language whatever thoughts they wish to express, and to perform the same language functions in each language, just as monolingual learners do.

This principle stands in partial opposition to principles (A) and (B), because achieving full expressive power requires sometimes more and sometimes less processing effort, for example, more complex as well simpler expressions, and less frequent ones, contra (B). Complex and less frequent items are also harder to learn, contra (A).

There is another general principle that is sometimes opposed in its predicted outputs to principles (A) and (B), and that involves not expressive power per se, that is, semantics, but the efficient delivery of meanings in real time between interlocutors.

(D) *Maximise efficiency in communication.* Bilingual speakers prefer to maximise efficiency in communication with their interlocutors (production and comprehension), just as monolingual speakers do.

Hawkins (2004, 2009, 2014) has argued that *efficiency* plays a central role in all communication and that grammars and cross-linguistic patterns show clear evidence for its conventionalisation in their rule systems and lexicons. Efficiency, as he defines it, relates to the basic function of language – which is to communicate information from the speaker (S) to the hearer (H) (Hawkins, 2014, pp. 34–35):

1. Communication is efficient when the message intended by S is delivered to H in rapid time and with minimal processing effort.
2. Acts of communication between S and H are generally optimally efficient; those that are not occur in proportion to their degree of efficiency.

There are a number of specific principles that make for efficient communication and for efficiency in grammars in Hawkins' system (see his Minimise Domains, Minimise Forms and Maximise Online Processing) and that lead to testable predictions for both usage and grammars. These principles are supported by performance and grammatical data showing a clear preference for: *speed* in delivering linguistic properties in online processing; *fine-tuning* of structural selections to frequency of occurrence, accessibility and inferencing; and *few online errors* or garden paths. These factors interact, sometimes reinforcing, sometimes opposing one another. The main point to be made in this context is that efficiency is a measure of how quickly, with how little processing effort and with how much online error reduction, the intended message can be delivered by the speaker and understood by the hearer.

There is an interesting relationship between efficiency and complexity (see Hawkins, 2004): efficiency generally results in a preference for structural and grammatical simplicity, and for simplicity of all linguistic items. But sometimes communicative efficiency requires the use of structures that are more complex; for example, when the hearer needs more detailed and explicit information about a referent or event. If the referent is unfamiliar to the hearer, the speaker must use a more complex phrase to identify it rather than a simple pronoun; for example, ‘the man I talked to last night’ instead of a simple ‘he’. Efficiency can be maximised by using simple forms when this is possible, and it can also be maximised by using complex forms if necessary. This kind of efficiency-driven model involves considerable hearer sensitivity on the part of the speaker, in accordance with the insights about cooperation in conversation that have been formulated by Grice (1957) and developed further in neo-Gricean (Levinson, 2000) and post-Gricean theories (Sperber & Wilson, 1986).

Certain uses of language and different genres, such as poetry or narratives, will purposefully delay key information, or exploit ambiguity and other rhetorical devices and will not strive primarily for efficiency in communication. Grice (1957) was very clear that these are departures from normal communicative practices that adhere to his Quantity and Quality maxims, and they involve implicatures of various kinds whereby the hearer recognises the speaker’s intention to give too little or too much information or to say something that is only partially true or downright false. Similarly, in this context, efficiency is the norm, and departures from it need special contexts and conventions.

With respect to bilingualism, an efficiency approach will describe not only its role in communication, as just illustrated, but also, and primarily, the relationship between different linguistic systems in the bilingual mind and their observable outputs in communication. It is our contention that bilingual speakers strive for an efficient resolution of two grammars in one mind.

(E) *Maximise common ground.* Bilingual learners and speakers maximise common grammatical and lexical representations and their associated processing mechanisms in two languages, L1a and L1b, within the grammatical constraints and conventions imposed by each. Specifically:

1. If L1a and L1b share a given construction, grammatical rule or word meaning, and associated processing mechanisms, then these shared entities will be used more frequently in both languages. (These entities may be the preferred or majority pattern in one language and a minority or dispreferred in the other, but they will still be the pattern of choice in the bilingual speaker’s use of both languages.)
2. If L1a and L1b do not share a given construction, grammatical rule or word meaning, and associated processing mechanisms, then common ground will be created by introducing entities from one language into the other. New shared entities will be introduced wherever possible within the constraints of current grammatical and usage conventions for the relevant language.
3. Any violations of a grammatical or usage convention in L1a or L1b that occur when maximising common ground will be in proportion to the environmental and psycholinguistic factors enumerated in **Discussion**.

Bilingual linguistic behaviors: Possible outcomes

The principle of maximising common ground enables us to understand numerous phenomena that have been previously described as ‘convergence’, ‘positive transfer’, ‘bidirectional influence’,

'bilingual-specific language use' or 'in-between performance' (e.g. see Pavlenko, 2014, p. 73), and as 'matching L1 and L2 patterns' in Muysken (2013). They are all driven by efficiency, as we see it, and specifically by the efficiency principles we define and exemplify here. 'Maximise common ground' can also explain why some outputs previously classified as 'negative transfer' can arise and why in most cases there is nothing *negative* about them (unless they impede communication by being severely ungrammatical, for example).⁵ They stem from the interaction of the same principles put forward here. Importantly, efficiency also helps us understand those outcomes in bilingual performance where there does not seem to be any apparent influence of one language onto another, even though it could be expected (e.g. as in Kupisch, 2014).

The most efficient way for bilingual speakers to deal with two linguistic systems at the same time is to use as many of their shared resources as possible. This makes sense because bilingual speakers need to alleviate the cognitive load of simultaneously processing two languages, as Silva-Corvalán (1994, p. 206) explains: 'in language-contact situations, bilingual speakers develop strategies aimed at lightening the cognitive load of having to remember and use two different linguistic systems.'

It is therefore important for our 'maximise common ground' principle to determine as a first step whether common ground is available or not in the grammatical and lexical representations of the two languages; compare (E.1) and (E.2) above, within our 'maximise common ground' principle. By focusing initially on the typological relationship between them, we do not neglect other relevant factors; see (E.3). It is impossible to isolate a bilingual speaker from his personal language history and acquisition background and socio-economic environment, because these factors, as is well-known, play an important role in how the two languages manifest themselves and affect one another. Our point of departure is justified, however, by the fact that the first step for *predicting logically possible outcomes* has to be the establishment of the typological background about the two languages, namely what structures and meanings are available in each for the bilingual mind to use.

When two languages have a word or structure in common, the most likely outcome is for them to be used in both. Sometimes this does not happen, as Kellerman (1986) points out: L2 learners may not believe that a word or a structure from their L1 can be perfectly mapped onto a corresponding word or structure in L2. With balanced bilingual speakers this should be less of an issue because more of the correspondences between languages will have been made due to equal or near-equal exposure to contexts in which words and structures from both are applicable.

More often however, a perfect match will not be available between linguistic systems due to a lack of equivalence in meaning, in form or in form-meaning mapping, or due to a difference in general frequency of use (i.e. a given pattern may be dominant in one language but minor in the other). Even similar languages differ in many subtle yet important ways (see e.g. Filipović & Ibarretxe-Antuñano, 2015 on some intratypological distinctions). In addition, when it comes to frequency of use, we know that bilingual speakers may underuse or overuse certain structures (see Hawkins & Filipović, 2012 for a succinct overview and discussion). Such speakers exhibit different linguistic behavior from monolingual speakers, and this is to be expected because bilingual speakers are not two monolingual speakers in one (Grosjean, 2001). But what are they then, if not two monolingual speakers in one? And why can they be more monolingual-like on some occasions, but not on others?

Most often there are some lexical and grammatical distinctions that are made in one but not the other language of the bilingual speaker. The logical options in this case are two: maximise common ground, or keep the two systems separate and do not maximise common ground. When common ground is made, there are again two options: express the same meaning in both, or do not express the relevant meaning in either. Forms and meanings can be added to the relevant language that does

not have them with the result that both languages have them, or they can be ‘lost’ in the language that has them with the result that both languages end up not having them. We will discuss all these situations below.

CASP for bilingualism at work

In this section we illustrate how many apparently different outcomes reported in different studies, both at the lexical and grammatical levels, can be shown to follow from the same underlying principles of bilingual processing.

Ways to maximise common ground

Maximise common ground means resorting to the same meanings in both languages if they exist and using them with possibly greater frequency than in monolingual populations. Frequency of use will be particularly high when the bilingual speaker is actively using both languages in the so-called bilingual mode (see **Discussion**). If structures and meanings are shared they can become connected in bilingual minds, more reinforced and more readily accessible in both languages than language-specific alternatives.⁶ Whether they will actually be accessed or not in specific situations will vary, as we explain in **Discussion**. If a specific form exists only in the one language but not the other, the most straightforward way to maximise common ground might seem to be by directly importing forms from one language into another (i.e. calquing). Trudgill (2011, p. 49) reports on an example of this in the context of sociolinguistic language change driven by bilingualism. It seems that Afghan Arabic has imported some structural features from its co-territorial neighbors, such as the interrogative suffix *-mi*, an identical form and function of which is available in Turkish but not in mainstream Arabic. But this kind of direct structural borrowing does not happen frequently, judging by published reports. The much more widespread strategy for maximising common ground involves making the same kind of form-meaning distinctions or using the same patterns in both languages. For instance, if certain meanings are expressed that use grammatical devices found in only one of their languages, bilingual speakers can often express these meanings lexically in the other language. This results in a more frequent use of certain lexical items or combinations of lexical items, which can lead to their narrower specialisation, and finally to the grammaticalisation of these lexical items and constructions (see Heine & Kuteva, 2005 for a discussion of many case studies that exemplify this).

A different situation is when a common pattern does exist in both languages but does not work equally or equally well in both, and bilingual speakers expand or diminish usage in the one language to match or approximate to the other. The pattern in question may be the majority one (i.e. preferred/default/high-frequency) in one case and the minority one (dispreferred/marked/low-frequency) in the other, or there may be greater specificity in one language and greater generality in the other; for example, in some lexical field covered by both languages.

Various phenomena have been identified of these different types and discussed under the rubrics of *cross-linguistic influence*, *interference* or *unidirectional transfer*; or in the event that the accommodation works in both directions, the terms *convergence* and *bidirectional transfer* have been used. We illustrate below how they can be explained using our ‘maximise common ground’ principle and how our model predicts the different outputs in this context.

Finally, common ground can be made through a loss of distinctions in both languages. This seems to be the preferred and the most attested option in different experiments on lexical categorisation. In grammar, however, losing distinctions can lead to ungrammaticality and is thus likely to be avoided as much as possible (in proportion to the proficiency-driven ability to avoid ungrammaticality). This outcome may be efficient (if it does not breach grammaticality in either language)

but the loss of some lexical and constructional meanings then leads to weaker expressive power. We exemplify all these situations in detail in the following sections.

Maximising common ground at the level of form

At the level of grammatical form, both addition and loss are logically possible, but we see from the available evidence that the loss of grammatical structures and paradigms is mainly characteristic of bilingual speakers who have lower proficiency in one of the two languages. The addition of grammatical forms is normally characteristic of proficient bilingual speakers rather than of second language learners (see Trudgill, 2011 for details from a historical language change perspective).

If both languages have the same option at the level of form, this option will be the preferred output in both languages even though one of the languages may have other structures, perhaps even more frequent ones. Consider first so-called ‘pro-drop’ versus ‘non-prodrop’ languages (for alternations corresponding to *went to the movies* versus *she went to the movies*). If one language of the bilingual speaker has the option to drop the subject (pro-drop) and the other does not (non-pro-drop), we tend to see more overt subjects in the pro-drop language, this being the common ground between the two systems. This is a well-confirmed finding in many studies and is discussed as ‘pro-drop resetting’ among immigrant communities in the USA (Heine & Kuteva, 2005, p. 99) such as Russian (Schmitt, 2000), Spanish (Myers-Scotton, 2003) and Serbian immigrants (Savić, 1995). It is also documented elsewhere within heritage linguistics research (Polinsky, 1995 for Polish, Tamil and Kabardian; Rappaport, 1990 for American Polish and Fenyvesi, 1994 for American Hungarian). We have a situation here in which a language with two options reduces one of them in order to be aligned with the other, with the result that one common form (the overt subject) can be regularly used in both languages. The principles of ‘minimise processing effort’ (B), ‘maximise efficiency in communication’ (D) and ‘maximise common ground’ (E) are all at work here and lead to the use of common grammatical forms for the expression of common meanings in both languages. On this occasion there is no competition among these principles.

Similarly, for adjective positioning, Nicoladis (2006) has shown that French-English bilingual children are more likely to depart from the dominant French pattern of noun before adjective ordering (*étudiant intelligent* ‘student intelligent’, i.e. *intelligent student*) when speaking both French and English. French also allows the adjective-before-noun pattern, which is the only acceptable pattern in English, although in French it is much less frequent. Therefore, bilingual children who have both languages active use a pattern that works in both, even though it is the less frequent ordering in one of the languages, resulting in less common or even unlicensed patterns compared to those of monolingual French. The key point here is that French allows both patterns, whereas English is more restrictive and allows only one – adjective before noun. Thus, the non-shared pattern is avoided, and common ground is made using the pattern that is shared. Children acquiring both languages thus stretch common ground beyond its limits and this may result on occasion in erroneous uses.

Erroneous uses are rarer with adult bilingual speakers. For instance, adult French-German bilingual speakers generally avoid erroneous uses in French if their French proficiency is very high and French, their stronger language. If it is not, then they do produce erroneous adjective placement in French although the errors are restricted to certain specific contexts only (see Kupisch et al., 2013). This is not surprising. Bilingual speakers can indeed perform like monolingual speakers in their respective languages on certain occasions, especially if they are required in experiments to use only one of the languages in which they are highly proficient (see again Kupisch et al., 2013). Highly proficient bilingual speakers can control their outputs better and they ‘know’ exactly when common ground can or cannot be made (see **Discussion** for further details).

Syntactic priming studies with bilingual speakers (Hartsuiker, Pickering, & Veltcamp, 2004; Hartsuiker & Pickering, 2008; Loebell & Bock, 2003) provide evidence that bilingual speakers are aware of what is or is not shared across two grammars. Processing in one language has been shown to have an effect on processing syntactically similar sentences in the other. Traxler (2012) explains that reactivating the same syntactic structure representation is easier than activating an entirely new representation (see also Tooley, Traxler, & Swaab, 2009; Traxler, 2008; Weber & Indefrey, 2009). However, this sharing appears to be sensitive to the consistency in the order of elements in a sentence. German and English passives do not appear to prime one another because the passive verb in German appears at the end of the clause, whereas in English it is typically at the beginning of the verb phrase following the auxiliary (Loebell & Bock, 2003). On the other hand, German and English double object/prepositional datives do prime one another because they have the same types of words in the same order in both languages. Thus, the bilingual mind seems to ‘know’ what is and what is not shared when it comes to the common ground in word order processing.

These psycholinguistic priming studies help us to understand what happens when word orders are complete mirror image opposites of one another. Maximising common ground is almost impossible in these cases, for instance when one language is English, rigorously SVO and head-initial and the other is Japanese, SOV and head-final. Bilingual speakers of these languages do not maximise common ground, even in the early stages of SLA when their respective language proficiencies are unbalanced.⁷ However, diachronically we know that things can happen that lead precisely to the making of common ground even in this extreme case, namely under conditions of intense language contact and social influence. If a language does not have a ready grammatical construction that is the direct or unambiguous counterpart of a construction in the other language, it can generally come up with one. There is generally some pattern, no matter how minor or obscure in every language, which can serve as a match for a major pattern in another language, or else there are lexical resources that can be recruited to capture the meaning in question. Heine and Kuteva (2005) cite a number of case studies of this type from their research on historical language change to illustrate this point. In New Guinea there have been widespread shifts from VO in the coastal Austronesian languages to the OV of the indigenous New Guinea languages (Ross, 1996, 2001). The reverse shift from OV to VO can be seen among Uto-Aztecan languages like Nahuatl in contact with VO Mesoamerican languages (Gast, 2007), see Hawkins (2014, pp. 85–89 for general discussion). Heine and Kuteva (2005) point out that a minor word order in one language can be generalised into a major pattern to match the other language. Similarly, Polinsky (1995) reports that in two strict SOV languages, Tamil and Kabardian spoken by heritage speakers in America, ‘there seems to be a variation between SVO and SOV, which in itself is an indication that the verb-final order is weakening’ (Polinsky (1995, p. 108). Maximising common ground in bilingual usage has become a pathway for major language change under language contact in these cases.

Maximising common ground for grammatical and lexical meanings

As with the grammatical forms discussed in the previous section, form-meaning mappings and distinctions that are regularly made in one language but not the other can be introduced by bilingual speakers from one into the other. Alternatively, one of the languages in bilingual usage can ‘lose’ relevant semantic distinction in order to align better with the other language. At the lexical level, we find evidence for both addition and loss of distinctions so that both languages either end up with the same distinctions or else neither has them because the one that had them lost them in order to maximise common ground.

For instance, we find that evidentiality in an L1a like Turkish (which makes regular grammatical distinctions between statements depending on their source and reliability and on the speaker’s

confidence in their truth) can be expressed using more general constructions and lexical items that are present in L1b (e.g. English). This can be done in different ways; for example, by prefacing statements with 'it seems that ...' or 'it may be that...' or by using certain modal verbs and constructions with much greater frequency (e.g. *may have been*) or by adding adverbials such as 'apparently' etc. A bilingual speaker might use multiple strategies on different occasions or just stick to a single strategy. Heine and Kuteva (2005, pp. 73–74) give an example of Tariana-Portuguese bilingual speakers. Tariana is a North Arawak language of northwest Amazonia that has an obligatory paradigm for tense and evidentiality marking. Tariana speakers all use Portuguese, the official language of Brazil, as a lingua franca. Heine and Kuteva point out that Tariana speakers tend to replicate their evidentiality system when speaking Portuguese using evidential markers 'more frequently and developing them into what appear to be incipient categories for which there is no equivalent in Standard Portuguese' (Heine & Kuteva, 2005, p. 73). Aikhenvald (2002, pp. 315–316) provides examples from Tariana Portuguese showing how evidentiality marking is proceeding from being an incipient category to becoming a full category. Slobin (2016) reports that the introduction of evidentiality from Turkish into English is driven by recency effects. In our terms, maximising common ground by extending English usage in this way is particularly common after a holiday to Turkey or after a trip to a place where use of both languages was constant and intense. 'Maximise common ground' operates here, along with 'maximise efficiency in communication', to create a common and efficient set of grammatical and lexical representations for the two languages in the bilingual mind that can serve as a basis for 'thinking for speaking' (to use Slobin's term). This makes possible similar encoding strategies in the production of both languages, and whatever extra processing effort is then required in the language (here English) that adds extra evidential distinctions that would not normally be expressed by monolingual speakers is offset by the greater advantages of making the common ground. 'Maximise common ground' operates constantly to reconcile the grammatical and lexical representations of two whole languages and finding and maximising their common ground makes for more efficient mental representations of both. Thinking for speaking then becomes more efficient for bilingual speakers, because only one common system needs to be accessed and produced. The result is plausibly less processing effort overall, and greater efficiency, even when it leads to the expression of some non-obligatory forms that require additional processing effort on specific occasions.

Putting this in terms of our CASP model of cooperating and competing principles, 'maximise common ground', 'maximise efficiency in (bilingual) communication' and 'minimise processing effort' all work together to encourage shared mental representations and shared thinking for speaking among bilingual speakers. These principles confer general benefits for bilingual storage and usage, and evidential marking then becomes favoured in a system where it would not normally be used by English monolingual speakers. The processing effort required for this particular structure (and other non-obligatory form-meaning mappings like it) is clearly trumped by the general efficiency advantages afforded by the other principles. The extent to which common ground will be made in cases such as this depends again on the specific internal and external factors that we lay out in **Discussion**, but the key claim is that a range of possible outputs is predicted in our model. Bilingual speakers exhibit this common ground behavior to different degrees based on their circumstances. The fact that they make common ground at all when they do not have to (when speaking a language without the relevant obligatory distinctions) indicates that habitual bilingual language use favours the common ground pattern.

There are many further examples in the literature of bilingual speakers adjusting their form-meaning mappings and distinctions from one language to another. This has been observed among Spanish-English bilingual speakers who introduce a clear-cut distinction between intentional and non-intentional constructions into English to capture an important meaning difference that is

regularly made in Spanish. Spanish has two different constructions, one for intentional events, (*Ella rompió la botella* ‘She broke the bottle’) and the other for non-intentional events, *Se le rompió la botella* ‘It happened to her that the bottle broke’). They are used consistently to map onto the two different event types by monolingual Spanish speakers and by bilingual speakers who have either nearly equal proficiency in both languages (Filipović, under review) or stronger proficiency in Spanish than in English (Filipović, 2018). English monolingual speakers as well as English L1/Spanish L2 bilingual speakers do not draw these distinctions consistently and they use constructions that are ambiguous (e.g. *She broke the glass* – on purpose or not?) or underspecified constructions (*The glass broke* – no information on agency or intentionality). Note that in English both *She broke the glass* and *The glass broke* can be used in descriptions of either event type (intentional or non-intentional). English speakers may assign less intentionality to the inchoative construction *The bottle broke* than to *She broke the bottle* but inchoative constructions are still used by native speakers to describe stimuli that depict intentional events (as in *She crushed the plastic cup and it broke*) in the same way that SVO constructions can be used for non-intentional events (*The woman knocked the bottle off the table*, see Filipović, 2013a, 2013b, 2018 for details).

As in the case of grammatical forms, bilingual speakers can maximise common ground by using the pattern that is applicable in both languages. In a study of caused motion lexicalisation in English and French, Engemann, Harr, & Hickmann (2012) point out that the verbalisations of bilingual speakers differ significantly from those of monolingual speakers. Namely, bilingual speakers over-use cross-linguistically overlapping patterns, in this case the common strategies available to both English and French for describing caused motion events, which in this case is the English-style ‘manner verb + direction particle’ pattern as in: *X pulls the case to the bottom* = *X tire la valise jusqu’en la bas* (see Engemann, Harr, & Hickmann, 2012 for details).

Therefore, the patterns that work in both languages seem to be reinforced. This finding is confirmed in numerous other studies in various domains in addition to that of motion events; for example, see Müller & Hulk, 2001; Nicol, Teller, & Greth, 2001; Lai, Garrido-Rodriguez, & Narasimhan, 2014; Filipović, 2011, 2014; Flecken, 2011; Torbio, 2004. The findings from these studies all support ‘maximise common ground’ in our terms and interestingly its effects can be seen not only in preferred language usage, but in bilingual event encoding and memory as well, indicating more generally that grammatical and lexical representations strive for common ground in the bilingual mind. Filipović (2011) has shown that regardless of whether L1a or L1b is the stronger language, the pattern for motion lexicalisation that is shared by speakers of both languages is the one that guides their behavior in experiments and also their memory for motion events. Bilingual speakers in this study used the Spanish pattern (*‘salió corriendo’* = ‘*exited running*’) and not the English one (*‘ran out’*) both for encoding in production and as an aid for memory regardless of which language they were speaking during the experiment (i.e. English or Spanish). Clearly these bilingual speakers were maximising common ground because it is the Spanish pattern that is licensed in both languages, whereas the English pattern is blocked in Spanish (see Filipović, 2007, 2008; Slobin, 1996; 1997, 2003; Talmy, 1985).

Similar patterns stemming from ‘maximise common ground’ are found in bilingual lexical categorisation and lexical use whereby bilingual outputs have often been observed to illustrate a kind of ‘in-between performance’ in comparison to the outputs of monolingual populations. These outcomes are usually termed ‘convergence’. For example, Ameel, Storms, Malt, & Sloman (2005) have shown that if one language uses a general term when the other prefers to make more fine-tuned distinctions, the most common specific term will be extended in the direction of the general term in the other language, thus making the mapping from words to concepts more similar in online language use. Thus, whereas the referents of Dutch *fles* are typically differentiated by French monolingual speakers into *bouteille* (larger bottle) and *flacon* (smaller bottle),

Dutch-French bilingual speakers assign most bottles to the category *bouteille* in French and only a few to the category *flacon*. Pavlenko (2014) states that it may be language dominance in Dutch that drives this outcome. Similar findings have been reported by Pavlenko & Malt (2011) in the naming of kitchen objects in English and Russian whereby the English L2 affects the categorisation in Russian L1. Alfernik & Gullberg (2014) present similar results with regard to bilingual use of posture verbs in Dutch and French, where the more distinctive Dutch categories become less so under the influence of a single category in French. This kind of shared bilingual lexical categorisation is driven by the ‘minimise processing effort’ principle and by ‘maximise efficiency in communication’ in addition to ‘maximise common ground’. Processing effort is minimised by having a single and shared set of lexical categories, and the resulting mental representations and their outputs in communication are more efficient because one set of distinctions works for both languages. These principles may sometimes act against ‘maximise expressive power’ in one of the languages because of the loss of distinctions that monolingual speakers have at their disposal. But because more general terms like Dutch *fles* can always be used for more specific types of bottles, the disadvantage of not drawing finer distinctions is relatively weak and the advantages of ‘maximise common ground’, ‘minimise processing effort’ and ‘maximise efficiency in communication’ evidently prevail.

What is the basic cause of these changes in the ways bilingual speakers recategorise objects and events across different semantic systems? Alfernik & Gullberg (2014) note that different approaches have been proposed in attempts to define the kinds of overlaps that are necessary between meanings, such as closeness in conceptual space (Gathercole & Moawad, 2010), conceptual equivalence (Berthele, 2012), or rough translation equivalence (Ameel, Malt, Storms, & van Assche, 2009). Alfernik & Gullberg (2014) also point out that such general descriptions do not say much about the degree of change, such as whether there is a slight shift in category boundaries versus a complete drop of a semantic feature, or about the precise type of change, such as a shift in distributional frequencies of use versus merging two categories into one.

A potential underlying principle has been proposed by Muysken (2000), who suggests that ‘processing economy’ is a possible force behind the move toward a more general system (also voiced in Ameel et al., 2009). We suggest that *it is processing efficiency rather than processing economy* that underlies this ‘uniformising tendency’ (Muysken, 2000, p. 277), or these cases of ‘maximising common ground’, as we call them here. As we explained in **General Principles**, processing simplicity or economy is not always more efficient. More complex and less economical structures for processing are sometimes more efficient; for example, when using *the man I talked to last night* instead of *he* as a referring expression, and also plausibly when adding evidential distinctions to a language like English in bilingual interactions with Turkish. Efficiency does not equal economy, therefore. Economy is only one part of efficiency.

Overall, ‘losing’ distinctions in both languages in order to maximise common ground is a less optimal alternative to gaining or adjusting meaning distinctions in one or the other because it can lead to a loss in expressive power, and also to not sounding authentic in either language. More severely, it can lead to ungrammaticality. Some distinctions are easier to ‘lose’ than others, however. For instance, major word order differences cannot be swept under the carpet, but bilingual speakers can avoid using certain grammatical structures that are available in only one of their two languages (e.g. avoidance of Spanish-specific non-agentive affective dative *se*-constructions by English learners of L2 Spanish; see Filipović, 2018).

There are other types of situations where two patterns are equally available to bilingual speakers. In such cases the preference can be driven by external factors (see further details in **Discussion**), such as language dominance or the language of the environment in which they live. An example of this can be seen in studies of the processing mechanisms that are associated with

grammatical representations in the different languages bilingual speakers use. In syntactic processing, Dussias (2001, 2003) and Dussias & Sagarra (2007) found that Spanish-English bilingual speakers use a single, shared parsing mechanism (favoured by the language of their living environment) for relative clause attachment in sentences such as *the waiter shot the brother of the actress who was standing on the balcony*. This is significant because different parsing preferences have been observed in monolingual populations (high attachment in Spanish, i.e. to the noun *brother*, vs. a low attachment preference in English, i.e. to the noun *actress*). All these outcomes are compatible with our model. It is more efficient to use one and the same syntactic processing mechanism (regardless of which one in the case of syntactic attachment, because both can succeed in resolving ambiguities through a default preference) rather than having to constantly switch between them, especially when both are available in different parts of the processing system.

Discussion: Variability and adjustability in bilingual learning and use

The principles we have put forward are modulated by both internal and external factors. Internal factors include age of acquisition, proficiency, learning environment, type of input, and frequency of use, which can all sway our predictions in different ways and make bilingual outputs *variable* (see especially Jarvis & Pavlenko, 2008 for detailed discussion of a number of these factors). External factors are driven by the inherently *adjustable* nature of bilingual linguistic behavior, which depends on the interlocutor types involved (i.e. who bilingual speakers are talking to) or the type of communicative situation a bilingual is involved in (e.g. formal vs. informal; see Dewaele, 2001). For instance, the same bilingual speaker will produce different outputs when talking to another bilingual speaker of the same two languages compared to communicating with a monolingual of one of the two languages (see also discussion in Muysken, 2013, p. 714 on different factors that impact outputs in language contact situations).

We have put forward a CASP model for Bilingualism with the initial premise that both languages in the bilingual mind are readily available and accessible in any given communicative situation. This assumption is supported by substantial evidence from contemporary research on bilingualism, whereby both comprehension and production studies have shown that bilingual speakers (highly proficient and second language learners) ‘activate information about both languages when using one language alone’ (Kroll & Bialystok, 2013, p. 497). Our focus has been on the *logically possible outcomes when multiple linguistic options are each ‘vying for supremacy’ on-line*. We have discussed the empirical manifestations of the competition and collaboration among principles in our model. We have noted repeatedly that the reported outcomes and the ways in which the principles operate seem to vary across the different situations that involve different kinds of bilingual speakers and different circumstances in which they interact. We believe, and we propose here, that the same principles still apply in all bilingual language processing, except that the degree to which they are active is modulated by the linguistic profile of our bilingual speaker (which is variable; bilingual speakers have different backgrounds and competencies) and who our bilingual is talking to (bilingual speakers adjust to the situation as best they can, depending on their linguistic profile). In the terminology of this paper, we need to consider bilingualism both *within and across minds*. Long-term language change is also crucially dependent on these two factors because, when certain types of bilingual speakers, for example, L2 users, interact with monolingual speakers of the bilingual speakers’ L2 and crucially, if the bilingual speakers outnumber the monolingual speakers in a community, then the monolingual speakers’ language may change under the influence of those that speak it as an L2 (see Trudgill, 2011).

Our model can readily accommodate and incorporate internal factors such as relative language proficiency in each language, and it does so in the definition of the ‘maximise common ground’ principle above (see (E.3) within this principle). Proficient bilingual speakers are expected to make greater and more accurate maximisation of the common ground because they have more (unconscious) knowledge of what is shared and grammatical in both languages. In less proficient bilingual speakers, such as L2 learners, exposure to and knowledge of the L2 is still incomplete and so the full potential of this principle may not yet be realised and thus the efficiency benefits deriving from it may not be there yet either. Some bilingual speakers can still try to maximise common ground even though it can lead to ungrammaticality and communication breakdown because their language proficiency is uneven. Perhaps with explicit instruction that raises awareness about the typological similarities and differences between languages the beneficial processing strategies for L2 learners can be enhanced (see Filipović, 2018 for discussion). The best pedagogy for second language teaching is one of the applications of our model that merits future research. Our model also makes predictions for different outcomes in ‘balanced’ versus ‘unbalanced’ bilingual speakers that can be tested psycholinguistically (see below).

External factors such as who the bilingual speaker is talking to and with what purpose also affect how our principles operate. For instance, we expect fewer instances in which the ‘maximise common ground’ principle is active when our bilingual speaker is talking to a monolingual speaker, and more instances when he or she is talking to two monolingual speakers in the two different languages. The profile of the communicative situation can also affect the outputs, as Dewaele (2001) has pointed out. A more formal situation can result in fewer instances of common ground than an informal one because the output monitor (see de Groot, 2011) may be on a higher alert in the former rather than the latter. Finally, these situation-driven outputs are also compounded by the internal factors of relative proficiency: bilingual speakers with equal proficiency in both languages can be more successful in output monitoring than those whose proficiency in both languages is unequal.

The construct of *language mode* is of special interest in this context because it captures the need to acknowledge that bilingual speakers’ performance varies depending on the communicative situation they are in. Language mode was first proposed and defined by Grosjean (2001) as ‘the state of activation of the bilingual’s language and language processing mechanism’, and it can be affected by a number of different factors. It is important to note again that proficiency plays a role when it comes to the level of control that can be exercised over language mode. Bilingual speakers who are highly dominant in one language may simply not be able to control language mode in the same way as balanced bilingual speakers. Although they may deactivate their stronger (L1) language in a monolingual environment that requires them to use only their weaker (L2) language, their L2 language may simply not be active or developed enough to allow them to stay in a monolingual mode (Grosjean, 2001, p. 21).

There has been some criticism addressed to the notion of language mode and how to determine it, and more importantly, a question has been raised about whether it is the fluctuating language mode that is in charge of variation in bilingual linguistic behavior or something else; for example, a kind of *conscious output monitor* that regulates the output rather than language mode per se (see de Groot, 2011, p. 288, and also p. 290 on the circularity of language mode argumentation).⁸ As Dewaele (2001) has noted, students in a bilingual language mode can show very different language behavior (including code-switching types and instances) depending on the current communicative context (e.g. formal vs. informal). Green and Wei (2014) have thrown light on the role of cognitive control processes in different types of bilingual code-switching. It remains to be seen whether this adaptability concerns fluctuations in the degree of activation of the bilingual speaker’s two language subsets or fluctuation in the attentiveness of a mental monitor that watches over the output

of the language system. In either event, our CASP for Bilingualism model is compatible with both options. In fact, the two theoretical positions (mode fluctuation vs. mental monitor fluctuations) need not necessarily be viewed as diametrically opposed. Whichever mechanism is responsible for the control over bilingual output, the fact remains that bilingual outputs will depend on the specifics of the communicative situation (e.g. formal vs. informal) and interlocutor profile (e.g. monolingual vs. bilingual).

It is important to point out that what we are talking about here goes beyond saying that speakers are in different language modes (Grosjean, 2001) when they communicate in different situations. Our bilingual speaker can be in a bilingual mode when he is interpreting in a formal context, when speaking to two monolingual speakers in each language in an informal context, when speaking to a relatively incompetent L2 speaker, when speaking a weaker language if proficiency is uneven, and when speaking to bilingual speakers in the same languages. All these situations may result in different outputs even though the mode is always characterised as bilingual. We need to capture more fine-grained distinctions here that stem from our key factor: adjustability. Who are our bilingual speakers interacting with and adjusting their outputs to?

We must also refer in this context to Communication Accommodation Theory, developed by Howard Giles in the 1970s. The basic idea behind this theory is that people change their behavior in order to attune their communication to their partner. This theory is grounded in social psychology and social identity theory (see Giles & Smith, 1979 for details). Although this approach to the study of communication addresses a wider set of issues that concern communication in general and not bilingualism specifically, we must draw a parallel between some of the notions from this framework that bear relevance to our model. Namely, speakers can employ ‘strategies of convergence’ whereby individuals adapt to each other’s communicative behaviors to reduce social differences. In contrast, ‘strategies of divergence’ are employed when individuals accentuate speech and non-verbal differences between themselves and their interlocutors (Giles, Coupland & Coupland, 1991). Too much convergence can lead to over-accommodating that can be interpreted as condescending. This theory has been applied to the study of communication between native and non-native language speakers in SLA (Zuengler, 1991). Interestingly, native speakers have been found to engage in so-called ‘foreigner talk’ when interacting with second language learners, adjusting their language by adopting features such as slower speech rates, shorter and simpler sentences, greater pronunciation articulation etc. (Zuengler, 1991).

There is a parallel here with what bilingual speakers do when engaging with different interlocutors. For instance, they may code-switch more when talking to another bilingual speaker of the same two languages than when speaking to a monolingual speaker. By the same token, their code-switching may be more controlled when talking to other bilingual speakers on a formal occasion compared to informal communicative situations. The conscious output monitor (de Groot, 2011) mentioned earlier may guide this accommodation to the specifics of the communicative situation. Our model captures this adjustability of bilingual speakers and explains how it may depend on factors of variability in individual speakers (such as proficiency in each language), which together determine which of the possible outputs predicted by our model will materialise.

We also need to account for some monolingual-like performances by bilingual speakers. These can occur *when a bilingual speaker is talking to a monolingual speaker in one of the two languages*. Monolingual-like performance is also documented in some psycholinguistic studies. This outcome can be due to the fact that bilingual speakers are kept in monolingual mode (e.g. Kousta, Vinson, & Vigliocco, 2008) or when access to one of the two languages is disabled in the experiment (e.g. Athanasopoulos et al., 2015) or when the proficiency is higher in one of the two languages. Monolingual-like performances are one extreme of our cline, modulated by internal factors

such as proficiency and external factors such as whether or not there is a need to keep both languages constantly active and accessible throughout the communicative situation. The other extreme of the cline is making common ground excessively even when this is not grammatically licensed in one of the two languages, which is more characteristic of SLA outputs. The outputs at either end of the cline are also in line with our principles, modulated again by the same individual factors and depending on who the bilingual speaker is talking to. The most proficient bilingual speakers will also be the most efficient: they know when and how best to make common ground with the least required learning and processing effort whilst achieving maximum expressive power fit for the specific communicative goals. They will have overcome any falsely perceived language distance (see Kellerman's (1986) psychotypology discussion) and they will also not make ungrammatical attempts at making common ground, although some lexical or pragmatic usage conventions may not be adhered to (e.g. through the overuse of more marked and less frequent patterns in one of the two languages due to habitual bilingual use). CASP can help us make concrete predictions here. For example, *maximise efficiency and minimise processing effort* would go against *maximise common ground* and *maximise expressive power* when a proficient bilingual speaker is involved in a monolingual communicative situation and we can therefore expect fewer instances of common ground than in a situation when both languages are active to a similar degree.

Whether and how bilingualism affects other cognitive functions still awaits a holistic scientific account although some evidence is emerging about how bilingualism affects memory for events (e.g. Filipović, 2011, 2013a, 2018) and categorisation (e.g. Lai et al., 2014) as well as having some benefits for executive function (see Kroll & Bialystok, 2013, for a focused overview). What kinds of benefits are possible and realised may vary in different cognitive domains, language combinations and individual bilingual characteristics. These questions can all be investigated in the context of the CASP for Bilingualism model.

Finally, we must address a critical question that arises for all multi-factor models, including the one we are proposing here, and involving their overall power and the falsifiability of their predictions. When numerous factors and principles are proposed, there is a danger that at least one of them can be found to account for any set of bilingual data, making the overall predictions vacuous. We are mindful of this danger. However, we believe that any successful model of bilingualism must be a multi-factor one, and we can show further that our model is readily falsifiable.

The problem with most theoretical work in the field hitherto is that it has not been general enough, and has focused on either social or psychological aspects of bilingualism, on poor versus good proficiency, on some linguistic levels rather than others (e.g. lexical vs. syntactic), often ignoring crucial typological differences between languages, whose potential for common ground is vastly different for different pairs of languages, and which interact in different ways with the social and proficiency factors of the overall model. We believe that the field is full of invaluable data and case studies and insights, but they need to be brought together into a more general and more complete context, and this is what CASP is trying to do. The model has to be general and multi-factor, therefore, and we think it is right to err on the side of power, if need be, in order to capture the multifaceted phenomena at hand.

So how then do we ensure falsifiability? It is indeed possible for a multifactor model to make constrained and falsifiable predictions. The crucial requirement is that we clearly *delimit the relevant factors and principles that apply in any given empirical domain*. So, if a pair of languages is such that common ground in their word orders can be created without offending a grammatical constraint in one or the other language (recall the adjective ordering data among French-English bilingual speakers discussed above and taken from Nicoladis, 2006), then we predict that common ground will be maximised, whereas overriding a grammatical constraint (leading to language change) requires other

relevant factors, e.g. of a social nature, to be considered as well. More generally, if our bilingual speakers are interacting with other bilingual speakers, or with monolingual speakers, or in one language mode versus another, then the relevant social factors (and their independently established general consequences for bilingualism) are invoked, in conjunction with the grammatical contrasts in question, to make one set of predictions rather than another.

Similarly, if proficiency in one language is low rather than high, then ‘maximise common ground’ and the other CASP principles predict different outcomes. In short, it is crucial to identify that subset of factors from the multi-factor model that makes predictions for the case at hand, and to assess how they work together. Sometimes one factor will not interfere with the predictions of another. Sometimes it will. Sometimes they will co-operate and pull in the same direction, sometimes they will compete. Having multiple factors to consider does not make the overall predictions easy, therefore, but it does not make them impossible, as long as we delimit the empirical domain of applicability for each factor and its expected consequences. In any case, we have no choice. Any adequate model of bilingualism must be a multi-factor one.

The reader is referred to Hawkins (2014, ch. 9) for general discussion of co-operation and competition among multiple factors in a grammatical and language processing context, to Filipović & Hawkins (2013) for detailed illustration of the interaction among multiple factors in an SLA context, and to the many papers on grammar, processing and learning in the competing motivations volume edited by MacWhinney, Malchukov, & Moravcsik (2014). The same issue of the relevance of different factors in any given domain, their relative strength and their co-operation or competition, arises in all of these studies.

For example, if we look at typological predictions for the ordering of relative clauses and head nouns across languages, two of Hawkins’ (2014) principles (‘minimise domains’ and ‘maximise online processing’) work together to predict that head-initial languages (like English) will have only post-nominal relatives. This prediction is straightforwardly falsifiable. For head-final languages these principles compete, and are compatible with both pre-nominal and post-nominal relative clauses. But a prediction is still made, despite the competition. Specifically, pre-nominal relatives are predicted in proportion to how strictly and rigidly head-final the language is (like Japanese with its obligatory verb-final rule in clauses) – that is, this is a gradient prediction. Thus, if we discover more pre-nominal relatives in non-rigid head-final languages than in rigid head-final ones, this is a clear counter-example and the prediction has been falsified. This illustrates how multi-factor models make falsifiable predictions, as long as we delimit the relevant domain of their applicability (to e.g. head-initial vs. head-final languages, rigid vs. non-rigid head-final ones, etc.), and clarify the nature and strength of the prediction (e.g. we have an absolute prediction for head-initial languages vs. a gradient one for head-final ones).

Similar reasoning applies to the CASP Model for Bilingualism. Because there is a lot of possible variation, we need to *first restrict the relevant universe of applicability*, as we suggested above; namely, we need to state what the languages are, who is talking to whom (and why), and with what level of proficiency. It is then that we can start to make falsifiable predictions and these predictions will be relative to the grammatical, social and psycholinguistic universe to which they apply and within which they work together, possibly constraining each other or not, as the case may be. For instance, English-Turkish bilingual speakers equally competent in both languages who use both languages in their daily life and work are more likely to express evidentiality in both English (lexically) and Turkish (grammatically), as we discussed in the previous section. This is the most efficient way to be communication-ready in both languages, by making common ground and being able to express all meanings in both languages. If they use one language more than the other habitually, or when they are speaking to a monolingual person, this impacts the extent to which

common ground is made (e.g. when speaking English only to an English monolingual person there will be no consistent lexical expression of evidentiality).

Conclusions and further research

In this paper we have introduced a new multi-factor model, CASP for Bilingualism. As indicated by the name of the model, it assumes that the principles underlying bilingual language behavior reflect a complex adaptive system whereby a number of factors interact in competition or in collaboration. The general principles of CASP for Bilingualism define the mechanisms that characterise language learning, usage and grammar by bilingual speakers, which can then be modulated by the specific circumstances of the bilingual individual (their respective levels of proficiency, acquisition histories and frequencies of using one or the other language) as well as by the specific features of a communicative situation (e.g. bilingual speaking to other bilingual speakers, one monolingual or two monolingual speakers, one in each language). Crucially, as we point out, all bilingual behavior can be described and explained as the consequence of an interaction among general processing and learning principles that are adaptable depending on the internal and external factors that we outline in **Discussion**. These general principles are the underlying driving force of the model, operating at all times, sometimes collaborating, sometimes competing. They apply to all instances of bilingual linguistic behavior, although the degree to which they are manifested varies based on the internal and external factors.

This model creates a common platform for the study of bilingualism that does not describe the same or similar outcomes by different names just because different bilingual participants are involved. In fact, it makes it possible to study bilingualism as the cline that it actually is. For instance, we expect to see the same tendencies in the linguistic outputs of both more and less proficient bilingual speakers, namely those driven by the general principle of ‘maximise common ground’. The degree of success may vary (less proficient bilingual speakers may produce more ungrammatical outputs than more proficient ones in their attempts to stretch the common ground beyond the limits of grammaticality), but the underlying processing mechanism remains the same for both populations.

Finally, we can explain different outcomes in a wide variety of bilingualism research strands, both psycholinguistic and sociolinguistic, and make the desired connections between the various processes within and across bilingual minds that affect language production. We can also explain other aspects of bilingual speakers’ linguistic behavior, such as the relationship between language and memory and historical language change on a wider scale. We make predictions about future developments and the linguistic, psychological and social conditions that underlie different outcomes. For instance, we can calculate which of the possible bilingual outputs is more likely to occur based on its relative efficiency in comparison with other options, and then assess how individual or social circumstances may encourage or discourage the preferred option in the output. The research realm of code-switching and also historical language change can both further inform, and benefit from, applications of the CASP model. Outputs in bilingual interactions are conditioned by the dynamic, complex and adaptive relationships among *general language processing principles and the specific characteristics of communicative situations based on who speaks in which language to whom and why*. From a long-term perspective, these relationships are also the determinants of the directions in which languages in contact evolve. CASP for Bilingualism can help us test, predict, and explain more- versus less-likely outcomes.

It is our hope that other investigators, regardless of theoretical persuasion, will find CASP for Bilingualism useful as a framework for developing and testing their ideas and for contributing to our knowledge about bilingualism.

Acknowledgements

This paper has benefited considerably from anonymous *IJB* reviews and from the comments and encouragement of *IJB* editor Li Wei. We are grateful for this feedback, which has resulted in several revisions and additions being made to the original version. We are also most grateful to Dan Slobin and Jan Hulstijn for their comments on an earlier version, and to faculty and graduate students in linguistics and psycholinguistics at the University of California Davis for their feedback to a presentation of this work in April 2018. All remaining errors and shortcomings are very much our own.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. See also Dörnyei, 2009 for an extensive critical overview of different models, in particular with reference to SLA; see also Basnight-Brown, 2014 for an overview and discussion of models of lexical access and bilingualism.
2. The notion of proficiency is not uncontroversial itself. Hulstijn (2012, p. 423) points out that it is often taken for granted, although there is a great variety of approaches in classifying informants (see also Treffers-Daller, 2015, p. 236 for discussion on this topic, and also Treffers-Daller, 2015, p. 239 on the difference between proficiency and dominance).
3. We bear in mind, however, that perfectly balanced bilingual speakers with entirely equal proficiency in both languages may be somewhat of a myth, a fictional rather than a realistic construct (Cook, 1995; Romaine, 1989).
4. See Treffers-Daller, 2015, p. 248 on defining balanced bilingualism not as an overall category but with respect to specific criteria; for example, vocabulary scores, percentage of use across different domains, or similar.
5. The term ‘transfer’ does not quite reflect what is going on in the bilingual mind because nothing is actually *being transferred* (i.e. ‘taken from one place and placed in another’).
6. The nature and degree of overlap can vary, as Pavlenko (2009) pointed out, e.g. it can be complete or only partial (see Duyck & Brysbaert, 2004 for an example of connectionist modelling for semantic overlap at word level; see also de Groot, 1992; van Hell & de Groot, 1998, and Basnight-Brown, 2014 on the different degree of overlap in translation between concrete and abstract nouns).
7. Ungrammatical outputs in VP word order are costlier to understanding than those in NP word order because they impact the communication of the whole message more severely (see Filipović & Hawkins, 2013 for details). NP word order errors are reported for both children (Nicoladis, 2006; Yip & Matthews, 2007) and adults with one stronger language (Kupisch et al., 2013), although adult bilingual speakers exhibit ungrammaticality in the NP word order in more restricted contexts (Kupisch et al., 2013) than children (Yip & Matthews, 2007).
8. The monitor need not be purely linguistic, but rather the one that is in charge of cognitive control in general. Namely, bilingualism has been demonstrated to afford advantages in executive control across the board (e.g. in working memory and inhibition; see further discussion in the section on cognitive consequences of bilingualism, further down in this chapter). Even if the languages are typologically close, such as two different dialects, the advantages across the executive control system is evident (see Antoniou, Grohmann, Kambanaros, & Katsos, 2016).

References

- Abutalebi, J., Cappa, S. F., & Perani, D. (2009). What can functional neuroimaging tell us about the bilingual brain? In J. F. Kroll, & A. M. B. de Groot (Eds.), *Handbook of bilingualism* (pp. 497–515). Oxford, UK: Oxford University Press.

- Aikhenvald, A. (2002). *Language contact in Amazonia*. Oxford, UK: Oxford University Press.
- Alfernik, I., & Gullberg, M. (2014). French-Dutch bilinguals do not maintain obligatory semantic distinctions: Evidence from placement verb. *Bilingualism: Language and Cognition*, 17, 22–37.
- Ameel, E., Malt, B. C., Storms, G., & van Assche, F. (2009). Semantic convergence in the bilingual lexicon. *Journal of Memory and Language*, 60, 270–290.
- Ameel, E., Storms, G., Malt, B. C., & Sloman, S. A. (2005). How bilinguals solve the naming problem. *Journal of Memory and Language*, 53, 60–80.
- Antoniou, K., Grohmann, K. K., Kambanaros, M., & Katsos, N. (2016). The effect of childhood bilingualism and multilingualism on executive control. *Cognition*, 149, 18–30.
- Athanasopoulos, P. (2011). Cognitive restructuring in bilingualism. In A. Pavlenko (Ed.), *Thinking and speaking in two languages* (pp. 29–65). Bristol, UK: Multilingual Matters.
- Athanasopoulos, P., Bylund, E., Montero-Melis, G., Damjanovic, L., Schartner, A., Kibbe, A., . . . Thierry, G. (2015). Two languages, two minds: Flexible cognitive processing driven by language of operation. *Psychological Science*, 26, 518–526.
- Basnight-Brown, D. M. (2014). Models of lexical access and bilingualism. In R. H. Heredia, & J. Altarriba (Eds.), *Foundations of bilingual memory* (pp. 85–107). New York, NY: Springer.
- Berthele, R. (2012). On the use of PUT verbs by multilingual speakers of Romansh. In A. Kopecka, & B. Narasimhan (Eds.), *Events of 'putting' and 'taking': A crosslinguistic perspective* (pp. 145–166). Amsterdam, Netherlands: John Benjamins.
- Cook, V. (1995). Multi-competence and the learning of many languages. *Language, Culture and Curriculum*, 8(2), 93–98.
- de Bot, K. (1992). A bilingual production model: Levelt's 'Speaking' model adapted. *Applied Linguistics*, 13, 1–24.
- de Groot, A. M. B. (1992). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 1001–1018.
- de Groot, A. M. B. (2011). *Language and cognition in bilinguals and multilinguals: An introduction*. New York, NY: Psychology Press.
- Dewaele, J.-M. (2001). Activation or inhibition? The interaction of L1, L2 and L3 on the language mode continuum. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives* (pp. 69–89). Clevedon, UK: Multilingual Matters.
- Dörnyei, Z. (2009). *The psychology of second language acquisition*. Oxford, UK: Oxford University Press.
- Dussias, P. E. (2001). Sentence parsing in fluent Spanish-English bilinguals. In J. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 159–176). London, UK: Blackwell Publishing.
- Dussias, P. E. (2003). Syntactic ambiguity resolution in second language learners: Some effects of bilinguality on L1 and L2 processing strategies. *Studies in Second Language Acquisition*, 25, 529–557.
- Dussias, P. E., & Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish-English bilinguals. *Bilingualism: Language and Cognition*, 10, 101–116.
- Duyck, W., & Brysbaert, M. (2004). Forward and backward number translation requires conceptual mediation both in balanced and unbalanced bilinguals. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 889–906.
- Ellis, N. C. (1998). Emergentism, connectionism and language learning. *Language Learning*, 48, 631–664.
- Ellis, N. C., & Larsen-Freeman, D. (2009). Constructing a second language: Analyses and computational simulations of the emergence of the linguistic constructions from usage. In N. Ellis, & D. Larsen-Freeman (Eds.), *Language as a complex adaptive system* (pp. 90–125). Chichester, UK: Wiley.
- Engemann, H., Harr, A.-K., & Hickmann, M. (2012). Caused motion events across languages and learner types: A comparison of bilingual first and adult second language acquisition. In L. Filipović, & K. M. Jaszczolt (Eds.), *Space and time in languages and cultures* (pp. 263–288). Amsterdam, Netherlands: John Benjamins.
- Fenyvesi, A. (1994). *Language contact and language death in an immigrant language: The case of Hungarian* (MA thesis). University of Pittsburgh, Pittsburgh, PA.
- Filipović, L. (2007). Language as a witness: Insights from cognitive linguistics. *Speech, Language and the Law*, 14, 245–267.

- Filipović, L. (2008). Typology in action: Applying insights from typological contrasts. *International Journal of Applied Linguistics*, 18, 42–61.
- Filipović, L. (2011). Speaking and remembering in one or two languages: Bilingual vs. monolingual lexicalisation and memory for motion events. *International Journal of Bilingualism*, 15, 466–485.
- Filipović, L. (2013a). Constructing causation in language and memory: Implications for access to justice in multilingual interactions. *International Journal of Speech, Language and the Law*, 20, 1–19.
- Filipović, L. (2013b). The role of language in legal contexts: A forensic cross-linguistic viewpoint. In M. Freeman, & F. Smith (Eds.), *Law and language: Current legal issues 15* (pp. 328–343). Oxford, UK: Oxford University Press.
- Filipović, L. (2014). Efficiency of the bilingual mind: Clues from processing, memory and second language acquisition studies. In L. Filipović, & M. Pütz (Eds.), *Multilingual cognition and language use: Processing and typological perspectives, Human cognitive processing series 44* (pp. 205–227). Amsterdam, Netherlands: John Benjamins.
- Filipović, L. (2018). Speaking in L2 but thinking in L1: Language-specific effects on memory for causation events in English and Spanish. *International Journal of Bilingualism*, 22, 180–198.
- Filipović, L. (under review). Bilingual memory advantage: Bilinguals use linguistic resources from both languages at once as aid to recall memory.
- Filipović, L., & Hawkins, J. A. (2013). Multiple factors in second language acquisition: The CASP model. *Linguistics*, 51, 145–176.
- Filipović, L., & Ibarretxe-Antuñano, I. (2015). Motion. In E. Dąbrowska, & D. Divjak (Eds.), *Handbook of cognitive linguistics* [Handbooks of Linguistics and Communication Science 39] (pp. 527–545). Berlin: Mouton de Gruyter.
- Fishman, J. (1965). Who speaks what language to whom and when? *La Linguistique*, 1(2), 67–88.
- Flecken, M. (2011). Event conceptualisation by early Dutch-German bilinguals: Insights from linguistic and eye-tracking data. *Bilingualism: Language and Cognition*, 14, 61–77.
- Gast, V. (2007). From phylogenetic diversity to structural homogeneity: On right-branching constituent order in Mesoamerica. *Sky Journal of Linguistics*, 20, 171–202.
- Gathercole, V. C. M., & Moawad, R. A. (2010). Semantic interaction in early and late bilinguals: All words are not created equal. *Bilingualism: Language and Cognition*, 13, 385–408.
- Gell-Mann, M. (1992). Complexity and complex adaptive systems. In J. A. Hawkins, & M. Gell-Mann (Eds.), *The evolution of human languages* (pp. 3–18). Redwood City, CA: Addison-Wesley.
- Giles, H., Coupland, J., & Coupland, N. (1991). *Contexts of accommodation: Developments in applied socio-linguistics*. Cambridge, UK: Cambridge University Press.
- Giles, H., & Smith, P. M. (1979). Accommodation theory: Optimal levels of convergence. In H. Giles, & R. N. St. Clair (Eds.), *Language and social psychology* (pp. 45–65). Oxford: Blackwell.
- Green, D. W., & Wei, L. (2014). A control process model of code-switching. *Language, Cognition and Neuroscience*, 29, 499–511.
- Grice, H. P. (1957). Meaning. *The Philosophical Review*, 66, 377–388.
- Grosjean, F. (2001). The bilingual's language modes. In J. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 1–22). Oxford, UK: Blackwell.
- Hartsuiker, R., & Pickering, M. (2008). Language integration in bilingual sentence production. *Acta Psychologica*, 128, 479–489.
- Hartsuiker, R. J., Pickering, M. J., & Veltcamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish/English bilinguals. *Psychological Science*, 15, 409–414.
- Hawkins, J. A. (2004). *Efficiency and complexity in grammars*. Oxford, UK: Oxford University Press.
- Hawkins, J. A. (2009). An efficiency theory of complexity and related phenomena. In D. Gil, G. Sampson, & P. Trudgill (Eds.), *Complexity as an evolving variable* (pp. 252–268). Oxford, UK: Oxford University Press.
- Hawkins, J. A. (2014). *Cross-linguistic variation and efficiency*. Oxford, UK: Oxford University Press.
- Hawkins, J. A., & Filipović, L. (2012). *Criterial features in L2 English: Specifying the reference levels of the Common European Framework*. Cambridge, UK: Cambridge University Press.

- Heine, B., & Kuteva, T. (2005). *Language contact and grammatical change*. Cambridge, UK: Cambridge University Press.
- Hulstijn, J. H. (2012). The construct of language proficiency in the study of bilingualism from a cognitive perspective. *Bilingualism: Language and Cognition, 15*, 422–433.
- Hulstijn, J. H. (2013). Is the second language acquisition discipline disintegrating? *Language Teaching, 46*, 511–517.
- Hulstijn, J. H. (2015). Discussion: How different can perspectives on L2 development be? *Language Learning, 65*, 210–232.
- Jarvis, S., & Pavlenko, A. (2008). *Cross-linguistic influence in language and cognition*. New York, NY: Routledge.
- Kellerman, E. (1986). An eye for an eye: Cross-linguistic constraints on the development of the L2 lexicon. In E. Kellerman, & M. Sharwood Smith (Eds.), *Cross-linguistic influence in second language acquisition* (pp. 35–48). New York, NY: Pergamon Press.
- Kousta, S.-T., Vinson, D. P., & Vigliocco, G. (2008). Investigating linguistic relativity through bilingualism: The case of grammatical gender. *Journal of Experimental Psychology: Learning, Memory and Cognition, 34*, 843–858.
- Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology, 25*, 497–514.
- Kroll, J. F., van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The revised hierarchical model: A critical review and assessment. *Bilingualism: Language and Cognition, 13*, 373–381.
- Kupisch, T. (2014). Adjective placement in simultaneous bilinguals (German-Italian) and the concept of cross-linguistic overcorrection. *Bilingualism: Language and Cognition, 17*, 222–233
- Kupisch, T., Lein, T., Barton, D., Schroder, D. J., Stangen, I., & Stoehr, A. (2013). Acquisition outcomes across domains in adult simultaneous bilinguals with French as weaker and stronger language. *Journal of French Language Studies*. (Advance online publication.) Available from: doi:10.1017/S0959269513000197 (accessed 22 January 2018).
- Lai, V. T., Garrido-Rodriguez, G., & Narasimhan, B. (2014). Thinking-for-speaking in early and late bilinguals. *Bilingualism: Language and Cognition, 17*, 139–152.
- Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics, 18*, 141–165.
- Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levinson, S. (2000). *Presumptive meanings*. Cambridge, MA: MIT Press.
- Long, M. H. (1993). Assessment strategies for second language acquisition theories. *Applied Linguistics, 14*, 225–249.
- Loebell, H., & Bock, K. 2003. Structural priming across languages. *Linguistics, 41*, 791–824.
- MacWinney, B., Malchukov, A., & Moravcsik, E. (Eds.). (2014). *Competing motivations in grammar and usage*. Oxford, UK: Oxford University Press.
- Mellow, J. D. (2008). The emergence of complex syntax: A longitudinal case study of the ESL development of dependency resolution. *Lingua, 118*, 499–521.
- Müller, N., & Hulk, A. (2001). Cross-linguistic influence in bilingual language acquisition: Italian and French as recipient languages. *Bilingualism: Language and Cognition, 4*, 1–53.
- Myers-Scotton, C. (2003). *Multiple voices: An introduction to bilingualism*. Oxford, UK: Blackwell.
- Muysken, P. (2000). *Bilingual speech: A typology of codemixing*. Cambridge, UK: Cambridge University Press.
- Muysken, P. (2013). Language contact outcomes as the result of bilingual optimisation strategies. *Bilingualism: Language and Cognition, 16*, 709–730.
- Nicol, J., Teller, M., & Greth, D. (2001). Production of verb agreement in monolingual, bilingual and second-language speakers. In J. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 117–158). Oxford, UK: Blackwell.
- Nicoladis, E. (2006). Cross-linguistic transfer in adjective-noun strings by preschool bilingual children. *Bilingualism: Language and Cognition, 9*, 15–32.

- O'Grady, W. (2005). *Syntactic carpentry: An emergentist approach to syntax*. Mahwah, NJ: Lawrence Erlbaum.
- O'Grady, W. (2008). The emergentist program. *Lingua*, 118, 447–464.
- Pavlenko, A. (2009). Conceptual representation in the bilingual lexicon and second language vocabulary learning. In A. Pavlenko (Ed.), *The bilingual mental lexicon: Interdisciplinary approaches* (pp. 125–160). Clevedon, UK: Multilingual Matters.
- Pavlenko, A. (2014). *The bilingual mind and what it tells us about language and thought*. Cambridge, UK: Cambridge University Press.
- Pavlenko, A., & Malt, B. (2011). Kitchen Russian: Cross-linguistic differences and first language object-naming by Russian-English bilinguals. *Bilingualism: Language and Cognition*, 14, 19–45.
- Polinsky, M. (1995). Cross-linguistic parallels in language loss. *Southwest Journal of Linguistics*, 14, 87–123.
- Rappaport, G. C. (1990). The language situation of Americans of Polish descent in Texas [*Sytuacja językowa Amerykanów Polskiego pochodzenia w Teksasie*]. In W. Modunka (Ed.), *Polish language in the world [Język Polski w świecie]* (pp. 159–178). Krakow, Poland: PAN.
- Romaine, S. (1989). *Bilingualism*. Oxford, UK: Blackwell.
- Ross, M. D. (1996). Contact-induced change and the comparative method: Cases from Papua New Guinea. In M. Durie, & M. D. Ross (Eds.), *The comparative method reviewed: Regularity and irregularity in language change* (pp. 180–217). Oxford, UK: Oxford University Press.
- Ross, M. D. (2001). Contact-induced change in oceanic languages in North-West Melanesia. In A. Y. Aikhenvald, & R. M. Dixon (Eds.), *Areal diffusion and genetic inheritance: Problems in comparative linguistics* (pp. 134–166). Oxford, UK: Oxford University Press.
- Savić, J. (1995). Structural convergence and language change: Evidence from Serbian/English code-switching. *Language in Society*, 24, 475–492.
- Schmitt, E. (2000). Overt and covert codeswitching in immigrant children from Russia. *International Journal of Bilingualism*, 4, 9–28.
- Silva-Corvalán, C. (1994). *Language contact and change: Spanish in Los Angeles*. Oxford, UK: Clarendon Press.
- Slobin, D. I. (1996). Two ways to travel: Verbs of motion in English and Spanish. In M. Shibatani, & S. A. Thompson (Eds.), *Grammatical constructions – Their form and meaning* (pp. 195–219). Oxford, UK: Clarendon Press.
- Slobin, D. I. (1997). Mind, code, and text. In J. Bybee, J. Haiman, & S. Thompson (Eds.), *Essays on language function and language type* (pp. 437–467). Amsterdam, Netherlands: John Benjamins.
- Slobin, D. I. (2003). Language and thought online: Cognitive consequences of linguistic relativity. In D. Gentner, & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 157–191). Cambridge, MA: MIT Press.
- Slobin, D. I. (2016). Thinking for speaking and the construction of evidentiality in language contact. In M. Güven, D. Akar, B. Öztürk, & M. Kelepir (Eds.), *Exploring the Turkish linguistic landscape: Essays in honor of Eser Erguvanlı-Taylan, Studies in language companion series 175* (pp. 105–120). Amsterdam, Netherlands: John Benjamins.
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Oxford, UK: Blackwell.
- Talmy, L. (1985). Lexicalisation patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description, Vol. 3: Grammatical categories and the lexicon* (pp. 57–149). Cambridge, UK: Cambridge University Press.
- Tooley, K. M., Traxler, M. J., & Swaab, T. Y. (2009). Electrophysiological and behavioural evidence of syntactic priming in sentence comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 19–45.
- Torbio, A. J. (2004). Convergence as an optimisation strategy in bilingual speech: Evidence from code-switching. *Bilingualism: Language and Cognition*, 7, 165–173.
- Traxler, M. J. (2008). Structural priming among prepositional phrases: Evidence from eye-movement. *Memory and Cognition*, 36, 659–674.

- Traxler, M. J. (2012). *Introduction to psycholinguistics: Understanding language science*. Chichester: Wiley-Blackwell.
- Treffers-Daller, J. (2015). The construct of language dominance, its operationalisation and measurement. In C. Silva-Corvalán, & J. Treffers-Daller (Eds.), *Language dominance in bilinguals: Issues of measurement and operationalisation* (pp. 235–265). Cambridge, UK: Cambridge University Press.
- Trudgill, P. (2011). *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford, UK: Oxford University Press.
- van Hell, J. G., & de Groot, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition, 1*, 193–211.
- van Hell, J. G., & Kroll, J. F. (2013). Using electrophysiological measures to track the mapping of words to concepts in the bilingual brain: A focus on translation. In J. Altarriba, & L. Isurin (Eds.), *Memory, Language and Bilingualism: Theoretical and Applied Approaches* (pp. 126–160). Cambridge, New York: Cambridge University Press.
- Weber, K., & Indefrey, P. (2009). Syntactic priming in German-English bilinguals during sentence comprehension. *Neuroimage, 46*, 1164–1172.
- Yip, V., & Matthews, S. (2007). *The bilingual child: Early development and language contact*. Cambridge, UK: Cambridge University Press.
- Zuengler, J. (1991). Accommodating in native-nonnative interactions: Going beyond the ‘what’ to the ‘why’ in second-language research. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of accommodation* (pp. 223–244). New York, NY: Cambridge University Press.

Author biographies

Luna Filipović is a professor of Language and Cognition at the University of East Anglia. She received her PhD in Linguistics from the University of Cambridge and held an ESRC Postdoctoral Research Fellowship in Psychology at University College London (2005–2008) and a Leverhulme Trust & Newton Trust Postdoctoral Research Fellowship in Linguistics at the University of Cambridge (2008–2011). Her research programme is focused on language-specific effects on memory and bilingualism, and includes interdisciplinary work across the areas of psycholinguistics, forensic linguistics, language typology and the interaction between language and cognition. She has authored, co-authored or co-edited six books and published numerous articles and invited chapters.

John A Hawkins is a distinguished professor of Linguistics at the University of California Davis, and an emeritus professor of English and Applied Linguistics at Cambridge University. He has held previous positions at the University of Southern California, the Max-Planck-Institute for Psycholinguistics in Nijmegen, and the University of Essex, and visiting appointments at the University of California Berkeley, University of California Los Angeles, the Free University of Berlin, University of Potsdam, University of the Saarland, and the Max-Planck-Institute for Evolutionary Anthropology in Leipzig. He has broad interests in the language sciences and has published widely on language typology and universals, efficiency and complexity in language, psycholinguistics (language processing and language acquisition), semantics and pragmatics, the Germanic language family, and language change.